

**The effect of a training and clinical facilitation
programme for registered midwives in primary
maternity settings with respect to managing labour :
a pragmatic cluster randomised trial**

Sheila Elizabeth Clow

Thesis presented for the degree of
DOCTOR OF PHILOSOPHY (MATERNAL and CHILD HEALTH)
in the Department of Paediatric Medicine
UNIVERSITY OF CAPE TOWN

April 2015

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Abstract

Author : Sheila Elizabeth Clow

Thesis title : The effect of a training and clinical facilitation programme for registered midwives in primary maternity settings with respect to managing labour : a pragmatic cluster randomised trial

Date : April 2015

Intrapartum complications contribute to nearly half of all avoidable maternal and perinatal deaths nationally. Inadequate understanding of the labour process by midwives, poor documentation of labour monitoring and inadequate system-level support may lead to wrong diagnosis, incorrect management, and the potential for missed opportunities to prevent mortality.

Aim

To evaluate the effect of an intervention package of training and clinical facilitation on the quality of clinical management in labour by registered midwives in primary level public sector health facilities in rural South Africa.

Methods

Research design : Pragmatic cluster randomised trial with 12 month follow-up.

Setting and participants : Seventeen clusters stratified by geo-political region and size of service; 1020 labour records (60 per cluster / site; systematic random sample); and 154 registered midwives employed in the study sites during the study period. Participants were not blinded.

Intervention :

A package of clinical facilitation training for selected experienced midwife clinicians / managers, and an intrapartum educational update for midwives. Intervention and control sites continued receiving routine communication, all clinical guidelines and scheduled outreach activities.

Main outcome measures :

Primary outcome - clinical practice measuring partograph utilisation, using a modified partograph checklist, the testing of which is described in this study.

Secondary outcome - midwives' knowledge and skills, measured by written and clinical tests.

Outcomes were analysed at the individual level using regression methods that allowed for clustering.

The evaluator was blinded to the study allocation.

Findings

The mean scores for the total partograph were not statistically significantly different between arms; the mean difference was 1.55 points out of a possible score 47 (95% CI: -1.18 to 4.28) $p = 0.27$. At a score of 27 the estimated absolute difference was 13.6% (95% CI : 0.16 to 0.25) $p = 0.026$.

The total score for midwives' knowledge and skills was 7 points (out of a possible 119) higher in the intervention arm (95% CI : 2.1 to 12.3), $p = 0.006$.

Conclusions

Although there was no difference in the quality of the overall completion of the partograph, there was a statistically significant difference in those of better quality completions in the intervention arm. Midwives' knowledge and skills were higher in the intervention arm and those in the control arm deteriorated over time. This difference was statistically significant.

Recommendations and implications for practice

This indicates a critical need to provide continuing professional education to midwives and to arrange midwifery staffing that optimises clinical practice in settings where intrapartum care is offered.

In addition to regular, sustainable programmes to enhance partograph utilisation and midwife knowledge and skills, barriers to the utilisation of the partograph need to be investigated and addressed.

Declaration

I, Sheila Elizabeth Clow, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: signature removed

Date: 30 April 2015

*To my parents,
Gordon and Jean,
for their unfailing love, support and encouragement
in all my endeavours*

*and to my Heavenly Father
in whom I live and have my being.
To Him be the glory*

Acknowledgements

I acknowledge with grateful thanks the support and contributions of the following people and organisations, who all contributed to the successful completion of this project.

Funding

- ◆ South African Medical Research Council
- ◆ The National Research Foundation
- ◆ The University of Cape Town University Research Fund
- ◆ Carnegie Fund

Research supervision

- ◆ Prof George Swingler, UCT Department of Paediatric Medicine
- ◆ Prof Bob Pattinson, MRC Unit for Maternal and Infant Care Research

Statistical consultancy

- ◆ Dr Carl Lombard, Medical Research Council

Research participants

- ◆ Nurse managers, mentors and registered midwife staff at all the participating sites
- ◆ Emeritus Prof Herman A van Coeverden de Groot and Prof Sue Fawcus
- ◆ Judges : Dr Stefan Gebhardt, Ms Liesbeth Mangate, Ms Mickey Masasa, Ms Nokuzola Mzolo, Dr Jenny Nash, Ms Dolly Nyasulu, Prof Bob Pattinson, Prof Hugh Philpott, Dr Hannes Steinberg, Dr Anna Voce
- ◆ Midwife Evaluators : Ms Carol Adams and Ms Lenie Krynaauw

Support from Provincial Government of the Western Cape (PGWC) health service personnel for access to sites, fieldwork support and access to data

- ◆ Mrs Edna Arends, Deputy Director : Women's Health
- ◆ Stefan Gebhardt, provincial co-ordinating clinician for Obstetrics and Gynaecology, and regional obstetrician West Coast / Winelands region
- ◆ Dr Frans Krige, Director : Boland / Overberg region
- ◆ Ms Carine Bester, Director : West Coast / Winelands region
- ◆ Mrs Surina Neethling, programme manager, Maternal, Child and Women's Health, Boland / Overberg region
- ◆ Dr Charl Oettle, regional obstetrician Boland / Overberg region, and Perinatal Problem Identification Programme (PPIP) co-ordinator for the region
- ◆ Mrs Trudy Fredericks, programme manager, Maternal, Child and Women's Health, West Coast / Winelands region
- ◆ Administrative staff and records clerks in the participating sites
- ◆ Dr Dave Greenfield, Perinatal Problem Identification Programme (PPIP) co-ordinator, metro region
- ◆ Members of the PGWC Maternity Guidelines group
- ◆ Dr Joey Cupido, Deputy Director General : District Health Service and Health Programmes; Mr Steven Titus, Director : Facility based programmes; Dr Tracy Naledi, Director : Health Impact assessment; and Dr Indira Govender
- ◆ Dr Erasmus (Helderberg Hospital) and Dr B Kruger (Paarl Hospital)

For graphical design of certain figures

- ◆ Mr Michael Wyeth, Imago-Visual

Research assistance and administrative support

- ◆ Ms Cathy de Groot, research assistance, co-ordination, document retrieval, document management and coding, and technical editing
- ◆ Ms Sarah Seggie and Mrs Margaret Clarke, for typing
- ◆ Ms Zandile Gumbo, who did a preliminary literature search for the design of the partograph utilisation instrument before her untimely death
- ◆ Ms Abeeda Mollagee for assistance with translation of materials
- ◆ Dr Jenny Luxton, Ms Evelyn Els and Ms Elizabeth Michael, for editing assistance

University of Cape Town

- ◆ Staff of the Division of Nursing and Midwifery, particularly for facilitating my taking study and research leave
- ◆ Dr Juliet Wood, Mrs Ingrid Coldicott and Mr Jason Marcus – Midwifery colleagues who provided invaluable support in terms of teaching relief while I was writing up the dissertation
- ◆ For the privilege of study and research leave which enabled me to conduct my fieldwork
- ◆ For a PhD completion grant to enable me to have six months relief from core teaching functions to make progress towards finalising the dissertation. Funding was supplied by the Carnegie Fund, and is gratefully acknowledged
- ◆ The Research and Innovation office, particularly Dr Lyn Holness, Dr Mignonne Breier, Dr Dianne Bond and the emerging researcher programme, for the ongoing encouragement, support and assistance. Also Dr Nelleke Bak, Director of Postgraduate Studies for guidance and support.
- ◆ Ms Haajirah Esau for assistance with funding applications
- ◆ Librarians, particularly in the Health Science library, whose helpfulness and efficiency is unmatched

My family and friends for their love, patience, support and encouragement.

Table of Contents

	Page
Abstract	i
Declaration	iii
Dedication	iv
Acknowledgements	vi
Table of contents, including appendices and lists of figures and tables	ix
Acronyms, abbreviations and definitions	xxvi
Chapter 1 Introduction	
1.1 Overview	1
1.2 Background to study	2
1.3 Maternal and neonatal indicators	3
1.3.1 International indicators	3
1.3.2 National indicators	5
1.3.2.1 Maternal mortality ratio	5
1.3.2.2 Perinatal and neonatal mortality	7
1.3.2.3 Strategies to address maternal and perinatal mortality in South Africa	9
1.3.2.4 The impact of HIV on maternal and perinatal mortality in South Africa	9
1.4 Study setting	10
1.5 Maternity care in the public sector health service	13
1.6 Nursing and Midwifery	14
1.6.1 Midwifery and the scope of practice in respect of intrapartum care	15
1.6.2 Midwifery knowledge and skills	19
1.7 Management of labour	20
1.8 Defining the problem	21
1.9 Conclusion	23

Chapter 2 Literature Review

2.1	Introduction	24
2.2	Literature search strategies and results.....	24
2.2.1	Changing practise.....	24
2.2.1.1	Common approaches	24
2.2.1.2	Knowledge translation	26
2.2.1.3	Practice Improvement interventions.....	26
2.2.1.4	Implementation methods	27
2.2.2	Mentorship.....	27
2.2.2.1	Search strategy	30
2.2.2.2	Search results	32
2.2.3	Instruments to evaluate partograph utilisation	32
2.2.3.1	Search strategy	32
2.2.3.2	Search results	33
2.3	Describing the field of improving the quality of clinical practice.....	34
2.3.1	Knowledge translation.....	35
2.3.2	Practice Improvement	38
2.4	Presentation of literature	40
2.4.1	Context and understanding of the problem / issue	41
2.4.2	Participation and Participatory process	41
2.4.3	Behaviour change.....	42
2.4.4	Implementation strategies	44
2.4.4.1	Educational outreach.....	45
2.4.4.2	Audit and feedback.....	47
2.4.4.3	Local opinion leaders.....	49
2.4.4.4	Distribution of educational materials	50
2.4.4.5	Continuing professional education.....	52
2.4.4.5.1	Continuing education meetings and workshops.....	52
2.4.4.5.2	Perinatal Education programme (PEP).....	53
2.4.4.5.3	Training-of-trainers and Face-to-face training approaches.....	55
2.4.4.5.4	Critical appraisal.....	58
2.4.4.6	Package of interventions	58
2.4.5	Standardisation.....	61
2.4.6	Evaluation.....	62
2.4.7	Health system issues	62
2.4.7.1	System change / support.....	62
2.4.7.2	Nature of the professional practice context.....	63
2.4.7.3	Research utilisation by nurses	65
2.4.7.4	Barriers to change	67
2.4.7.5	Organisational culture.....	69
2.4.7.6	Integrated care pathways (ICP)	71
2.4.7.7	Health system level interventions	72

2.4.8	Mentorship.....	74
2.4.8.1	Defining the field.....	74
2.4.8.1.1	Targets of mentoring	74
2.4.8.1.2	Nature, dynamics and characteristics of the mentorship relationship	75
2.4.8.1.3	Setting	75
2.4.8.1.4	Preparation for mentorship	75
2.4.8.2	Effectiveness in terms of patient outcomes.....	76
2.4.8.3	Limitations of mentorship literature	79
2.4.9	Instruments for evaluating partograph utilisation	80
2.4.9.1	WHO Safe Motherhood questionnaire for assessment of partogram use	81
2.4.9.2	Partograph checklist with five criteria.....	82
2.4.9.3	Partograph utilisation in Indonesia.....	82
2.4.9.4	Comparison of partograph designs in India	82
2.4.9.5	Partograph utilisation in Kenya	83
2.4.9.6	Partograph checklist in Iran	83
2.4.9.7	Other sources identified.....	83
2.4.9.8	Scoring considerations	85
2.4.9.8.1	Completeness	86
2.4.9.8.2	Correctness	86
2.4.9.9	Length of a partograph utilisation checklist	87
2.4.10	Conclusions	88
2.4.10.1	Practice change	88
2.4.10.2	Mentorship	89
2.4.10.3	Instrumentation to evaluate partograph utilisation	89
2.5	What gaps exist	90
2.5.1	Practice change	90
2.5.2	Mentorship.....	92
2.5.3	Instrumentation to evaluate partograph utilisation	92
2.6	Implications for this study	93
2.6.1	Aim	94
2.6.2	Objectives.....	95
2.6.3	Hypotheses	96

Chapter 3 Methods

3.1	Introduction	97
3.2	Considerations for appropriate study design	97
3.2.1	Research method	97
3.2.1.1	Randomised controlled trial (RCT).....	97
3.2.1.2	Cluster randomised trial (CRT)	101
3.2.1.2.1	Rationale for using a CRT design	101
3.2.1.2.1.1	Setting	101
3.2.1.2.1.2	Type of intervention	102
3.2.2	Pragmatic approach.....	102
3.3	CRT design considerations	102
3.3.1	Characteristics of study design	103
3.3.2	Design strategies	103
3.3.2.1	Matched pairs.....	103
3.3.2.2	Restriction	104
3.3.2.3	Stratification.....	104
3.3.3	Sample considerations.....	105
3.3.3.1	Design effect	105
3.3.3.2	Relationships within clusters.....	106
3.3.3.3	Variation between clusters.....	107
3.3.3.4	Number and size of clusters	107
3.3.4	Randomisation.....	108
3.3.5	Analysis	108
3.3.5.1	Approach to analysis	108
3.3.5.2	Statistical modelling.....	109
3.3.5.3	Intention-to-treat versus Per-protocol analysis.....	109
3.3.5.4	Addressing bias	110
3.3.6	Reporting requirements	111
3.4	Considerations for the design of the intervention.....	111
3.4.1	Context and understanding the problem	112
3.4.2	Participatory process	112
3.4.3	Behaviour change.....	113
3.4.4	Implementation strategy.....	113
3.5	Considerations regarding choice of indicators	114
3.5.1	End points.....	114
3.5.2	Proxy indicators	115
3.5.2.1	Intrapartum care	115
3.5.2.1.1	Neonatal indicators	115
3.5.2.1.2	Format and utilisation of the partograph.....	116
3.5.2.1.3	Suitability as a proxy indicator.....	117
3.5.2.1.4	Some debates about the use of the partograph	118
3.5.2.2	Midwifery knowledge and skills.....	119

3.6	Instrumentation – partograph utilisation checklist	120
3.6.1	Developing and validating the instrument.....	120
3.6.2	Testing of the instrument in this population	123
3.6.2.1	Population, sample size and sampling.....	123
3.6.2.1.1	Population.....	123
3.6.2.1.2	Sample size determination.....	123
3.6.2.1.3	Sampling.....	123
3.6.2.2	Evaluation by two midwives (inter-reader reliability).....	123
3.6.3	Analysis to determine concordance	124
3.6.3.1	Lin's Concordance Correlation Coefficient.....	124
3.6.3.2	Bland and Altman's limits of agreement.....	124
3.6.3.3	Correlation between difference and mean	124
3.6.3.4	F-test of equality of mean and variants	124
3.6.3.5	Results and interpretation (inter-reader reliability)	124
3.6.4	Evaluation by multiple midwives (inter-reader reliability)	125
3.6.4.1	Estimating difference and evaluating agreement	125
3.6.4.1.1	Estimating difference	125
3.6.4.1.2	Evaluating agreement.....	125
3.6.4.1.3	Results and interpretation	126
3.6.5	Conclusion.....	126
3.7	Research methods used for this study	126
3.7.1	Design rationale and description	127
3.7.2	Setting, sites and population	127
3.7.3	Sample size, sampling and randomisation	129
3.7.3.1	Clinical practice (Partograph utilisation audit)	130
3.7.3.2	Registered midwives' knowledge and skills	131
3.7.3.3	Sites / clusters	133
3.7.3.4	Comparator	134
3.7.4	Intervention package.....	134
3.7.4.1	Clinical facilitator development programme	136
3.7.4.1.1	Objectives.....	137
3.7.4.1.2	Target group and selection	137
3.7.4.1.2.1	Preparation of clinical facilitators prior to development programme.....	137
3.7.4.1.3	Time period.....	138
3.7.4.1.4	Programme outline	138
3.7.4.2	Registered midwife educational update	139
3.7.4.2.1	Objectives.....	139
3.7.4.2.2	Time period.....	139
3.7.4.2.3	Programme outline	140
3.7.4.3	Timing of intervention and data collection.....	140
3.7.5	Outcomes	142

3.7.6	Instruments and data collection.....	142
3.7.6.1	Clinical practice (Partograph utilisation audit)	142
3.7.6.1.1	Validity	143
3.7.6.1.2	Reliability	143
3.7.6.1.3	Pilot study	143
3.7.6.1.4	Data collection	144
3.7.6.2	Midwives' knowledge and skills	144
3.7.6.2.1	Validity	145
3.7.6.2.2	Reliability	145
3.7.6.2.3	Pilot study	145
3.7.6.2.4	Data collection	146
3.7.7	Data management	147
3.7.8	Data analysis	148
3.7.8.1	Clinical practice – primary outcome	148
3.7.8.2	Midwife tests – secondary outcome.....	149
3.7.9	Managing potential bias	151
3.7.10	Reporting the study.....	153
3.8	Ethical considerations	153
3.8.1	Scientific soundness of a study	154
3.8.2	Investigator preparedness.....	154
3.8.3	Ethics approval	155
3.8.4	Permission to conduct the study	155
3.8.5	Informed consent	155
3.8.6	Distributive justice	156
3.8.7	Privacy, confidentiality and data protection	157
3.8.8	Risk benefit analysis	157
3.8.8.1	Benefits	157
3.8.8.2	Potential risk.....	158
3.8.8.3	Risk benefit analysis conclusion	159
3.8.9	Dissemination of results	159
3.9	Conclusion	160

Chapter 4 Results

4.1	Introduction	161
4.2	Study profiles	161
4.2.1	Profile of sites	161
4.2.2	Profile of partograph records.....	163
4.2.3	Profile of Registered midwives	164
4.2.3.1	Professional education and experience	164
4.2.3.2	Knowledge and skills	165
4.3	Clinical practice (partograph utilisation audit) (Objective one)	168
4.3.1	Distribution of partograph records	169
4.3.2	Primary outcome : Clinical practice (partograph utilisation)	170
4.3.2.1	<i>Post hoc</i> subgroup analysis.....	176
4.3.2.1.1	The effect of size on the service	176
4.3.2.1.2	The effect of time	177
4.3.2.1.3	The effect of region.....	178
4.3.2.1.4	The intra-cluster correlation coefficient (ICC)	179
4.3.2.2	Conclusion for partograph utilisation score	180
4.3.3	Sector scores.....	181
4.3.3.1	Combined observations.....	182
4.3.3.2	Adherence to clinical management guidelines	184
4.3.3.3	Conclusions regarding sector scores	185
4.3.4	Scores related to specific innovations	185
4.3.4.1	Fetal heart observations	186
4.3.4.2	Critical comprehensive assessment and management plan ...	186
4.3.5	Contextual reasons for not completing the partograph	189
4.3.5.1	Reasons recorded on the partograph	190
4.3.5.2	Time of birth in relation to the last observation.....	190
4.3.5.3	Time of birth in relation to the commencement of the partograph.....	191
4.3.5.4	Conclusion regarding contextual factors	192
4.3.6	Partograph utilisation audit conclusion.....	192
4.4	Midwives' knowledge and skills (Objective two).....	192
4.4.1	Study population dynamics	193
4.4.2	Secondary outcome : Comparison of midwives' final tests between arms	195
4.4.2.1	Total scores.....	197
4.4.2.1.1	Test three complete – Intention-to-treat	197
4.4.2.1.2	Test three complete where test one was also Complete - Per-protocol analysis.....	199
4.4.2.1.3	Test three Component completions	201
4.4.2.1.4	Conclusion regarding midwives' total scores	202
4.4.2.2	Sub-scores of midwives' tests	202
4.4.2.2.1	Types of questions	203

4.4.2.2.1.1	Test three complete – ‘Intention-to-treat’	203
4.4.2.2.1.2	Test three complete where test one was also complete – ‘Per protocol analysis’	205
4.4.2.2.1.3	Theory or clinical component completers.....	206
4.4.2.2.1.4	Conclusions regarding types of questions.....	208
4.4.2.2.2	Cognitive levels	208
4.4.2.2.2.1	‘Intention-to-treat’ analysis group.....	209
4.4.2.2.2.2	‘Per-protocol’ analysis group	210
4.4.2.2.2.3	Conclusion for cognitive level results...	211
4.4.2.2.3	Practice aspects	211
4.4.2.2.3.1	Conclusion for practice aspects.....	213
4.4.2.3	Determining sensitivity of results	213
4.4.2.3.1	Theory scores.....	214
4.4.2.3.2	Clinical scores	215
4.4.2.4	Midwives’ tests conclusion.....	216
4.4.3	Changes within intervention arm over time (Objective three).	216
4.4.3.1	Total score.....	217
4.4.3.2	Sub-scores	218
4.5	Comparison between structured versus unplanned <i>ad hoc</i> approach to staff training (Objective four)	221
4.6	Conclusion	222

Chapter 5 Discussion

5.1	Introduction	223
5.2	Partograph utilisation audit.....	223
5.2.1	Key findings	223
5.2.1.1	Hypothesis testing	224
5.2.1.2	Bias	224
5.2.1.3	Adverse events.....	224
5.2.1.4	Dangers associated with multiplicity of analysis and outcomes	224
5.2.2	Possible mechanisms and explanations for these results	225
5.2.2.1	Non-utilisation of partographs.....	225
5.2.2.2	Non-completion of partographs.....	226
5.2.2.2.1	Staffing levels for intrapartum care	226
5.2.2.2.2	Inappropriate deployment of staff	227
5.2.2.2.3	Failure to value the importance of observations and documenting thereof	227
5.2.2.2.4	Birth or transfer.....	228
5.2.2.2.5	Lack of second stage observations recorded.....	228
5.2.2.3	Low scores	228
5.2.2.3.1	Clinical skills deficit.....	228
5.2.2.3.2	Lack of understanding of labour / clinical relevance	229
5.2.2.3.3	Compliance with clinical guidelines.....	229
5.2.2.4	Specific innovations.....	230
5.2.2.5	Institutional profiles.....	231
5.2.2.6	Clinical facilitation	232
5.2.3	Comparison with relevant findings from current available literature....	232
5.2.3.1	Utilisation of partograph.....	232
5.2.3.2	Policy in to practice needs to be actively mediated	234
5.2.3.3	Use of multifaceted intervention.....	235
5.2.3.4	Under-resourced settings	235
5.2.3.5	ICC values.....	236
5.2.3.6	Professional / workplace culture	236
5.2.3.6.1	Professional differences	236
5.2.3.6.2	Staff utilisation and support	237
5.2.3.6.3	Practice settings	238
5.3	Midwife tests	239
5.3.1	Key findings	239
5.3.1.1	Hypothesis testing	240
5.3.1.2	Bias	240
5.3.1.3	Adverse events.....	242
5.3.1.4	Dangers associated with multiplicity of analysis and outcomes.....	242

5.3.2	Possible mechanisms and explanations.....	242
5.3.2.1	Pattern of change in scores	243
5.3.2.2	Low scores	243
5.3.2.2.1	Cardiotocograph skills	244
5.3.2.2.2	Cognitive levels	244
5.3.2.2.3	Alignment with clinical guidelines.....	244
5.3.2.3	Tracking test performance of a subgroup of midwives at different timepoints	245
5.3.2.4	Human resource management	245
5.3.3	Comparison with relevant findings from current available evidence ..	246
5.3.3.1	Location of intervention.....	246
5.3.3.2	Role of expert trainer	247
5.3.3.3	Accessing and applying knowledge	247
5.3.3.4	The extent of change	248
5.3.3.5	Number of interactions.....	249
5.3.3.6	Barriers and supports for professional change.....	249
5.3.3.7	ICC values.....	251
5.4	Strengths and weaknesses (including limitations)	251
5.4.1	Strengths	251
5.4.1.1	Application of this research design to this issue.....	251
5.4.1.2	Application of this research design and question to this professional group	252
5.4.1.3	Application of this research design in this setting.....	252
5.4.1.4	Pragmatic approach	252
5.4.1.5	Application of this reseach design to this educational intervention.....	253
5.4.1.6	Design of intervention.....	253
5.4.1.7	Partograph utilisation instrument	255
5.4.2	Weaknesses / limitations.....	255
5.4.2.1	Design issues	255
5.4.2.1.1	Sample size calculations	255
5.4.2.1.2	Cluster versus individual level analysis	255
5.4.2.1.3	Statistical power	256
5.4.2.1.4	Blinding	256
5.4.2.2	Intervention issues.....	256
5.4.2.2.1	Number of clinical facilitators	256
5.4.2.2.2	Attributes of clinical facilitators	256
5.4.2.2.3	Role of regional MCWH programme coordinator ...	257
5.4.2.3	Instrument issues	257
5.4.2.4	Testing and analysis issues.....	257
5.4.2.4.1	Partograph utilisation audit	257
5.4.2.4.1.1	Proxy indicator	257
5.4.2.4.1.2	Lack of baseline scores	258
5.4.2.4.1.3	Non-utilisation of the partograph.....	258

5.4.2.4.2	Midwife tests.....	258
5.4.2.4.2.1	Different participants at different timepoints.....	258
5.4.2.4.2.2	Within-arm comparisons.....	259
5.4.2.4.2.3	Intervention arm analysis – missing timepoint.....	259
5.4.2.5	Health service issues.....	260
5.4.2.6	Method issues	260
5.5	Generalisability and external validity	261
5.5.1	Key aspects of the setting which determined the trial results	263
5.6	Reflexivity and pragmatic research.....	264
5.6.1	The role and purpose of reflexivity and its application to a pragmatic trial	265
5.6.2	The impact of the researcher and her presence on the research process.....	266
5.6.2.1	Identity and power	266
5.6.2.2	Research design and power	267
5.6.2.3	Power and interaction	269
5.6.2.3.1	Positive professional engagement	269
5.6.2.3.2	Changing the nature of the interaction	270
5.6.2.4	Integrity – minimising bias	272
5.6.3	The impact of the research process on the researcher	272
5.6.4	Reflection and conclusion	273
5.7	Conclusion	274

Chapter 6 Recommendations and conclusion

6.1	Introduction	276
6.2	Summary and reflection of the study	276
6.3	Contribution of this study to the field.....	279
6.4	Research and Clinical implications	280
6.4.1	Research implications	280
6.4.1.1	Establishment of ICCs for this population	280
6.4.1.2	Availability of a partograph evaluation instrument.....	280
6.4.2	Clinical implications.....	280
6.4.2.1	Lack of recording of observations	280
6.4.2.2	Low scores	281
6.4.2.3	Lack of clinical judgement.....	281
6.4.2.4	Second stage observations	281
6.5	Recommendations	282
6.5.1	Research	282
6.5.1.1	Instrument design.....	282
6.5.1.2	Design of educational outreach	282
6.5.1.3	Barriers to utilisation of the partograph	282
6.5.1.4	Barriers and supports for professional practice change	282
6.5.1.5	Research amongst non-physicians and in developing countries.....	283
6.5.1.6	Conceptual clarity in the field of mentorship.....	283
6.5.1.7	Effectiveness of mentorship	284
6.5.1.8	Costing analysis	284
6.5.2	Strengthening Midwifery capacity.....	284
6.5.2.1	Continuing Professional Development	284
6.5.2.2	Clinical facilitation and clinical leadership	285
6.5.2.3	Access to appropriate evidence.....	285
6.5.3	Health system strengthening.....	286
6.5.3.1	Clinical practice	286
6.5.3.2	Clinical records.....	286
6.5.3.3	Clinical audit	286
6.5.3.4	Institutional culture	287
6.5.3.5	Human resource management	287
6.6	Dissemination of findings.....	287
6.7	Conclusion	288
	References.....	289

Appendices

1.	Map of area (1.4 and 3.7.2)	324
2.	Example of Western Cape partograph (1.6.2)	325
3.	a) Summary of Systematic reviews (2.4)	327
	b) Effect of different implementation strategies (2.4.4)	334
4.	Validation documentation (3.6.1).....	346
5.	Partograph utilisation checklist (3.6.1 and 3.7.6.1)	364
6.	Information sheet and consent form for mentors (3.7.4.1.2.1 and 3.8.5)	369
7.	Labour record review (Philpott and Voce) (3.7.4.1.4)	371
8.	Intervention package programme (3.7.4.1. and 3.7.4.2).....	373
	a) Clinical facilitator programme (3.7.4.1.4 and 5.5.1).....	373
	b) Midwife training programme – session overview (3.7.4.2.3)	378
9.	Midwife knowledge and skills evaluation (3.7.6.2).....	379
10.	Midwife tests (3.7.6.2.1)	381
	a) Test 1	381
	b) Test 2	385
	c) Test 3	389
11.	Midwives' tests data collection tools (3.7.6.2.4)	393
	a) Test 1	393
	b) Test 2	395
	c) Test 3	397
12.	Alignment of reporting with CONSORT statement requirements (3.7.10)	399
13.	Ethics approval, University of Cape Town (3.8.3)	407
14.	Permission from Provincial Government (3.8.4)	409
15.	Information sheet and consent form for midwives (3.8.5)	413
	a) Midwives in intervention sites.....	413
	b) Midwives in control sites	415
	c) Midwives – all sites after month three	416
16.	Training evaluation by clinical facilitators (5.4.1.6)	419

List of Figures

1.1	Maternal mortality ratio (MMR) estimates for South Africa, with 2015 target	6
2.1	Literature search results	28
2.2	Knowledge to action framework	37
2.3	The Model for Improvement.....	39
2.4	The ingredients of readiness to change	43
3.1	Algorithm to determine the appropriate components of the study design and analysis	98
4.2.1	Scatter plot illustrating the years of experience (practice) in relation to the years qualified as a registered midwife.....	165
4.3.1	Flow chart indicating the sites/clusters and partograph records enrolled in the study	169
4.3.2	Map of presentation of results and findings - partograph utilisation audit	171
4.3.3	Boxplot of total partograph scores	172
4.3.4	Distribution of the partograph score, reflected for both arms	172
4.3.5	Distribution of the partograph scores with estimated data-dependent distribution.....	175
4.3.6	Box plot of partograph scores per arm and per size of service.....	177
4.3.7	Comparison of scores between arms over time	178
4.3.8	Box plot of partograph scores per arm and per region	179
4.3.9	Boxplot of partograph scores for all 17 sites	180
4.3.10	Distribution of combined observations scores	183
4.3.11	Boxplot of 'alignment to clinical management guidelines' score ...	184
4.3.12	Distribution of scores for the comprehensive assessment	188
4.4.1	Participant (registered midwife) level flow diagram	194
4.4.2	Map of presentation of results and main findings - midwife tests ..	196
4.4.3	Comparison of test three total scores in intervention and control arm (intention-to-treat)	198
4.4.4	Comparison of test three total scores in intervention and control arm (per-protocol)	200
4.4.5	Scatterplot of test three versus test one total scores for the two arms	201
4.4.6	Changes in total percentage score, and theory and clinical components, at three timepoints	218
4.4.7	Changes in percentage scores for question types	219
4.4.8	Changes in percentage scores for cognitive levels of questions ..	219
4.4.9	Changes in percentage scores for practice aspects	220

List of Tables

3.1	Differences between three readers	126
3.2	Stratification framework	133
3.3	Timeline for intervention and data collection	141
3.4	Variables for the partograph utilisation audit	142
4.2.1	Description of the sites (clusters) and their study allocation	162
4.2.2	Randomisation of the registered midwives in relation to the sites	162
4.2.3	Profile of the records of women and their babies included in the study	163
4.2.4	Profile of the registered midwives in each arm in terms of educational preparation and experience	164
4.2.5a	Pattern of completions for Test one - theory component	166
4.2.5b	Pattern of completions for Test one - clinical component	166
4.2.6	Descriptive statistics for baseline theory and clinical scores for all midwives enrolled at baseline	167
4.2.7	Descriptive statistics for baseline theory and clinical scores where the tests were complete ('intention to treat')	167
4.2.8	Descriptive statistics for baseline theory and clinical scores ('per protocol group')	168
4.3.1	Comparison of the utilisation of the partograph between the two arms	170
4.3.2	Descriptive statistics for the partograph score	170
4.3.3	Comparison of mean partograph scores between the two arms	173
4.3.4	Simultaneous quantile regression at deciles 10 – 90 for the total partograph score	174
4.3.5	The risk difference at a score of 27 on the partograph	174
4.3.6a	Comparison of the partograph scores for the total trial population at the cut point score of 27	174
4.3.6b	Comparison of the partograph scores for the total trial population	175
4.3.7	Linear mixed effects model of partograph score (adjusted for 17 clusters)	176
4.3.8	Partograph scores per arm and per size of service	177
4.3.9	Descriptive statistics for scores for the completion of the partograph in respect of sector scores	181
4.3.10	Odds Ratios for various subsector scores (with standard error adjusted for 17 clusters)	182
4.3.11	Difference in combined observations (maternal, fetal and progress of labour) between arms, adjusted for 17 clusters	183
4.3.12	Comparison of scores obtained in respect of alignment to clinical guidelines and frequency of observations	184
4.3.13	Odds Ratios for innovations	186
4.3.14	Comparison between arms of pre- AND post-contraction fetal heart observations in the records with partographs utilised	186
4.3.15	Descriptive statistics for scores for the critical comprehensive assessment	187

4.3.16	Comparison of comprehensive assessment for those records where this was indicated	188
4.3.17	Partographs where reasons were recorded for non-completion of the partograph	189
4.3.18	Comparison between arms of the study in respect of recording reasons for non-completion of the partograph	189
4.3.19	Reasons recorded for observations not being done	190
4.3.20	Births occurring within 1.5 hours of the last observation	191
4.3.21	Births occurring before a second set of observations could be done ..	192
4.4.1	Registered midwives who enrolled in the trial (per site) and participated at each timepoint	193
4.4.2	Descriptive statistics of numbers of midwives enrolled per arm.....	193
4.4.3a	Intervention arm - Test three theory vs clinical completions	195
4.4.3b	Control arm - Test three theory vs clinical completions	197
4.4.4	Descriptive statistics for total scores for test three	197
4.4.5	Total scores for test three ('intention to treat')	198
4.4.6	Descriptive statistics for test one (baseline) and test three (month 12) total scores ('per protocol')	199
4.4.7	GEE result for total score for test three 'per protocol' adjusted for baseline	200
4.4.8a	Pattern of completions for Test three - theory component	201
4.4.8b	Pattern of completions for Test three - clinical component.....	202
4.4.9	Descriptive statistics for types of questions for test three ('intention to treat')	204
4.4.10	GEE results for types of questions for test three ('intention to treat') ...	204
4.4.11	Descriptive statistics for types of questions for test three ('per protocol').....	205
4.4.12	GEE results for types of questions for test three ('per protocol')	206
4.4.13	Descriptive statistics for types of questions for those who completed theory or clinical components	207
4.4.14	GEE results for types of questions for test three for those who completed theory or clinical components	207
4.4.15	Descriptive statistics for knowledge and application scores ('intention to treat')	209
4.4.16	GEE model results for knowledge and application scores for test three ('intention to treat')	209
4.4.17	Descriptive statistics for test one (baseline) and test three (month 12) knowledge and application scores ('per protocol').....	210
4.4.18	GEE model results for knowledge and application scores for test three ('per protocol' analysis) (adjusted for baseline score)	210
4.4.19	Descriptive statistics for Test three theory component and practice aspect scores ('intention to treat')	212
4.4.20	GEE results for Test three theory component and sector scores ('intention to treat')	213

4.4.21	All subjects with valid baseline scores Test one (Baseline randomisation).....	214
4.4.22	Subjects with valid baseline and follow-up scores –Test one (Theory T1) and Test three (Theory T3) ('per-protocol' analysis)	214
4.4.23	Subjects with valid follow-up scores – Test three.....	214
4.4.24	All subjects with valid baseline clinical scores – Test one (Baseline randomisation)	215
4.4.25	Subjects with valid baseline and follow-up scores – Test one (Clinical T1) and Test three (Clinical T3) ('per-protocol' analysis)	215
4.4.26	Subjects with valid follow-up clinical scores – Test three.....	215
4.4.27	Descriptive statistics for total score (reflected as a percentage) at three timepoints	217
4.4.28	Descriptive statistics for theory and clinical components at three timepoints	217

Acronyms, Abbreviations and Definitions

AIDS : Acquired Immune Deficiency Syndrome

ANOVA : ANalysis Of VAriance

ANCOVA : ANalysis of CO-VAriance

ART : Anti-Retroviral Therapy

BANC : Basic ANtenatal Care

CCC : Concordance Correlation Coefficient described by Lin (1989)

CD4 : Cluster of Differentiation 4 - an immunoglobulin, used as a marker where low levels indicate immune suppression

CINAHL : Cumulative Index to Nursing and Allied Health Literature (database)

CIOMS : Council for International Organizations of Medical Sciences¹

Clinical Facilitator : For the purpose of this study a clinical facilitator is an experienced registered midwife who has responsibility for intrapartum clinical care and who works alongside less experienced registered midwives. Her / his role is to facilitate a positive learning climate in the labour ward, to assist colleagues to improve clinical insight, clinical judgment and clinical skills in order to improve the standard of patient care. The clinical facilitator may also be the nurse manager who is responsible for quality assurance and clinical audit of the labour ward, and who supports and complements the senior midwife in the clinical facilitation role. **Clinical facilitation** relates to the activities and behaviours of the clinical facilitator to address the clinical development needs of the registered midwives with whom s/he works.

CME : Continuing Medical Education

CPD : Continuing Professional Development

CRT (or C-RCT) : Cluster randomised trial. This is the recommended term, but is used interchangeably with **cluster randomised controlled trial, group randomised trial** or **community randomised trial**.

¹ UK English spelling is used throughout. However, US English spelling has been reflected where a direct quote, a reference, or the formal name of an organisation is spelt according to this convention, e.g. Council for Organizations of Medical Sciences

CVI : Content Validity Index

EOV : Educational Outreach Visit

EPOC : Cochrane Effective Practice and Organisation of Care group

ERIC : Educational Resource Information Centre (database)

DCST : District Clinical Specialist Team

DHS : Demographic and Health Survey

DSS : Demographic Surveillance Sites

ENND : Early neonatal death: Death of a neonate within the first 7 days of life

F2F : Face-to-face training

FIGO : International Federation of Gynecologists and Obstetricians

GEE : Generalised Estimation Equations

HIV : Human Immune deficiency Virus

ICC : Intraclass Correlation Coefficient

ICM : International Confederation of Midwives

IMAI : Integrated Management of Adolescent and Adult Illness

IMCI : Integrated Management of Childhood Illnesses

ITT : Intention to treat

KT : Knowledge translation

MCWH: Maternal, Child and Women's Health: A subdirectorate of the clinical services directorate in the Department of Health of the Provincial Government of the Western Cape at the time that the field work was done.

MDG : Millennium Development Goal

MOU : Midwife Obstetric Unit: A freestanding primary level midwife-run midwifery service offering continuum of care for pregnancy and childbirth

NCCEMD: National Committee for the Confidential Enquiry into Maternal Deaths: A committee appointed by the national Minister of Health

NICU: Neonatal Intensive Care Unit

NS : Not Significant

OR : Odds Ratio

OSCE: Objective Structured Clinical Evaluation

Partograph : A recording tool used to record observations of labour, and which assists in determining prolonged labour. In South Africa this term is used interchangeably with the term **partogram**

Partograph utilisation : This term is used to refer to the correctness and completeness of the record of observations made on a partograph. The partograph is recommended by the World Health Organization for all labours. (World Health Organization 1989, World Health Organization 2003)

PEP : Perinatal Education Programme

PPA : Per protocol analysis

PI : Practice Improvement

PPIP : Perinatal Problem Identification Programme

PGWC: Provincial Government of the Western Cape

RCT : Randomised Controlled Trial

RD : Risk difference

RN : Registered nurse

RM : Registered midwife

SANC : South African Nursing Council : The regulatory body for nurses and midwives in South Africa, established under the Nursing Act No 33 of 2005.

SD (or sd) : Standard Deviation

Skilled attendant : “An accredited health professional - such as a midwife, doctor or nurse – who has been educated and trained to proficiency in the skills needed to manage normal (uncomplicated) pregnancies, childbirth and the immediate postnatal period, and in the identification, management and referral of complications in women and newborns” (World Health Organization 2004:1).

TB : Tuberculosis

TOT : Training of trainers

UK : United Kingdom

UN : United Nations

USA : United States of America

USN : Unilateral Spatial Neglect

VBAC : Vaginal Birth After Caesarean

WHO: World Health Organization

*“A health system that delivers for women
when women are ready to deliver,
is a strong health system that will benefit all people.”*

(Ban Ki-Moon 2009)

Chapter 1 Introduction

1.1 Overview

The South African national 'Saving Mothers' and 'Saving Babies' reports over the past decade have noted that nearly half of all maternal deaths were possibly or probably avoidable and nearly half of perinatal deaths due to labour related complications were probably avoidable (Pattinson 2007, Department of Health 2009). Of those that related to health care providers, the inadequate utilisation of the partograph and inappropriate response to abnormal observations was reported (Department of Health 2006a, Macdonald, Bartlett *et al.* 2007).

Recommendations have been made for the correct utilisation of the partograph in all institutions conducting births, and for a quality assurance programme to be implemented using an appropriate tool (Department of Health 2002, Pattinson 2003, Department of Health 2006a, Pattinson 2007, Department of Health 2009, Pattinson 2011). The purpose of this study was to contribute to this national priority for intrapartum care.

Registered nurse-midwives render the majority of the care to the pregnant population of South Africa. Nurse-midwives are empowered by legislation to take responsibility for the care of pregnant women with normal pregnancies, and where there are complications associated with a pregnancy, there are clear parameters as to their professional responsibility (South African Nursing Council 1987; South African Nursing Council 1990).

In caring for a pregnant woman, every clinical interaction between the midwife and pregnant woman has implications for the wellbeing of at least two individuals - the mother and the fetus, and ultimately the neonate. Labour is a dynamic process which needs to be monitored using specialised assessment skills, so that any indication of prolonged labour with no progress is speedily recognised and managed appropriately. Prolonged labour has implications both for the mother (as this may indicate an underlying obstetric pathology)

(Pattinson 2003) and for the fetus which might develop intrapartum hypoxia. This in turn could result in neonatal encephalopathy, which may have serious long term sequelae, or even death (Greenfield, Rhoda *et al.* 2011). Therefore, it is essential for midwives to maintain professional competence and expertise for all aspects of pregnancy and newborn care. In particular, the monitoring and care of the pregnant woman and fetus during labour is critical. Further, as 31% of child deaths occur amongst neonates in South Africa, attention to the early neonatal outcomes will make a substantial impact on the rate of childhood deaths (Lawn and Kerber 2006).

This study specifically targeted the primary level of maternity care as this is where midwives are required to exercise the greatest autonomy. Further, the study was conducted in two rural regions where there is less medical support than in the metro region. A pragmatic cluster randomised trial was conducted where the unit of randomisation was the cluster. The intervention consisted of two components (training of on-site clinical facilitators and targeted training of registered midwives in intrapartum care), and the effect was measured by evaluating the quality of recording of observations in labour (partograph utilisation audit) and testing the knowledge and skills of midwives.

In addition, because none of the instruments measuring the utilisation (completion) of the partograph that were reported in the literature had been subjected to rigorous validation and testing, such a tool was developed, validated and tested for reliability in this population.

1.2 Background to the study

The rest of this chapter describes the background to the study. This includes an overview of relevant health indicators, the context and setting in which the study was conducted, maternal and perinatal health care provision and the scope of practice of the registered midwife, management of labour, as well as a description of the sources of the problem associated with poor intrapartum care in South Africa. The chapter concludes by presenting the motivation for the study.

1.3 Maternal and Neonatal indicators

1.3.1 International indicators

Five years before the adoption of the Millennium Development Goals (MDG) by the United Nations (UN) in 2000, it was estimated that worldwide, approximately 515 000 women died each year as a result of pregnancy related causes, giving an estimated global Maternal Mortality Ratio of 397/100 000. More than half of these deaths occurred in Africa (Hill, AbouZhar *et al.* 2001). Hogan, Foreman *et al.* (2010) put the annual number of maternal deaths in 2008 at 342 900 with a global Maternal Mortality Ratio (MMR) of 251/100 000, down from 320/100 000 in 1990 (the foundation on which the target for MDG5 is based). The World Health Organization announced a fall in the worldwide number of maternal deaths for 2013, giving a MMR of 289/100 000 (World Health Organization 2014).

The UN Millennium Declaration has amongst its goals and targets the reduction of maternal mortality by 75% by 2015 (goal 5) (United Nations 2009). However by focussing on maternal death, this ignores the burden of suffering caused by pregnancy-related complications, for women, their families, and their communities (Graham and Hussein 2004). Maternal disability or death may impact more widely than on the woman herself. She may be neither able (through disability) nor present (through death) to look after her children who become orphaned, and dependent on their communities or the state. Further, compromised maternal health status (including nutritional status) as well as poor prevention and treatment of maternal conditions during pregnancy contribute to stillbirth and neonatal mortality (Darmstadt, Yakoob *et al.* 2009).

The highest burden of maternal mortality is experienced in Africa (van den Broek and Falconer 2011) where women are often poor, have limited education, live in rural areas or places with inadequate social service infrastructure, and are disenfranchised or live in a patriarchal society. In South Africa, race has had an additional impact so that black women have been described as quadruply oppressed (Pick, Hoffman *et al.* 1990). Although the legacy of apartheid of unequal resourcing and inadequate health service provision to certain areas or

population groups is in the process of being addressed, the legacy on the health status and health capacity of women remains.

Added to this, the scourge of Human Immune deficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS) in Sub-Saharan Africa and in South Africa in particular, has reversed the gains in reducing maternal mortality made between 1980 and 1990 (Bradshaw and Dorrington 2012). However, the evidence suggests that the peak of the epidemic might have been reached (Department of Health 2009, UNAIDS 2013). The arrest and control of the spread of HIV/AIDS is recognised as a key issue for sustainable global development and has been included in MDG six, along with halting the spread of malaria and other diseases (United Nations 2009).

Millennium Development Goal four, which aims to reduce the number of childhood deaths by two-thirds (United Nations 2003), is also of relevance to this study. This MDG4 will not be achieved unless there is a substantial reduction in neonatal deaths. In 2000 it was estimated that 38% of all child deaths occurred in the neonatal period, and three-quarters of all neonatal deaths occur in the first week of life – the highest risk being on the first day (Lawn, Cousens *et al.* 2005). In addition the number of third trimester stillbirths, a third of which are estimated to have occurred in the intrapartum period (Darmstadt, Yakoob *et al.* 2009), is similar to the three million early neonatal deaths (Lawn, Blencowe *et al.* 2011). In low to middle income countries these are caused largely by complications occurring in labour and childbirth, e.g. prolonged or obstructed labour or umbilical cord accidents.

The cost of poor pregnancy outcomes related to neonatal morbidity affects health service provision (the use and allocation of scarce resources - equipment, staff, time) as well as having long-term consequences for the capacity of the affected individual to learn and ultimately gain employment and support him/herself (Lawn and Kerber 2006) .

International figures currently show that despite the many international commitments of nations to improving maternal and child health, the health status of mothers and children remains poor in most developing countries. The same is true for South Africa despite its middle income status (Shung King, de Pinho *et al.* 2006). Concerted efforts must be made to rectify and address the core contributing factors to poor maternal and child health. Whilst many of these lie outside of the health sector, the health sector has a significant contribution to make, and at the very least health interventions with known effectiveness must be implemented in the best possible way (Shung King, de Pinho *et al.* 2006).

1.3.2 National indicators

1.3.2.1 Maternal mortality ratio

The backdrop to this study was an estimated national MMR of 175–200/100 000 live births (Department of Health 2002)². This suggests a picture of a developing country. The estimated MMR for the Western Cape Province in which this study was located was 54/100 000 (Department of Health 2002).

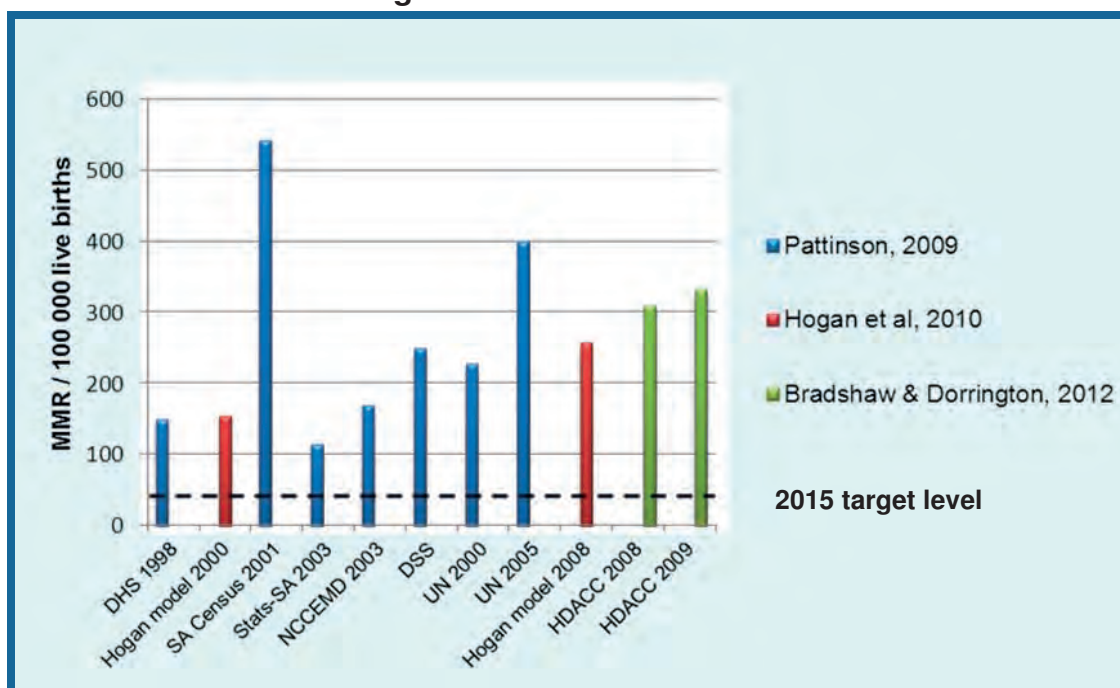
It is recognised that tracking the level of maternal mortality in developing countries, where vital registration and other population data is inadequate, is challenging (Bradshaw and Dorrington 2012). It should be noted that there are a number of estimates of the population-based national MMR where different denominators are used, e.g. demographic and health survey (DHS), census, demographic surveillance sites (DSS), and the UN interagency group for mortality estimates (Bradshaw and Dorrington 2012). The accuracy of such a ratio is further compromised where the birth data is not necessarily complete, because out-of-hospital births (particularly those where a stillbirth or neonatal death has occurred) are under-registered (Graham and Newell 2009, Lawn, Yakoob *et al.* 2009).

In the large 181-nation study undertaken by Hogan, Foreman *et al.* (2010) from the Institute for Health Metrics Evaluation (IHME), which used robust modelling to try and develop comparative data, it was estimated that the South

² This was the most recent 'Saving Mothers' report at the time the study was planned.

African MMR was 155/100 000 in 2000 and 258/100 000 in 2008. However, this method, and its revised estimate of 91/100 000 for 2011, has been described as unrealistic as there was unclear case definition and the resultant inclusion of HIV-related deaths which were pregnancy-related deaths but not maternal (Garenne and McCaa 2010, Shelton and Gray 2010, Bradshaw and Dorrington 2012). The UN interagency group on maternal mortality estimation (MMEIG) uses a model which takes into account *inter alia* HIV deaths, Gross Domestic Product, fertility rate, and the proportion of women assisted during birth by a skilled attendant. The various estimates are presented in Figure 1.1.

Figure 1.1 Maternal mortality ratio (MMR) estimates for South Africa, with 2015 target indicated



Such models facilitate some comparison across large numbers of countries with minimum datasets. However, in order to be able to monitor more closely the situation at a national level and to inform and effect policy, the [South African] Health Data Advisory and Co-ordinating Committee (HDACC) recommended basing the indicator of maternal mortality on vital registration, adjusted for under-reporting (8%) and misclassification of causes (14%). This resulted in the MMR for 2008 being revised from 167/100 000 to just over 200/100 000

(Bradshaw and Dorrington 2012). Further taking into account the MMEIG recommendation to add 50% for under-reporting of maternal deaths as being maternal, this figure was set at 310/100 000 for 2008. Applying the same method for 2009 the estimate was 333/100 000 (Bradshaw and Dorrington 2012).

The major source of South African maternal data is from the National Committee for the Confidential Enquiry into Maternal Deaths (NCCEMD) which publishes its findings in the 'Saving Mothers' reports. When the first 'Saving Mothers' report was released in 1999, the Minister of Health noted, "In almost half of all the maternal deaths reported there was an opportunity to prevent that death, but that opportunity was missed" (Minister of Health 1999:1).

The majority of missed opportunities by health providers (nurses, midwives and doctors) occur at the primary level. The most frequent health care provider avoidable factors were failure to follow standard protocols (which include use of the partograph to monitor women in labour), and poor problem recognition and initial assessment (Department of Health 2002, Department of Health 2006a, Department of Health 2009, Department of Health 2012). These have remained largely unchanged over the twelve years of the reports. While it is possible to monitor trends in these categories over time, the parameters may be difficult to define and measure reliably, particularly in the population of labours that does not result in a maternal death.

In addition to the need to ensure adequate screening and treatment of the major causes of maternal death and skilful management of emergency care, recommendations to address the findings of the Saving Mothers report include improving health care provider education and health system strengthening (Department of Health 2012).

1.3.2.2 Perinatal and Neonatal mortality

The major source of perinatal data in South Africa is provided by a Perinatal Care survey (Perinatal Problem Identification Programme – PPIP) which publishes the 'Saving Babies' reports. The PPIP gathers national information

from sentinel sites and reports triannually. In rural areas 13.5% of all perinatal deaths were regarded as health worker related where missed opportunities and substandard care contributed to deaths that were probably avoidable. Further, delays in referral or calling for expert assistance would be expected to have had a direct effect on the management and course of labour (Pattinson 2003).

The 'Saving Babies' reports state that the early neonatal mortality rate was 8.5/1000 in infants with a birth weight of more than 1000g. This has been largely unchanged over the ten years (2000-2009) that this data has been collected (Greenfield, Rhoda *et al.* 2011). In 'Saving Babies 2008–2009', intrapartum asphyxia and birth trauma were the leading causes of neonatal deaths and were also the leading cause of perinatal death for babies $\geq 1000\text{g}$ (excluding unexplained deaths) and for district hospitals³ (Pattinson 2011). The most common cause of death in neonates over 2000g was due to hypoxia (Greenfield, Rhoda *et al.* 2011). On-site reviewers of the deaths felt that 46% of the deaths due to labour-related complications were probably avoidable had the health care provider acted appropriately (Pattinson 2011).

The 'Saving Babies V' report highlights that the most frequent avoidable factors are health worker related and states that the overall substandard care in the intrapartum period reflects a dismal picture of mismanagement, demonstrating a lack of basic skills or an unwillingness to follow clinical protocols (Macdonald, Bartlett *et al.* 2007). Specific health worker avoidable factors applicable to intrapartum care included lack of appropriate monitoring in labour, incorrect interpretation of findings (specifically of the partograph), and inappropriate action in relation to observations (Greenfield, Arends *et al.* 2005, Macdonald, Bartlett *et al.* 2007). These too were consistent over the same ten year reporting period (Pattinson 2003, Pattinson 2007, Pattinson 2011).

Intrapartum hypoxia is one of the leading causes of death in the perinatal period (Pattinson 2003). Hypoxia in labour will cause various levels of stress / distress

³ The majority of district hospitals are in rural areas.

in the fetus, but severe hypoxia will cause neonatal encephalopathy. In the Western Cape, the province in which this study was situated, intrapartum hypoxia, which is directly related to intrapartum care, accounts for 10.6% of obstetric causes of neonatal deaths $\geq 1000g$, while hypoxia accounts for 28.7% of neonatal causes of death in this category of neonate (Greenfield, Arends *et al.* 2005). This latter figure has been dropping over the past few years but remains the highest rate after immaturity (Rhoda personal communication 2012).

1.3.2.3 Strategies to address maternal and perinatal mortality in South Africa

Among the five key strategies to decrease perinatal mortality is the following: was to ensure that all settings where intrapartum care was offered had the necessary equipment, clinical guidelines and suitably qualified health care providers who knew how to use a partograph, and that a quality assurance tool should be used to monitor the effective utilisation of the partograph (Pattinson 2003). Similar recommendations have been made consistently up to the 2008-2009 report (Pattinson 2011) which mirror those made in the 'Saving Mothers' reports (Department of Health 2002, Department of Health 2006a).

1.3.2.4 The impact of HIV on maternal and perinatal mortality in South Africa

In addition to the obstetric and neonatal challenges *per se* is that of HIV and AIDS (relating to MDG6). Amongst the pregnant population attending public health facilities in South Africa, the prevalence of HIV was 29.4% in 2009 (Padayatchi, Naidoo *et al.* 2011). Non-pregnancy-related infections accounted for 40.5% of all maternal deaths in South Africa in the 2008 – 2010 triennium, most of these deaths being due to HIV infection complicated by tuberculosis (TB), pneumocystis carinii pneumonia (PCP) and pneumonia (Department of Health 2012). HIV and AIDS are not major underlying causes of death in labour and babies. Other than the association found with intrapartum asphyxia and HIV infection (Fawcus, Kennedy *et al.* 2012), there is little impact on labour itself.

1.4 Study setting

In 2006 when this study was designed, the Western Cape Province was divided into 4 regions – 1 metro and 3 rural regions⁴. This study was conducted in two of the rural regions (see map **Appendix 1**).

Region 1 (Boland / Overberg) comprised the area to the east of the Cape Town metro boundary along the south coast as far as Swellendam (approximately 220km) and northwards to Ceres and Montagu, bordering the little Karoo.

Economy

This is a largely rural area with winemaking, deciduous fruit, wheat, sheep and some dairy farming accounting for the bulk of the economy (Government Communication and Information System (GCIS) 2008). There is quite a change in population dynamics at harvest time as seasonal workers arrive from other parts of the country. It is a dry region and is served by the Breede and Berg river systems.

Infrastructure

The region is served by two major national roads – the one travelling north-east in the direction of the Zimbabwean border and the other eastwards in the direction of the Mozambican border. There are good regional roads between the towns in the region, but transport is problematic for farmworkers.

Population

The Community Survey of 2007 indicated that there was a population of 507 587 (Statistics South Africa 2007).

Health service

The referral hospital at the time of the intervention and data collection was situated in Worcester and road journeys could take 1 - 2½ hours from the district hospitals. The administrative centre was also located in Worcester.

⁴ This has subsequently changed into 6 districts (and 4 sub-districts of the metro district). The two regions in this study became three districts, but for the purpose of this dissertation the situation as it existed at the time the study was implemented, is described.

There were two Midwife Obstetric Units (MOUs)⁵ in this region, one of which (site five) referred to a hospital in the metro region due to proximity and road access.

The district hospitals (with on-site medical staff) each have a distinct obstetrics and gynaecology ward (covering medical and surgical conditions) which incorporates a labour ward of two or three beds. On average there could be two to four births per day per hospital. Usually there would be one registered nurse-midwife per 12 hour shift assisted by one or two enrolled nursing auxiliaries, with a unit manager (also a nurse-midwife) available during the day⁶. Staff were appointed as registered midwives and tended to stay within this department but some staff rotated through from other units, e.g. trauma, as need dictated.

The smaller hospitals (utilising sessional medical practitioners) had one or two wards (usually split between male and female patients) and the staff was rotated through all sections. In some cases there was only one registered nurse-midwife on duty per shift (with assistance from an enrolled nurse or enrolled nursing auxiliary) who would have to attend to the ward (of approximately 30 adults and children with varying illnesses), casualty and trauma, theatre, and the labour ward (which consisted of a single dedicated room). Such services would attend to approximately one birth per day, but would also assess women (who presented themselves thinking that they were in labour) and decide if they were to be admitted or sent home to await onset of labour. The registered nurse in charge of these hospitals was also responsible for all administration of the hospital as well as fulfilling clinical responsibilities. In some of these hospitals registered nurses also fulfilled the role of pharmacist, radiographer, or ultrasonographer.

⁵ Similar to a freestanding (primary level) birthing unit in other countries.

⁶ Despite numerous attempts to obtain formal staff complements for the sites from institutional and provincial clinical and human resource departments, these were not available.

Outreach visits from the regional obstetrician occurred regularly. These visits covered all primary level sites and included perinatal review meetings (audit and feedback). These continued as usual throughout the study.

Region 2 (West Coast / Winelands) comprised the area to the north of the Cape Town metro border to Vredendal in the north (± 400 km) and extended north-eastwards to Stellenbosch, Malmesbury and the Swartland region bordered by the Cedarberg mountain range on the east and the Atlantic coast on the west.

Economy

It is a largely agricultural region with wheat, citrus and grapes being the main crops, sheep and goat farming, and winemaking. The local fishing industry (South African Government Information Service 2008) has been decimated in the past 10 years but there are still pockets of subsistence fishing communities. There are low levels of full time employment as much of the employment is seasonal and the population swells at harvest time. It is a very dry region but is served by the abundant Olifants River and its extensive canal system (South African Government Information Service 2008).

Infrastructure

The region is served by a national road which ends at the Namibian border. There is a reasonably good network of regional roads, but the sparsely populated mountainous areas have poor access.

Population

The Community Survey of 2007 indicated that there was a population of 704 364 (Statistics South Africa 2007).

There was a similar profile between the two regions in terms of household size, accommodation, access to electricity, piped water, toilet facilities, refuse removal and household appliances (Government Communication and Information System (GCIS) 2008).

Health service

The referral hospital at the time of the intervention and data collection was in the south-east corner of this region at Paarl, necessitating a journey of approximately three hours from the furthest point. The regional administrative centre was located at Malmesbury.

The health service had similar dynamics to those described for region 1, with the exception of the outreach visits by the regional obstetrician which did not take place according to a regular schedule at the time of the study.

1.5 Maternity care in the public sector health service

Maternity care in the public sector is offered at primary, secondary and tertiary levels. In pregnancies which are assessed as being low risk, women receive their antenatal care at the primary level. A separate risk assessment for labour is done and the woman may continue receiving care at the primary facility or be referred to another level of the health service for intrapartum care. Primary level maternity care may be located in an MOU, a comprehensive health centre, a level 1 district hospital, or a satellite clinic of the MOU or hospital. These services are staffed by registered midwives who have responsibility for managing the pregnancy and labour, and whose practice is regulated by the South African Nursing Council. The National Department of Health has published national guidelines for maternity care at the primary level since 2000 (Department of Health 2000, Department of Health 2007).

In general, the care in labour is undertaken by midwives. The clinical responsibility lies with them at primary level whereas at the higher levels of the service the responsibility for decision making lies with the medical professionals. Any aspect of maternal or fetal wellbeing that indicates a complication or abnormality should be referred to a secondary hospital and from there, if necessary, more severe complications or abnormalities are referred to a tertiary level hospital or academic health centre for clinical management. Referral is governed by regulations under the Nursing Act as well as provincial policies and clinical guidelines.

With the development of comprehensive health centres and the expansion of maternity services to all community health facilities, there are many registered nurses (who trained many years ago as midwives) who have not practised as midwives for some time who are now seeing pregnant women as part of the antenatal service. There is some concern that during this upscaling of the health service there might be some dilution of the quality of antenatal care which could impact those services offering intrapartum care.

1.6 Nursing and Midwifery

In South Africa the majority of registered nurses are also qualified as registered midwives (South African Nursing Council 2004). While minimum clinical requirements in terms of hours and skills are stipulated, there is a concern amongst educators that higher order skills of critical thinking and clinical problem-solving and clinical judgment are not well developed (Uys, van Rhyn *et al.* 2004, Uys and Meyer 2005). This is not limited to South Africa (Muir 2004, Tanner 2006). Indeed some studies suggest that the teaching approaches required to foster these skills are not widely practised (Uys and Meyer 2005, Tanner 2006).

Once a person attains registration as a nurse and / or midwife, there is no database that indicates which of these individuals retain currency of knowledge and clinical expertise. In the metropolitan areas, due to the population pressure and the size / capacity of health services, there are posts for registered nurses (who are registered as midwives) in services which are designated for maternal and perinatal care only. However, in the rural areas the registered nurse is expected to retain competence in all fields of Nursing and Midwifery. The staff allocation system favours rotation of staff through all areas. This does not facilitate the development and consolidation of expertise in areas that require in-depth knowledge and skills despite many calls for this practice to be stopped (Pattinson 2011). In effect this means that these professionals function in intrapartum settings as 'part-time midwives' because they do not have the opportunity to consolidate their expertise (Clow 2006).

Further, opportunities for professional development are constrained by shortages of staff and relative lack of opportunities in the rural regions to access appropriate professional development training. From a human resource development perspective, there appears to be little interrogation of health indicators in relation to the training / reskilling / upskilling required by health professionals, with an accompanying plan to meet these needs.

In addition to the special expertise in intrapartum care, all levels of the health service are struggling with the shortage of Nursing staff as nurses opt to move out of the public sector, move out of the health sector, or seek opportunities abroad. This is exacerbated in specialised areas such as Midwifery. This means that expertise can easily be lost from a system if one or two individuals leave. Building and maintaining clinical expertise can not be achieved in an *ad hoc* manner, but needs an intentional strategy at a health system level.

1.6.1 Midwifery and the scope of practice in respect of intrapartum care

The International Confederation of Midwives (ICM) recognises a midwife *inter alia* as someone who has successfully completed a midwifery education programme based on the ICM Essential Competencies for Basic Midwifery Practice (International Confederation of Midwives 2011a) and the framework of the ICM Global Standards for Midwifery Education (International Confederation of Midwives 2011b). Such a programme should be recognised in the country where it is located. In order to use the title ‘midwife’ a person should have acquired the requisite qualifications to be registered and/or legally licensed to practise midwifery; and have demonstrated competency in the practice of midwifery. The midwife is recognised as a responsible and accountable professional who works in partnership with women to give the necessary support, care and advice during pregnancy, labour and the postpartum period, who conducts births on her/his own responsibility and provides care for the newborn and the infant. Such care includes preventative measures, the promotion of physiological birth, the detection of complications in mother and child, the accessing of medical care or other appropriate assistance and the

carrying out of emergency measures (International Confederation of Midwives 2011a). The definition also specifies each of the aspects of practice s/he should be able to undertake and where this may be offered.

The Midwifery model of care, adopted by the ICM, is based on the premise that pregnancy and birth are normal life events. This model of care is woman-centred model and includes monitoring the physical, psychological, spiritual and social wellbeing of the woman and family throughout the childbearing cycle; continuous attendance during labour, birth and the immediate postpartum period; minimising technological interventions; and identifying and referring women who require obstetric or other specialist attention (International Confederation of Midwives 2002).

The ICM agreed on a set of competencies that could be expected of a trained midwife (International Confederation of Midwives 2002). The aim of this document is to give sufficient guidance so that the competencies can be used to establish appropriate and locally sensitive standards⁷. The competencies are divided into six broad areas covering : 1) generic knowledge, skills and behaviours, 2) pre-pregnancy care and family planning methods, 3) care and counselling during pregnancy, 4) care during labour and birth, 5) postnatal care of women, and 6) newborn care. Only those competencies relating to care during labour and birth that are of specific relevance for the scope of this study have been selected, as these informed the development of the training programme and evaluation of the midwives in this trial.

ICM competency four refers to midwives providing high quality, culturally sensitive care during labour, and conducting the delivery in a clean and safe manner, which includes managing selected emergency situations relating to the

⁷ A more recent descriptive document setting out the essential competencies for basic midwifery practice was adopted at the international congress of the ICM held in June 2011 (International Confederation of Midwives, 2011a) which was revised further in 2013, but this had not been updated at the time that the study was designed.

health of women and their newborn (International Confederation of Midwives 2002). Basic knowledge and skills required to support this competency include : appropriate anatomy and physiology to enable recognition of the commencement of labour and to assess the wellbeing of the woman and the fetus and the progress of labour through all stages; appropriate management of normal labour, birth and transition to extra-uterine life; and recognition of complications or emergencies. The competency specifically states that knowledge of and skills in how to use a partograph or similar tool are required. Other basic skills specified are the taking of vital signs, undertaking a screening physical examination, performing an abdominal assessment and pelvic examination, and assessing effectiveness of uterine contractions (International Confederation of Midwives 2002). Additional skills related to this competency include managing malpresentation, shoulder dystocia and fetal distress, identification of and managing a prolapsed cord, appropriate use of oxytocics for labour induction or augmentation and treatment of postpartum haemorrhage, and timely transfer of women for additional / emergency care (International Confederation of Midwives 2002).

Skills relating to fetal monitoring which are specified under the skills for the competency of care and counselling during pregnancy are also relevant to care in labour, viz. listening to the fetal heart rate and palpating the uterus for fetal activity pattern, and monitoring the fetal heart rate with Doppler (International Confederation of Midwives 2002).

In South Africa, the practice of registered midwives is regulated by the South African Nursing Council, under the authority of the Nursing Act. The 1978 Act (No. 50) was replaced by Act No. 33 of 2005, although this has not yet been implemented fully. The Scope of Practice regulations, R1469 (South African Nursing Council 1987) and the Practice regulations for midwives, R2488

(South African Nursing Council 1990)⁸ are still defined in terms of the 1978 Act, but these are likely to change in the near future once the new Act is implemented⁹.

The Scope of Practice regulation R1469 describes a 'midwifery regimen' as relating to any midwifery intervention which may have an influence on the course and management of pregnancy, all stages of labour and the puerperium. This should include the provision of care plans, their implementation and evaluation. Further any health problem identified, and the care received by the mother and child whilst in the care of the midwife, should be recorded (South African Nursing Council 1987).

The scope of practice also specifies that a registered midwife should apply scientifically based acts or procedures in her/his care of the mother and child in the course of pregnancy, labour and the puerperium (South African Nursing Council 1987). This includes diagnosing a health need and initiating an appropriate midwifery care plan, administering treatment as prescribed, monitoring the progress of pregnancy, labour and the puerperium (including the vital signs of the mother and child) (South African Nursing Council 1987). Although a partograph is not specified in the regulations, this would be relevant in terms of monitoring labour.

The conditions under which a registered midwife may carry on his (*sic*) profession (R2488) reiterates the requirement for keeping records of all acts and emergency acts performed in connection with the mother and child, and also specifies situations in which a registered midwife must summon medical assistance or make a referral (South African Nursing Council 1990). Those relevant to this proposal include : Illnesses, abnormalities or complications occurring during labour, undue prolongation of any stage of labour, disordered or abnormal uterine action, presentation or prolapse of the cord, fetal distress,

⁸ It is recognised that these regulations are dated, but they are the ones in current use.

⁹ This had no influence on the study. This information is included to clarify the legal standing of the regulations at a time of transition from one Act to another.

and while waiting for the arrival of the medical practitioner should deal with the emergency to the best of her/his ability (South African Nursing Council 1990).

Clearly, registered midwives in South Africa are recognised as having professional skills and are accountable for their practice in managing normal labour, recognising complications as they arise and dealing with them, as well as ensuring appropriate and timely referral to an appropriate practitioner / level of care. Thus the scope of this project is well within the scope of what a service might reasonably be able to expect of registered midwives.

1.6.2 Midwifery knowledge and skills

While there may be a legal or policy framework in place, it is essential that the people who are required to provide the care are adequately prepared educationally, are supported in their practice and have the opportunity for regular educational and skills updates.

In 2004, prior to this study being undertaken, a baseline assessment was done on the knowledge, skills and clinical decision-making abilities of a convenience sample of 102 registered midwives (who attended orientation and training) in all four regions of the Western Cape Province (across all levels of the public sector health service). This was carried out at the time of the rollout of the revised partograph (**Appendix 2**) which was accompanied by provincial guidelines for labour care (Clow 2005).

The mean scores at baseline showed that, while midwives could complete the partograph, more than half were not assessing the risk status, and nearly 80% had errors in the way in which observations were recorded which could lead to clinical misinterpretation. Questions set at the lower cognitive levels had a mean score of 56% while those at the higher cognitive levels (analysis and application) had a mean score of 38%. The implication of these findings are that it is this more complex thinking, necessary to make reasoned conclusions of the often complex developments in labour, which is lacking (Clow 2005).

Comparisons were done between the 88 midwives working in primary level public sector health services in the metro region (where they are more likely to be in posts which focus on maternity care) and the rural regions (where registered midwives work simultaneously as registered nurses). Overall there were lower scores in the rural regions. While there were improvements evident in the post-test, these were greater in the metro region. In general there was a lower than desired level of knowledge and analytical skill apparent amongst all the midwives in this earlier study (Clow 2006).

1.7 Management of labour

At the launch of the first report of the National Confidential Enquiry into Maternal Deaths the national Minister of Health (1999) noted that the high rate of deaths of mothers during pregnancy and childbirth in South Africa was not due to a lack of knowledge on how to manage severely ill pregnant women, but that the challenge was to focus our energies on every possible strategy and mechanism to prevent any death of woman in pregnancy. While the reasons for maternal and neonatal morbidity and mortality are many, and some fall outside the remit of the health service *per se*, there are some that can be addressed directly within the health sector and specifically by health care providers.

Scientifically sound evidence relevant to South Africa on what should be done is available. This was used to formulate proposals for managing intrapartum care in South Africa by a multiprofessional group of health professionals and administrators involved in intrapartum care (Medical Research Council Unit for Maternal and Infant Health Care Strategies 2006). These have subsequently been incorporated into national guidelines (Department of Health 2007). The challenge is to disseminate the evidence and guidelines as widely as possible to the appropriate areas of the health service, and to accompany these with targeted and effective training and monitoring to ensure that they are implemented effectively.

1.8 Defining the problem

At a **clinical level**, the common modifiable factors are *inter alia* poor intrapartum monitoring including lack of utilisation of a partograph, and prolonged abnormal monitoring without action (Greenfield, Arends *et al.* 2005, Woods, Pattinson *et al.* 2005, Macdonald, Bartlett *et al.* 2007, Department of Health 2009). Thus, improving the management of labour is critical for the wellbeing of both women and neonates.

Given that the management of labour in primary level health services is the responsibility of registered midwives, i.e. professionals who have already undertaken a basic training leading to registration and license to practise as a registered midwife, it is of concern that in both 'Saving Mothers' and 'Saving Babies' reports there is a consistent problem identified - that the partograph is not utilised appropriately and labour is poorly managed (Pattinson 2007, Department of Health 2009).

To date, interventions in this area have concentrated on completion of the partograph. In this study, this subject was approached from a fresh perspective – that of appropriate clinical decision making (based on the observations made and recorded on the partograph), and its effect on improved clinical outcomes. The interest of this research project was thus on the following two questions.

1. Is it that the process and progress of labour are inadequately understood, and thus observations are done incompletely resulting in premature or inaccurate conclusions?
2. Is it that clinical skills are poorly developed and that there is little confidence in the ability to make a firm clinical finding?

At an **educational level**, current training programmes tend to address the lower levels of cognitive functioning, i.e. knowledge and comprehension, rather than the more complex levels, e.g. application, analysis and synthesis which are necessary for quality clinical practice (Quinn 1988).

Unless there is clinical understanding and reasoning for its use, the partograph will remain just a recording instrument. Any efforts to increase its disciplined and complete utilisation must address the fundamental issues of clinical reasoning and decision-making, and for these one needs at least the knowledge and competencies (Fullerton, Thompson *et al.* 2013) to conduct an accurate and full clinical assessment, and to analyse and apply the theory to the clinical findings.

However, a passive transfer of information is not sufficient to galvanise action in individuals. Within each section of the **health system** (in this case the maternity unit) there need to be mechanisms that support and encourage appropriate and good quality clinical management, and a system that limits the possibilities for poor clinical management or error (Headrick 2000). To this end both individual knowledge and motivation, and system support are required.

In a trial of an active multifaceted information dissemination intervention Gülmezoglu, Villar *et al.* (2004) noted that was easier to deliver the intervention (workshops) to a group at the hospital level (unit) than to select individuals working in the same unit. When addressing the staff as a group, group dynamics and peer pressure can be utilised, which may facilitate the adoption of the recommended practices. One of the three workshops in that study specifically focussed on implementing change. However, Gülmezoglu, Villar *et al.* (2004) noted that the science and process of behaviour change was generally not familiar to the clinicians [obstetric consultants] who were tasked with delivering the intervention.

At the **health system management level** there appears to be no coherent response to addressing obvious shortcomings in clinical practice. Currently there is an *ad hoc* approach to continuing professional development with little, if any, consideration of its effectiveness; greater prominence should be given to strategic and ongoing human resource development planning in the (provincial) Department of Health.

1.9 Conclusion

The motivation for the project came from the realisation that improvements in the standard of care during the intrapartum period could reduce the avoidable factors implicated in maternal and perinatal morbidity and mortality. This resonates with the World Health Organization and other international agencies' calls for skilled attendance at birth as a crucial intervention strategy.

Furthermore, the wellbeing of mothers and babies are recognised as key development indicators as these are articulated in two of the eight United Nations Millennium Development goals (United Nations 2003).

This project was born out of the work done by the Provincial maternity guidelines group, and was supported by the province. Because clinical service managers were exposed to the findings of the 'Saving Mothers' and 'Savings Babies' reports and were involved in the ongoing PPIP surveillance, there was a high degree of awareness that 'something needed to be done'. This project was regarded as a positive response to the health challenges, and the clinical service managers were enthusiastic to participate.

Chapter 2 Literature Review

2.1 Introduction

This literature review presents the current knowledge of the topics relevant to this project and illustrates how they have influenced the design of the intervention and instrumentation.

The common approach to the literature search is presented (2.2.1.1) and then those that are topic-specific follow (2.2.1.2-2.2.3). A description of the theoretical constructs regarding practice change which guided the process, *viz.* Knowledge Translation (KT) and Practice Improvement (PI) will be presented (2.3), followed by the key features identified in the practice change / implementation literature (2.4). Thereafter the literature relating to the field of mentorship will be presented (2.4.8) followed by literature regarding instrumentation to evaluate partograph utilization (2.4.9). The chapter will conclude by identifying gaps in the literature (2.5) and the implications for this study (2.6).

2.2 Literature search strategies and results

2.2.1 Changing practice

2.2.1.1 Common approaches

The search strategy had the following characteristics in common for the changing practice topics considered :

- ◆ databases consulted – Pubmed, Medline, CINAHL, Cochrane Library (EPOC group and Reproductive Health Library), EBSCO Africa wide
- ◆ years of publications – 2006 onwards (the original literature review for the proposal incorporated work up to 2006). Where necessary or appropriate this was extended back to 2000
- ◆ language – no language restriction (English abstracts for non-English articles)

- ◆ inclusion criteria
 - ◇ research design - cluster randomised trials, randomised controlled trials, controlled trials, evaluation studies, review
 - ◇ type of participant - groups of health care professionals
 - ◇ setting - health service level
 - ◇ type of intervention - package of relevant interventions, e.g. dissemination of guidelines and implementation, establishment of practice standards, skills training
 - ◇ outcomes - objective measures of health care professional knowledge and clinical skills, objective measures of documentation of care
- ◆ exclusion criteria - studies with no midwife or nurse. This was not applied to reviews
- ◆ search strategy and sourcing of literature - identification of systematic reviews or major review articles; key government and WHO monographs and reports; relevant national legislation or international instruments; follow-up of key works which predated the search period which included books.

Grimshaw, McAuley *et al.* (2003) raise the issue of the breadth of a review, where the 'lumping' or 'splitting' decision needs to be made. While this literature review does not claim to be a systematic review, this has been resolved by lumping as it was found that there was a small proportion of literature where midwives or nurses were included as participants. However, this highlights the need for understanding and sensitivity to context as findings cannot necessarily be transferred without some interpretation.

There were a number of relevant and recent systematic reviews done through the Cochrane Collaboration. These were regarded as being comprehensive and good quality. However, as highlighted later in this chapter the studies included tended to be physician orientated and situated in urban setting in high income countries. For this reason, the systematic findings of the systematic reviews

were taken as a foundation. Where good quality individual studies which addressed Midwifery, Nursing, maternity care, intrapartum care and/or in resource constrained settings, these are also included.

Where there were additional criteria or differences these are indicated in the sections that follow, where relevant. A flow chart representing the search results for these three strategies is presented in Figure 2.1 (see pages 28 and 29).

2.2.1.2 Knowledge translation

In addition to the common approach described in 2.2.1.1, the search strategy included the following :

- ◆ Keywords - Knowledge translation and review
- ◆ Inclusion criteria - health care, health system
- ◆ Exclusion criteria - theoretical papers without application

In order to include studies that were relevant to Midwifery or Nursing, or to settings other than high income country settings, further studies with different designs were sought.

2.2.1.3 Practice Improvement interventions

In addition to the common approach described in 2.2.1.1, the search strategy included the following :

- ◆ Keywords - Practice improvement and (maternity or Midwifery or intrapartum)
- ◆ Inclusion criteria - implementation process for practice improvement, i.e. the 'how' rather than the 'what')
- ◆ Exclusion criteria – lack of an implementation process

In order to include studies that were relevant to Midwifery, Nursing or perinatal care, additional studies which did not comply with the research design criteria were included.

2.2.1.4 Implementation methods

Much work has been done by the Cochrane Effective Practice and Organisation of Care (EPOC) group to promote and conduct systematic reviews in the area of professional practice and the delivery of effective health services by evaluating and analysing the effectiveness of professional, organisational, financial and regulatory interventions.

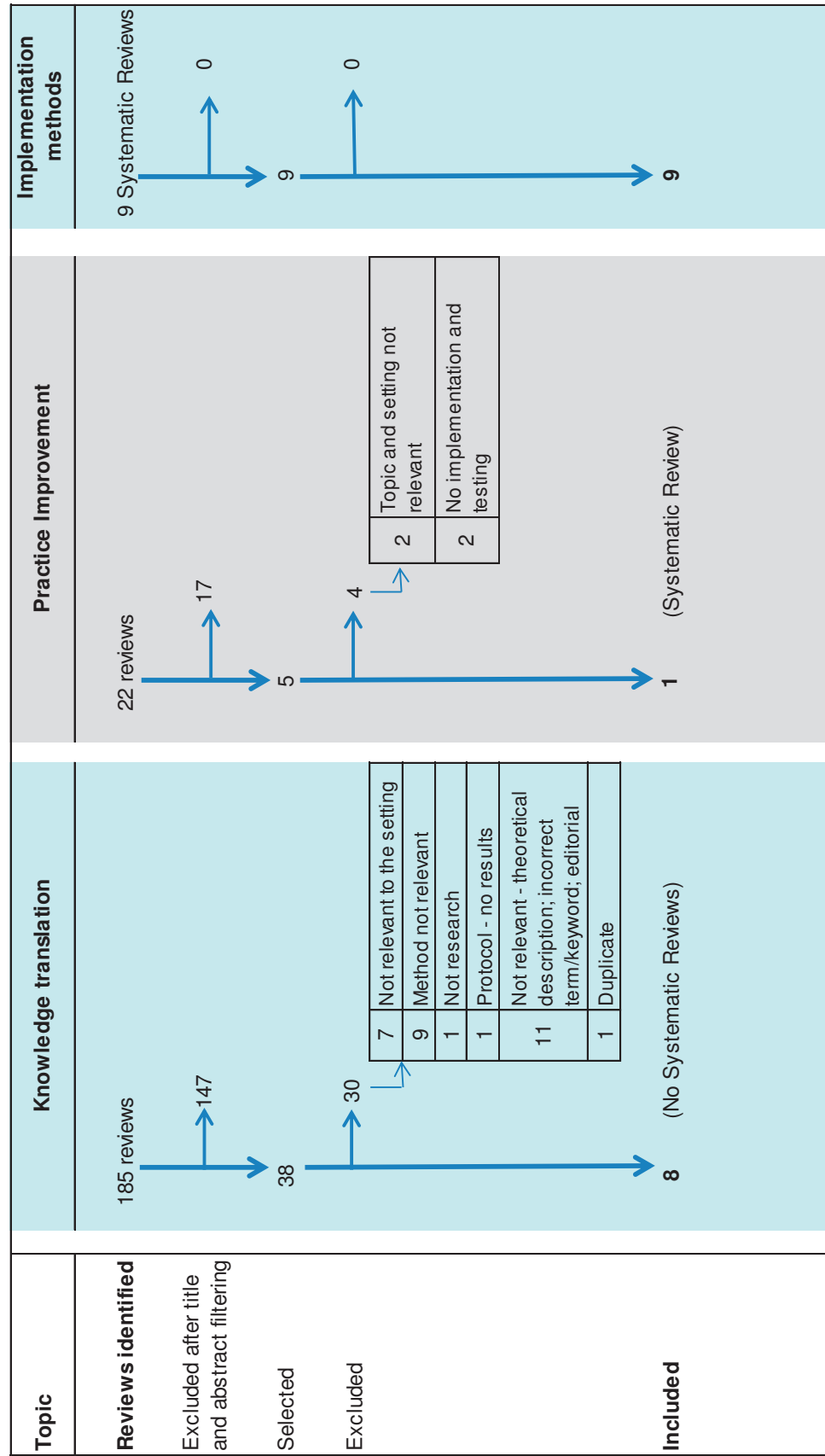
The field of professional quality improvement lacks a generally accepted classification system, making it difficult to source all existing research in this field (Grimshaw, McAuley *et al.* 2003). This area of practice has a specific focus in the Cochrane Collaboration through the EPOC group. This group has an extensive and comprehensive search strategy (which includes Medline, EMBASE, CINAHL and SIGLE) (Grimshaw, McAuley *et al.* 2003) and so systematic reviews emanating from this group are likely to give high quality evidence and cover.

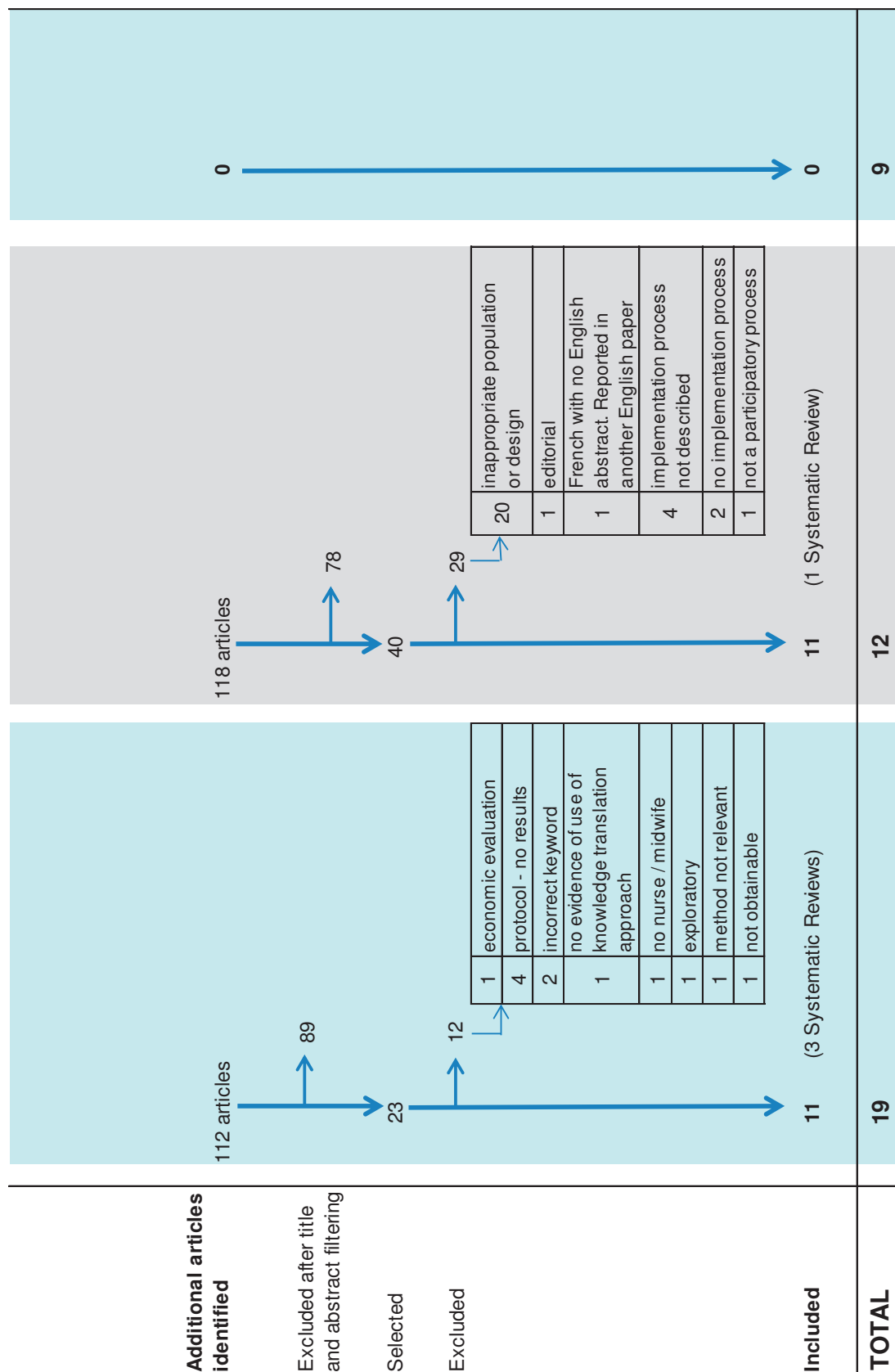
The literature search for this aspect was focussed on the systematic reviews in the Cochrane Library (EPOC group) and the WHO Reproductive Health Library. The most recent versions available of the systematic reviews relevant to the topics identified above were selected. In all, nine were included.

2.2.2 Mentorship

Literature consulted on practice improvement and knowledge translation did not identify mentorship *per se* although educational outreach / academic detailing and audit and feedback, which have been shown to be effective, have elements of the type of communication required in mentorship and coaching. This study's intervention was designed to strengthen the clinical supervision endeavour. Some of the strategies employed to facilitate the relationship between midwives and clinical supervisors were those which are used in coaching and mentoring, hence this subject was explored. The focus of the literature review on mentorship was to determine the effectiveness of mentorship programmes among registered midwives or nurses in terms of patient or health outcomes. To this end randomised controlled trials and systematic reviews were selected rather than descriptive or theoretical papers.

Figure 2.1 Literature search results





2.2.2.1 Search strategy

The search strategy includes the following :

- ◆ databases consulted – Pubmed, Medline, CINAHL, Cochrane Library (including EPOC group and Reproductive Health Library), Africa-wide, ERIC (Education Resources Information Center), Academic Search Premier and SCOPUS
- ◆ years of publications – 2006 onwards. Where necessary or appropriate this was extended back to 2000
- ◆ language – no language restriction (English abstracts for non-English articles)
- ◆ inclusion criteria
 - ◇ research design - randomised controlled trials, cluster randomised trials, controlled trials, systematic review
 - ◇ type of participant – registered midwives or registered nurses (also Midwifery or Nursing)
 - ◇ setting – clinical practice, primary level health care settings
 - ◇ type of intervention – mentor/ship, clinical supervision, clinical leadership, clinical facilitator/facilitation
 - ◇ outcomes – objectively assessed patient or health outcomes
- ◆ exclusion criteria - studies with no midwife or nurse, pre-registration programmes or pre-service learning, postgraduate academic programmes, research programmes, assessment, preceptor/ship, 'sign-off' mentor, supervisor/vision of midwives, regulated supervision, counselling, therapy, consultant, psycho-analysis/therapy, under-performance, assessment, pharmacological studies, descriptive studies, theoretical discussions, letters, editorials
- ◆ search strategy and sourcing of literature - identification of systematic reviews or major review articles; follow-up of key works which predated the search period which included books, as well as sourcing appropriate professional or government regulations. Additional references were found by reference checking. Initially studies were considered by title and

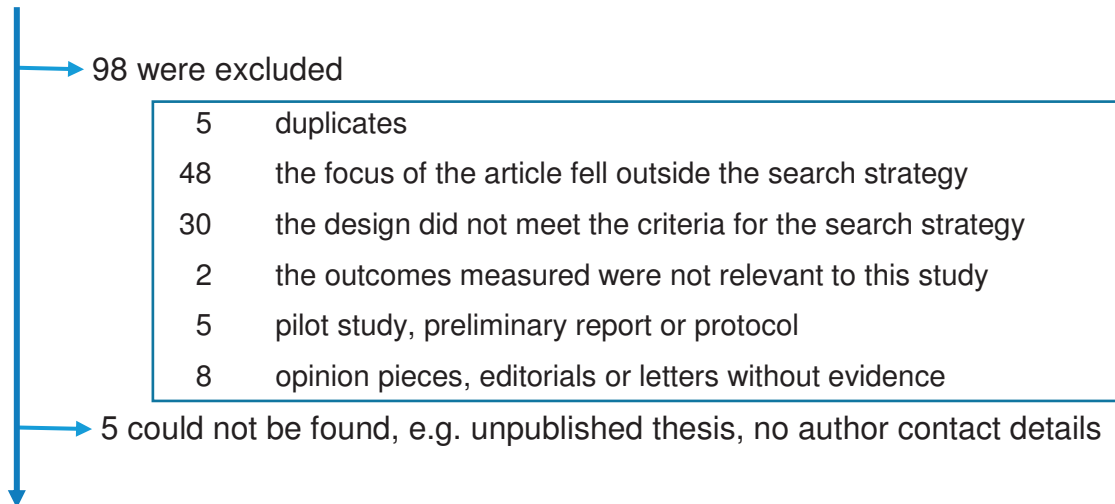
abstract (where available). Promising articles and those without abstracts were followed up for full text screening.

The majority of articles sourced were not intervention studies so did not yield definitive evidence of impact or effectiveness. It was therefore decided to consider articles from descriptive and qualitative studies that showed promise for informing this study as well as those which assisted in clarifying the terminology in this field and to delineate the field.

The terms 'clinical supervision' and 'clinical leadership' were not clearly defined in the MeSH system. Searches using these words yielded large numbers of studies which included 'clinical' or 'supervision' or 'leadership' but did not necessarily reflect the intended meaning. 'Clinical' in a research context frequently referred to pharmacological research and such articles were excluded. Two areas in the literature relating to supervision are supervision of midwives (which is context-specific in that it has a clearly defined role within the regulation of midwives in the United Kingdom)(Steele 2009)) and supervision of mental health professionals (which is a formal arrangement whereby a senior/qualified health practitioner supports, directs and guides the work of a junior colleague through case-based review and personal reflection (Brunero and Stein-Parbury 2008, Milne, Sheikh *et al.* 2011) and which may be accountable to a regulatory body (Clark, Jamieson *et al.* 2006)). These terms were excluded in the search strategy as they did not match the requirements for this study, and therefore will not be presented. Clinical facilitator/facilitation yielded only five articles and these were all included within the mentorship literature. Thus the search was narrowed to 'mentor' and 'mentorship' as keywords.

2.2.2.2 Search results

The search identified 130 references



The final number of references included was 27, of which only five addressed effectiveness of mentorship in terms of patient or health outcomes directly. Other studies, representing a variety of research designs were retained to assist in describing this field of research and practice.

2.2.3 Instruments to evaluate partograph utilisation

2.2.3.1 Search strategy

The search strategy for instruments that evaluated partograph utilisation included the following :

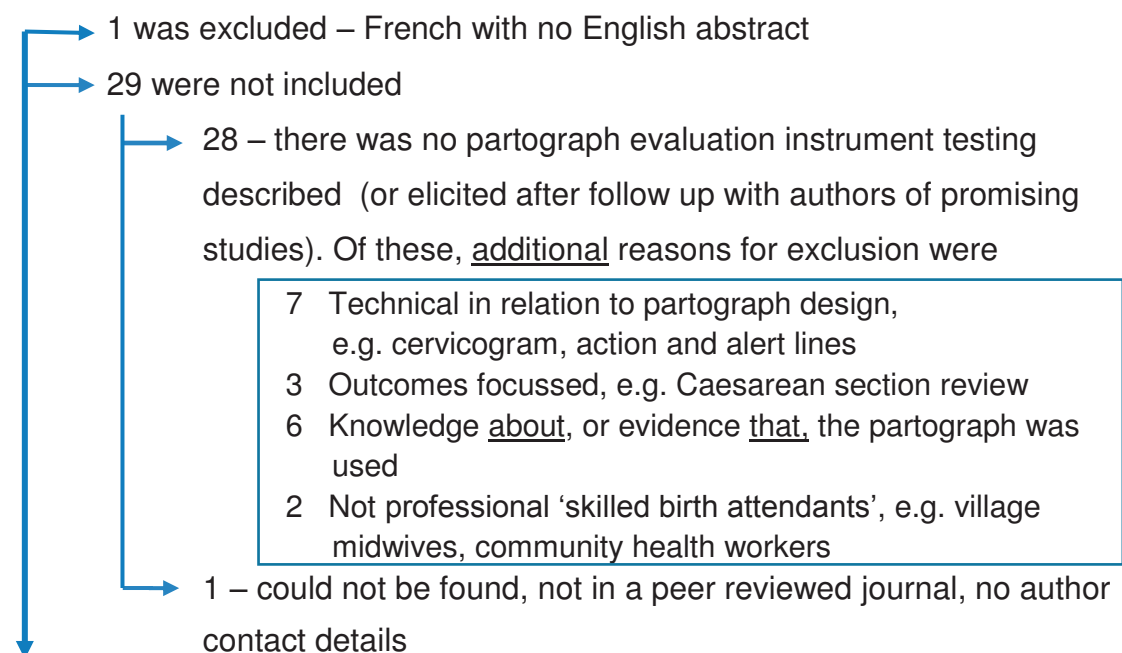
- ◆ databases consulted – Pubmed, Medline, CINAHL, Cochrane Library (including Reproductive Health Library), EBSCO Africa wide
- ◆ years of publications – 2006 onwards
- ◆ parameters of the search – English, English abstracts for non-English articles
- ◆ keywords – partograph or partogram; assessment / evaluation / audit
- ◆ inclusion and exclusion criteria
 - ◇ inclusion – partograph evaluation, hospital / clinic based
 - ◇ exclusion – impact on maternal or neonatal outcomes; absence of partograph evaluation instrument, or validation or reliability testing procedures; second stage; absence of professional ‘skilled attendants’

- ♦ search strategy and sourcing of literature – identification of systematic reviews or major review articles; key government and WHO monographs and reports; where known relevant research was not yet published, known local studies which were relevant to this topic even though they fell outside the time limit. Where the contents of an article held promise for the formal testing of an instrument, this was followed up with the corresponding (or other) author/s, where possible. Further, key researchers in this field were contacted to ascertain if they were aware of any literature that met the criteria.

The use of the term 'labour record' or 'intrapartum record' yielded many articles but these did not include the partograph, which was the focus of this sub-study. Similarly 'assessment', 'evaluation' and 'audit' even when combined with 'partogram' or 'partograph' yielded articles from many medical disciplines without adding value to what had already been obtained. Thus it was decided not to pursue this line further and to restrict the review to 'partogram' and 'partograph' as keywords.

2.2.3.2 Search results

The combined search yielded 35 publications



Thus five are reported in this review.

Some promising references falling prior to the specified period were followed up. None reported a validation process or reliability testing of an instrument to assess the use of the partograph, and therefore were effectively excluded according to the criteria stated in the search strategy. However, the findings from the search are summarised below.

2.3 Describing the field of improving the quality of clinical practice

Over the past 25 years, the field of improving practice has moved from inspection and retrospective review of discipline-specific core functions associated with quality assurance programmes towards a more concurrent and interactive approach. A paradigmatic shift occurred in the late 1990s when it was recognised that there was a need to improve processes in order to enhance outcomes (Page 1999). An example of this is the development of process indicators for basic and comprehensive maternity care. Achieving the indicators cannot be done in isolation and is dependent on various parts of a system supporting the work of the whole system. Thus the previous linear or vertical approach has shifted towards a holistic or system-orientated approach.

Numerous studies have attempted to demonstrate effectiveness in changing practice. These highlight that this is not a simple activity as there are many factors in the system which support or block effective change, e.g. practitioners, management systems, access to resources, infrastructure, organisational culture, bureaucracy, etc. Indeed Oxman, Thomson *et al.* (1995) concluded that there is no 'magic bullet' for provider behaviour change.

While the term 'quality of care' might be used quite liberally, it is interpreted variously depending on the perspective of the person evaluating this concept. Raven, Tolhurst *et al.* (2011) concluded there is no single and comprehensive definition. When considering approaches to quality of care for maternal and neonatal health care specifically, they suggested that these should include a rights-based approach, adopting care that is evidence-based, consideration of the mother-baby dyad as interdependent and the fact that pregnancy is

generally a healthy state. They propose a model of quality of maternal and newborn health care which can be used as a basis for developing quality improvement strategies and activities, and incorporating quality into existing programmes (Raven, Tolhurst *et al.* 2011).

Quality management in health has a large lexicon of terms that have evolved as the field has evolved. Sometimes these terms are used interchangeably, and can create confusion. A similar situation exists in the area of translational science which intersects with quality clinical management. Indeed a review of titles in the literature underscores this, e.g. ‘Lost in Translation’; ‘Tower of Babel’, ‘... wallowing in transformation’, to highlight a few. Mitchell, Fisher *et al.* (2010) suggest that the lack of standardised terminology reflects the fact that translational science is a young and developing multidimensional field incorporating many disciplines and organisations, both across health sciences and in the fields of marketing, communication, education and management. While many of these terms are inter-related, the focus of this study is on knowledge translation (and related implementation science) and practice improvement (and related clinical practice improvement).

Knowledge Translation and Practice Improvement approaches are complementary. The Knowledge-to-Action model used for Knowledge Translation is explicit in its intention to create knowledge and apply it. On the other hand, the Model of Improvement is a generic model which has been applied in various sectors including the health sector where it has a clinical focus. Although this model refers to building knowledge while changes are being made, the emphasis in health care appears to be in improving practice and the processes and systems that contribute to this.

2.3.1 Knowledge translation

Knowledge translation (KT) is the “exchange, synthesis and ethically-sound application of knowledge – within a complex system of interactions among researchers and users – to accelerate the capture of the benefits of research ... through improved health, more effective services and products, and a

strengthened health care system” (Canadian Institutes of Health Research 2009). Knowledge translation is used interchangeably by some authors with the term ‘knowledge transfer into action’. It moves beyond passive dissemination and transfer of knowledge to how knowledge is used, be this by participants or decision makers at any level in the health system so as to inform health-related decision making (Straus, Tetroe *et al.* 2011). Knowledge translation and implementation science appear to be similar, with the latter being the scientific study of methods to promote the systematic uptake of clinical research findings and other evidence based practices into routine practice and hence to improve the quality and effectiveness of health care (Foy, Eccles *et al.* 2001). The differentiation appears to depend on which side of the Atlantic the work is being conducted.

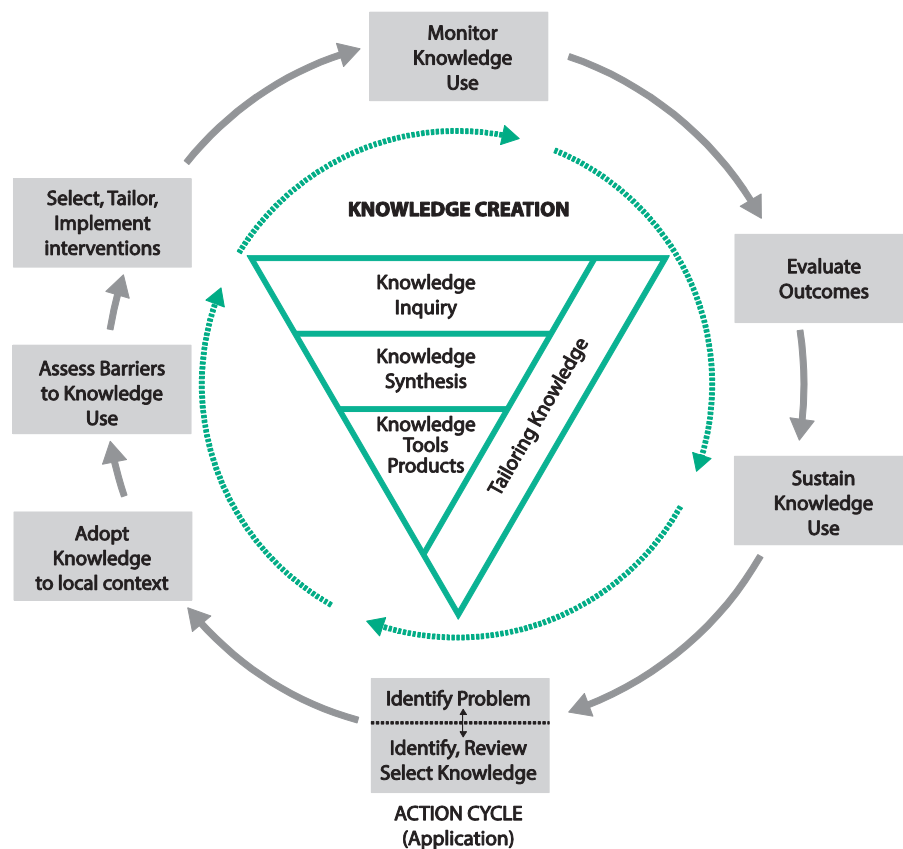
Mitchell, Fisher *et al.* (2010) have identified four thematic areas in the field of translational sciences which they derived from 47 models for knowledge translation found after an extensive literature search. These thematic areas are

1. Evidence based practice, research utilisation and knowledge transformation processes, e.g. development of clinical guidelines, technology assessment, standards of care
2. Strategic / organisational change theory to promote uptake and adoption of new knowledge, e.g. use of various activities such as audit and feedback, use of opinion leaders and facilitation to promote practitioners’ adoption of practices that are based on best evidence
3. Knowledge exchange and synthesis for application and enquiry, e.g. regular and ongoing interactions between researchers, clinicians, policy makers and consumers accelerate the application of new discoveries in clinical care
4. Designing and interpreting dissemination research, e.g. develop generalisable empirical evidence to determine the effectiveness of an intervention with widespread application (Mitchell, Fisher *et al.* 2010).

However, they acknowledge that each thematic area is not fully discrete and several of the models incorporate elements from more than one thematic area. The main focus of this study is in thematic area 2.

A comprehensive knowledge-to-action framework (illustrated in Figure 2.2) was developed on commonalities found in an assessment of more than 30 planned action theories (Graham, Logan *et al.* 2006, Straus, Tetroe *et al.* 2011).

Figure 2.2 Knowledge to Action Framework



Source : Graham ID, J. Logan *et al* (2006) *J. Continuing Education in the Health Professions* 26(1):19

This includes two major components – knowledge creation and knowledge application (action cycle). The knowledge creation component is described as knowledge being sifted through a funnel, from inquiry (asking the right questions) to synthesis (pulling together research and information from other sources) and to the development of products (delivering the right information in the right format), in order to become refined and presumably more useful for

end users. Within the action cycle there are seven phases which may occur sequentially or in parallel and at any time there may be interaction with the knowledge creation process (Graham, Logan *et al.* 2006, Straus, Tetroe *et al.* 2011). The main focus of this study is on the knowledge application component.

2.3.2 Practice Improvement

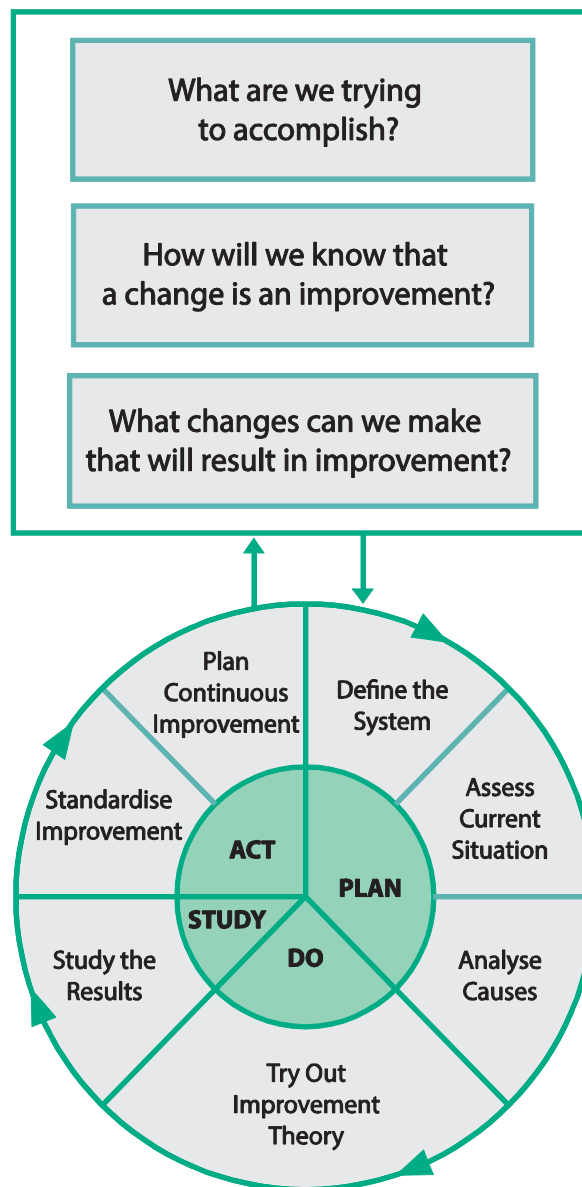
Clinical practice improvement “profiles the care process by analysing the content and timing of individual steps of a medical care process” (Horn 1997, cited by Haxton and Fahy 2009).

The Model for Improvement (Figure 2.3), which has been adopted by various health care organisations and units for clinical practice improvement, is based on the sequential building of knowledge. It is centred on three questions that are fundamental to all improvement activities and the Plan-Do-Study-Act (PDSA) cycle. These are :

1. What are we trying to accomplish?
2. How will we know that a change is an improvement?
3. What changes can we make that will result in an improvement? (Langley, Nolan *et al.* 2009)

The questions and the PDSA cycle allow for application to be as simple or as sophisticated as needed, depending on the situation and the people involved. The PDSA cycle employs a rapid cycle approach which allows for incremental changes (multiple cycles) to contribute to larger scale systems and process improvements, allowing for knowledge to be built while changes are being tested, thus reducing risk (NHS Scotland 2007, Langley, Nolan *et al.* 2009).

Figure 2.3 Model for Improvement



Sources : Institute for Health Care Improvement; Division of Specialized Care for Children, University of Illinois at Chicago

The Clinical Practice Improvement process has five recognised phases :

1. The project phase - where the problem is identified, goals developed and determination of who should be included;
2. The diagnostic phase – diagnosing, describing and measuring the problem, and identifying and implementing of a number of interventions to reduce the problem;

3. Intervention phase – this consists of the PDSA cycle when changes identified in the diagnostic phase are implemented;
4. Impact and evaluation phases - these measure and record the effects of the changes;
5. Sustaining improvement phase – this involves continuing monitoring and planning for future improvements (New South Wales Health Department 2002).

It was thus appropriate to search the literature on both Knowledge Translation and Practice Improvement.

2.4 Presentation of literature

The literature relating to practice change will be presented using the following concepts :

- ◆ Context and understanding the problem / issue (2.4.1)
- ◆ Participation and participatory process (2.4.2)
- ◆ Behaviour change (2.4.3)
- ◆ Implementation strategies (2.4.4)
- ◆ Standardisation (2.4.5)
- ◆ Evaluation (2.4.6)
- ◆ Health system issues (2.4.7)

Where possible, systematic reviews have been located. These are summarised in **Appendix 3a** and paragraph references given¹⁰. Because they cover various health outcomes, studies which are specific to intrapartum care or Midwifery complement the text where they have a specific contribution.

Thereafter the literature related to mentorship will be presented (2.4.8) followed by literature on instruments to evaluate partograph utilisation (2.4.9).

¹⁰ This will be noted in the text with an asterisk (*). They are presented in alphabetical order according to the first author's surname.

2.4.1 Context and understanding of the problem / issue

A critical piece for any quality improvement undertaking is a thorough understanding of the context in which this is planned. A variety of instruments may be used to identify and describe the extent of a problem. Some of these originated in the industrial setting, i.e. flow-mapping, flow-charting, cause and effect diagrams and Pareto charting. Others, such as clinical audit, have gained prominence in the clinical arena, while various strategies to ensure that all inputs are identified may be adopted, e.g. brainstorming and multi-voting (Colton 2000, Haxton and Fahy 2009).

Haxton and Fahy (2009) aimed to reduce the length of stay of pregnant women who presented as outpatients in the delivery suite during pregnancy, safely and effectively. A database of emergency and outpatients presenting in maternity was initiated (which had not existed previously) to identify the extent of this practice, and an out-patient flow chart to identify duplications and unnecessary delays was developed. Clinical practice improvement strategies such as brainstorming and multi-voting were used.

A study to reduce perineal trauma in a tertiary setting in New South Wales, Australia had a comprehensive team representing the various role players involved. They utilised flowcharts of the labour and birth process, a cause and effect diagram relating to the high incidence of perineal trauma, a Pareto distribution chart on the causes of perineal trauma, and a list of customer / patient and health care provider expectations (Nicholl and Cattell 2006).

Pitchforth, Lilford *et al.* (2010) combined clinical and social science perspectives and methods in a pilot study in Ethiopia to assess and understand issues affecting the quality of clinical care for labour, and to identify priorities for change.

2.4.2 Participation and participatory process

The Model of Improvement which has gained traction is that of PDSA quality improvement cycles (New South Wales Health Department 2002, Langley,

Nolan *et al.* 2009, Siassakos, Fox *et al.* 2011). This has similar characteristics to the Participatory Action Research and Action Learning approaches described by Zuber-Skerritt (Zuber-Skerritt 2002) and others (McTaggart 1997, Baum, MacDougall *et al.* 2006, Kemmis 2007), although these have a greater emphasis on reflection throughout the processes. These approaches all require active participation by relevant participants.

Several studies have reported inclusion of various combinations of Midwifery clinicians, educators and managers, medical staff including residents, registrars, general practitioners and specialist obstetricians and paediatricians, pharmacy departments and pharmacists, management, and patients / community to address a variety of challenges. These included quality of care in Ethiopia (Pitchforth, Lilford *et al.* 2010), service re-engineering in Australia (Haxton and Fahy 2009), reducing postpartum haemorrhage in a remote hospital near the India / Tibet border (Mercer, Sevar *et al.* 2006), reduction in perineal trauma (Nicholl and Cattell 2006), prescription practice in an intensive care unit (Rajamani, Suen *et al.* 2011), and teamwork and improvement in quality and safety in the United Kingdom (Siassakos, Fox *et al.* 2011). All reported substantial changes in processes, relationships and practice.

2.4.3 Behaviour change

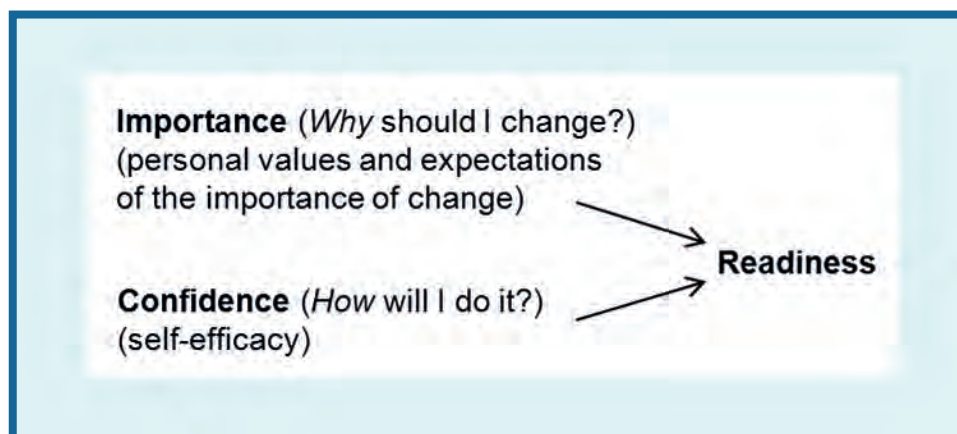
When introducing new ways of doing things or even revisiting familiar territory, one should to anticipate personal and organisational dynamics which may be ready to embrace change or may choose to resist change. In the health literature this often refers to changes in behaviour of the patient, particularly in the field of addiction, but it seems that there are similarities of human behaviours that can be applied to practitioners who are challenged to change entrenched practices.

In a pilot study in Ethiopia the reasons for not complying with good tenets of care related to motivation to change, practical knowledge and confidence to

change as well as perceived importance of some aspects of good practice, e.g. skin-to-skin care and the use of regional anaesthesia (Pitchforth, Lilford *et al.* 2010).

Rollnick, Mason *et al.* (1999) addressed behaviour change with motivational interviewing, which considered the need for information exchange and reducing resistance in order to build motivation for and confidence to make the desired change. This was based on the 'Stages of Change' transtheoretical model (Prochaska and DiClemente 1982) which described readiness for change and how people move towards making decisions and changing behaviour in their lives. The model defined five stages – pre-contemplation, contemplation, preparation, action and maintenance. If someone has not even contemplated the possibility of change, then they are unlikely to be ready to act on this. This impacts at both personal and group levels, so when entering a group process one should assess readiness for change. Different approaches were required depending on how receptive and ready for change the group was. Other factors that affect readiness are importance and confidence (Figure 2.4) (Rollnick, Butler *et al.* 1997). These might include costs and benefits to the individual or group, as well as their sense of being able to achieve the proposed change.

Figure 2.4 The ingredients of readiness to change



Source : Rollnick, Mason *et al.* (1999,22)

Cheater, Baker *et al.* (2006) evaluating both audit and feedback, and educational outreach on nurses' behaviour and patient outcomes, highlighted the need for future studies to be informed by individual theories of change and suggested the Stages of Change transtheoretical model of Prochaska and DiClemente (1982). Cheater, Baker *et al.* (2006) included training of link nurses (who provided the educational outreach) in motivational interviewing techniques to help resolve ambivalence and support nurses in their plans for change.

Michie, van Stralen *et al.* (2011) evaluated existing frameworks of behaviour change and developed the 'Behaviour Change Wheel' (BCW) that is comprehensive, coherent, and has a clear link to an overarching model of behaviour. At its centre is the **behaviour system** comprising **capability**, **opportunity** and **motivation** (COM-B). This is surrounded by nine potential interventions / strategies which in turn are surrounded by seven policy / infrastructural areas that could enable the interventions. This model recognises the contextual and system issues which are often ignored when an intervention is considered.

2.4.4 Implementation strategies

The classification of professional interventions from the Cochrane EPOC group taxonomy describes eleven methods / categories (Grimshaw, McAuley *et al.* 2003), not all of which are relevant to this study. Only relevant methods will be presented further, *viz.* educational outreach, audit and feedback, local opinion leaders, distribution of educational materials and continuing education. Given the complexity of presenting findings of various combinations of strategies, **Appendix 3b** summarises these with the related effects, and is complementary to the text. Paragraph references are indicated to link to the text¹¹.

Grimshaw, Thomas *et al.* (2004)*[§] found that analysis and reporting of multifaceted interventions was challenging due to the fact that there were 68 different combinations of interventions and 136 comparisons. Many of the multifaceted combinations included only one or two comparisons making it

¹¹ The studies will be noted in the text with [§]. These are presented by topic in Appendix 3b.

difficult to reach any substantial conclusion. No relationship was found between the number of component interventions and the effects of multifaceted interventions. However, more recent systematic reviews of certain implementation methods have shown a similar or greater effect when used in combination than alone (O'Brien, Rogers *et al.* 2007, Flodgren, Parmelli *et al.* 2011, Ivers, Jamtvedt *et al.* 2012). The authors of an earlier review concluded that multifaceted interventions tended to be more effective, although this was accompanied by increased cost (Wensing, van der Weijden *et al.* 1998). Among physical therapists, multi-component knowledge translation interventions were more effective than passive dissemination or a single knowledge translation intervention (Menon, Korner-Bitensky *et al.* 2009)[§]. This shows consistency with findings for other health professions. A number of interventions are *de facto* multiple interventions (e.g. educational outreach visits, local opinion leader, audit and feedback) as they may have various combinations embedded in their implementation.

2.4.4.1 Educational outreach

Educational outreach is a strategy where a trained person meets directly with providers in their practice settings to give information with the intent of changing the provider's practice. Information given may include feedback on the performance of the provider/s (Grimshaw, Thomas *et al.* 2004)*[§].

Academic detailing is used often as a synonym for educational outreach. This approach has been used extensively by the pharmaceutical industry targeting general practitioners and dentists to ensure that the use of the drug is rational and monitored correctly. The training of academic detailers includes effective communication and persuasion using various techniques (Soumerai and Avorn 1990). Educational outreach uses these in different combinations. What appears to distinguish it from other approaches are : persuasive techniques, targeted messages, face-to-face interactions, credible presenters and materials, and follow up or repeat visits. However, it tends to concentrate on prescribing behaviour and does not relate to the clinical skills of assessment and diagnosis (Soumerai and Avorn (1990), Bheekie 2001, Foy, Eccles *et al.* 2001).

A systematic review of educational outreach visits (EOV) and the effects on professional practice and health care outcomes reported a median adjusted risk difference in overall compliance with desired practice of 5.6%. The adjusted risk differences for prescribing were highly consistent for EOV but more varied for other types of professional performance (O'Brien, Rogers *et al.* 2007)* § .

There were eight trials (12 comparisons) where the intervention included an educational outreach visit and was compared to another intervention, usually audit and feedback. Interventions that included educational outreach visits appeared to be slightly superior to audit and feedback (O'Brien, Rogers *et al.* 2007). A similar comparison involving only two studies favoured educational outreach over audit and feedback in one and no difference in the other (Ivers, Jamtvedt *et al.* 2012)* § . When individual visits were compared to group visits the results were mixed (O'Brien, Rogers *et al.* 2007). A commentary on this systematic review suggested that despite there being so few included studies from developing countries, this was unlikely to affect the applicability of the findings in under-resourced settings. Further, it argued that because many health care professionals lack opportunities for self-learning and for attending professional meetings, educational outreach visits may be an effective way of providing new information to health care professionals in under-resourced settings (Costa and Khanna 2008).

There is an implicit assumption made in many studies that an improvement in health professional knowledge or behaviour will result in a change in patient outcomes. However, in only 14 of the included studies were patient outcomes reported and in very few were any patient level improvements reported, even if there was improvement in health professional practice. Hence O'Brien, Rogers *et al.* (2007) called for future studies to include patient outcomes as well as professional performance.

In the review of O'Brien, Rogers *et al.* (2007), the majority of studies was amongst physicians and one study had an obstetrician / senior midwife pair for

each of the study sites where educational outreach was given by a senior obstetrician on evidence for obstetric procedures. Most included studies were situated in high income countries (O'Brien, Rogers *et al.* 2007).

2.4.4.2 Audit and feedback

The quality assurance process for clinical practice includes a 5-step audit cycle – establish criteria of good quality of care, measure current practice, feed back findings and set targets, take action to change practice, and re-evaluate practice (Wagaarachchi, Graham *et al.* 2001). Audit and feedback is defined as “any summary of clinical performance of health care over a specified period” (Grimshaw, Thomas *et al.* 2004:8)* § . Audit can play various roles – monitoring and supporting clinically effective practice, as an educational tool to support improvement (Graham, Wagaarachchi *et al.* 2000), as well as contributing to developing better standards of care and improving quality in the health systems (Bhutta, Darmstadt *et al.* 2009). However, audit without action will not change practice thus it is necessary to ensure that any solutions are acted upon (Pattinson, Kerber *et al.* 2009)* § .

Single intervention studies using audit and feedback compared to no intervention report effects on professional practice ranging from small (Ivers, Jamtvedt *et al.* 2012)* § to modest (Grimshaw, Thomas *et al.* 2004). The effect on patient outcomes ranges from none (Lomas, Enkin *et al.* 1991)§ to moderate (Ivers, Jamtvedt *et al.* 2012) to a large effect of perinatal audit (Young, Hamilton *et al.* 2001, Pattinson, Kerber *et al.* 2009)§.

An overwhelming majority of studies focuses on physicians and is conducted in high income countries. However, Pattinson, Kerber *et al.* (2009), on reviewing perinatal audit and feedback in low-middle income countries, concluded that perinatal mortality audit has great potential to identify and address deadly delays and modifiable factors in peripartum care that lead to intrapartum stillbirths and intrapartum related neonatal deaths, as well as to maternal deaths. Further, it is critical that local solutions to address identified problems should be

implemented for the audit process to improve the quality of care and save lives (Pattinson, Kerber *et al.* 2009).

Comparing audit and feedback as a core essential feature of a combined intervention against various comparators yielded varied effects for professional practice but a moderate effect against usual care. When combined with other interventions, the effect size was larger than when audit and feedback was used alone. When compared directly with other implementation strategies these were equivocal for reminders and educational outreach (Ivers, Jamtvedt *et al.* 2012).

In single studies on audit and feedback the effect on patient outcomes demonstrated a large effect when combined with an educational intervention (to reduce postpartum haemorrhage and improve partograph recordings) (Mercer, Sevar *et al.* 2006)[§], whereas when compared with a local opinion leader to promote trial of scar and VBAC (Vaginal Birth After Caesarean) there was a large effect in favour of the local opinion leader (Lomas, Enkin *et al.* 1991).

Intensity of feedback and complexity of the targeted behaviour were identified as possible explanations for the variation in findings across studies (Jamtvedt, Young *et al.* 2007)^{*} as well as the role and level of engagement by health professionals for implementing the necessary changes (Jamtvedt, Young *et al.* 2006). Ivers, Jamtvedt *et al.* (2012) characterised the studies in this review according to a selection of variables which they considered important. In particular, the effect of the intervention appears to be larger when the baseline performance was low with recommended practice, when feedback was provided more than once and /or by a supervisor or senior colleague, when feedback was both written and verbal, when it aimed to decrease current behaviours, and when there were explicit targets and action plans. Only 56 of the 140 studies had explicit goals for change or action plans.

It has been suggested that not only the quality of data may be an effect modifier, but also the motivation and interest of recipients and the organisational support for quality improvement (Van der Veer, De Keizer *et al.* 2010).

Jamtvedt, Young *et al.* (2006) recognised this when calling for well designed process evaluations to be included in trials to explore and provide insights into the complex dynamics underlying the variable effectiveness of audit and feedback.

It seems that the approach of clinical audit and feedback is often that of seeking compliance. In addition, it may be done within a system where no other factors are modified to support positive change. Therefore, change is expected to occur as a result of health practitioner compliance or enthusiasm while other aspects of the system remain untransformed. Mancey-Jones and Brugha (1997) highlighted these dynamics when they stated that local internal audit is more likely to result in improvements in care if it is conducted in a structured and culturally sensitive way and everyone in the perinatal care process is involved in reviewing activities and in formulating recommendations so that there is a sense of ownership of the recommendations. This illustrates a distinction between a monitoring exercise and a practice improvement activity.

2.4.4.3 Local opinion leaders

Opinion leaders are those individuals who are regarded as educationally influential and who, by identifying the evidence underpinning best practice, may facilitate behaviour change (Doumit, Gattellari *et al.* 2007). They may use a variety of strategies common to academic detailing e.g. repeated contact, repetition of key messages and visually attractive presentation of information, which could influence the outcome of that intervention (Lomas, Enkin *et al.* 1991).

The review by Flodgren, Parmelli *et al.* (2011)* § included 18 cluster RCTs, four of which were in the reproductive health field. Most of the trials used a peer nomination process to identify educationally influential colleagues in their settings. The reviewers indicated that there was insufficient detail in the included studies to be able to identify such factors as the type of educational method used, the frequency and / or intensity of involvement.

Overall, the median adjusted risk difference was +0.12 representing a 12% absolute increase in compliance in the intervention group (Flodgren, Parmelli *et al.* 2011). The authors concluded that opinion leader interventions appear to improve performance but effectiveness is varied. This overall effect is similar to those of audit and feedback, distribution of educational materials, and multifaceted interventions including educational outreach. As has been found in other reviews, the majority of studies focussed on physicians (14 trials), with two trials focussing on nurses, and two including a combination of physicians, nurses and midwives.

Commentators on the 2007 systematic review of Doumit, Gattellari *et al.* (2007), expressed concerns which appear to be relevant to the current version (Flodgren, Parmelli *et al.* 2011). These were that there was no subgroup analysis according to setting or health field, and that in rural settings where health care providers work in relative isolation this intervention would not easily be applicable (Althabe 2007), and that it was not known if the effect of local leaders differs in developing country settings or varies according to health care profession or level of health care facility (Sguassero 2007, Souza 2007a). Sguassero indicated that before the intervention could be recommended for developing country settings, further research was needed on the best methods for selecting opinion leaders and the best ways of delivering opinion leader intervention at different health care levels in under-resourced settings (Sguassero 2007). Further questions raised related to the effectiveness of different formal (e.g. academic detailing, preceptorships, small-group teaching, lectures, seminars) and informal educational methods (e.g. one-to-one teaching, group discussions, informal consultations) when used as strategies for disseminating and implementing evidence-based practices, and how change in professional behaviour can be effected over time (Sguassero 2007).

2.4.4.4 Distribution of educational materials

This was defined as personally or mass delivered distribution of published or printed recommendations for clinical care. This could include clinical practice guidelines, audiovisual materials and electronic publication (Grimshaw, Thomas *et al.* 2004)* §.

An early study failed to support the efficacy of printed materials when used alone to improve medical decision making (Soumerai, Salem-Schatz *et al.* 1993). A more recent systematic review, where educational materials were the single intervention compared against a no-intervention control group, showed modest improvements in the process of care. However, there were indications that these changes may be short-lived (Grimshaw, Thomas *et al.* 2004).

Cheater, Baker *et al.* (2006)[§] conducted a cluster randomised trial specifically targeting nurses and Nursing practise in relation to managing urinary incontinence. Using a factorial design the nurses were allocated to one of four arms. The control arm consisted of educational materials only, but did not include evidence based recommendations on best practice. The three intervention arms added audit and feedback, educational outreach, or both. It appeared that printed educational materials alone might have been as effective as the other two strategies, but, in the absence of a true non-intervention arm, this was difficult to assess.

Kirshbaum (2008) [§] conducted a CRT to test if research-based educational materials (booklet) designed for breast care nurses could facilitate changes in knowledge, reported practise, and attitude outcomes regarding the benefits of exercise for breast cancer patients. It also addressed issues of motivation and confidence for change. This study showed a statistically significant improvement in favour of the intervention arm in ten out of the 17 knowledge questions, of reported practice where exercise would be advised in three out of 12 items, and in one out of the nine attitude questions. Kirshbaum (2008) concluded that providing nurses with educational material with relevant research based messages can be effective in transmitting research evidence to specific groups of nurses, and changing reported practice and related attitudes.

2.4.4.5 Continuing professional education

There is increasing emphasis being placed on continuing professional education, and in a number of countries this is linked to the ongoing license to practise. This can take a number of forms and delivery methods. Commonly, this takes the form of educational meetings or workshops, but other methods which have been developed and investigated in South Africa are presented, e.g. perinatal education programme, training-of-trainers and face-to-face training.

2.4.4.5.1 Continuing education meetings and workshops

A systematic review was undertaken to assess the effects of educational meetings on professional practice and health care outcomes (Forsetlund, Bjorndal *et al.* 2009)*[§]. Thirty-two trials reported multifaceted interventions which included reminders, patient education materials, supportive services, feedback reports, and educational outreach. Fourteen (17%) studies explicitly stated that the intervention was based on a known theory of behaviour change, learning theory or a diffusion of innovations theory (Forsetlund, Bjorndal *et al.* 2009). The effect appears to be larger with higher attendance at the educational meetings and with mixed interactive and didactic educational meetings. The effect for professional practice was greater than for patient treatment outcomes. Educational meetings did not appear to be effective for complex behaviours and they appeared to be less effective for less serious outcomes (Forsetlund, Bjorndal *et al.* 2009).

A commentary on this systematic review highlights the difficulty in generalising the findings to developing country contexts where financial and human resources required to conduct continuing education may be lacking and severe staff shortages (with the accompanying high staff turnover) mean that interventions aimed at changing clinical practice are difficult to sustain without continual external input (Smith, Brown *et al.* 2009).

One study targeting maternity care was selected as an example (Gülmezoglu, Langer *et al.* 2007). This was a cluster randomised trial of an active multifaceted intervention, based on the WHO Reproductive Health Library, to improve

obstetric practices and was undertaken in Thailand and Mexico. The intervention consisted of three interactive workshops over a period of six months, and ten targeted practices (which were not disclosed to the participants) which were measured starting four to six months after the third workshop. Participants included all staff (doctors, midwives, interns and students) although attendance varied due to staff turnover and participation from other departments. The awareness of the Reproductive Health Library increased substantially in both regions following the intervention. There was no consistency in the uptake of targeted practices, but in the Mexican sites there was an increase in the use of antibiotics for caesarean section and in the Thailand sites, a reduction in the routine use of episiotomy. Four practices showed no change (labour companionship, use of magnesium sulphate for eclampsia, use of corticosteroids for births before 34 weeks gestation, and vacuum extraction) and the remaining four could not be measured reliably (iron / folate supplementation, use of uterotonics after birth, breastfeeding on demand, and external cephalic version). The authors reflected that by focussing on knowledge access rather than on one or two interventions, this may have decreased the chances of a positive impact. This underscores the fact that information or knowledge in itself is not sufficient to produce behaviour change. Some degree of mediation is required.

2.4.4.5.2 *Perinatal Education Programme*

The Perinatal Education Programme (PEP) is a facilitated study programme for midwives, obstetricians and paediatricians designed to be undertaken in study groups. The co-ordination is done by a volunteer within the group. The programme has been developed locally by a team of midwives, paediatricians, and obstetricians (Woods and Theron 1994), based on the work of others (Kattwinkel, Cook *et al.* 1979 , Hesketh, Zhu *et al.* 1994). The PEP is divided into various manuals, *inter alia*, Maternal Care, and Newborn Care. Each chapter is set out in a question and answer format and has associated clinical skills workshops. A multiple choice question test is completed at the beginning and end of the course, as well as an OSCE (Objective Structured Clinical Evaluation) test for clinical skills. A pass mark of 80% is required to obtain

the PEP certificate, due to the fact that the participants have seen all the questions previously.

The effect of introducing the manual to different primary level sites, particularly for skills training, was tested prospectively. The practical skills of midwives improved significantly when following the instructions of the manual closely, regardless of their level of experience, and additional teaching was unnecessary (Theron 1997). The advantages of this programme are that it is an accurate set of information, can be used in various settings, does not require the practitioner to relocate to an educational centre (which could cause personal and workplace-related disruptions), and facilitates access to educational material for those practitioners who are distant from an educational centre. There is minimal cost and bureaucracy attached to the programme (Woods and Theron 1995). However, there are limitations in terms of knowledge transfer, cognitive skills development, lack of skilled educational facilitation, and the need for institutional support.

There is an assumption that improved cognitive knowledge will result in changed practice and improved outcomes. Further, there is an assumption that this learning style is one where there will be knowledge transfer into different contexts. This is confirmed by the conclusion of Haynes *et al.* (1984) cited by (Soumerai and Avorn 1990) who state that one failure of traditional continuing medical education has been the assumption that the transmission of rational information alone, independent of how it is provided, will predictably result in improved clinical decisions.

Bloom, in collaboration with college and university examiners, developed a taxonomy of cognitive skills which identified six categories, *viz.* knowledge, comprehension, application, analysis, synthesis and evaluation (Bloom 1956). Increasing complexity requires mastery of the simpler behaviours so that remembering (knowledge of) and understanding a literal message

(comprehension) are necessary for application, analysis, synthesis and evaluation to take place. The PEP programme tends to address knowledge and comprehension, whereas the other levels are necessary for quality clinical practice.

There has been no long-term evaluation published on the PEP programme in terms of the retention of knowledge and skills, neither has there been an assessment of whether or not this actually contributes to reduced maternal and perinatal mortality and morbidity. A small controlled study with a before-and-after design was conducted in a rural district in Mpumalanga¹². Midwives participating in the PEP programme, as well as patient records, were evaluated before and after the implementation of the PEP programme. While completion of the obstetric manual of the PEP improved the knowledge of the midwives, no alteration in practice was detected (le Roux, Pattinson *et al.* 1998).

While the PEP programme facilitates learning on-site (at the workplace) one should consider if there is any benefit in centralising continuing education programmes. In a project to implement Kangaroo Mother Care in 26 hospitals in Mpumalanga Province between 2004 and 2006, a comparison was done between those hospitals which had on-site facilitation and those which had off-site facilitation. The researchers concluded that there was no significant difference in effectiveness between the two sites of facilitation, although they conceded that more hospitals would need to be covered before conclusive evidence could be established (Davis, Bergh *et al.* 2006).

2.4.4.5.3 *Training-of-trainers and Face-to-face training approaches*

Training-of-trainers has become a popular strategy to scale-up the training of large groups of nurses in various aspects of clinical management and implementation of clinical guidelines. This also speaks to the dynamic of on-site or centralised location of training. While this cascaded approach is attractive to policy makers and health service managers, the capacity of the trainers requires both content knowledge of the subject at hand and related clinical competence, as well as confident and skillful educational delivery.

¹² A largely rural and poor province in the north-east of South Africa

A pragmatic cluster randomised trial was conducted in the Free State¹³ (Fairall, Zwarenstein *et al.* 2005). The intervention consisted of a five-day educational outreach programme covering the content of the syndromic management of respiratory disease in adults and of opportunistic infections associated with HIV, as well as the principles of interactive educational outreach. The participants were eight senior nurses who were responsible for the TB programme. The intention was for them to conduct four to five sessions of three hours each to all staff in the clinics over a period of three months in each of the study sites to which they were assigned. They managed to conduct a median of two in-service training sessions at clinics, which fell short of the intended cover, due to logistical constraints¹⁴. Of the four primary outcomes measured, two (case detection of TB and prescription of inhalational corticosteroids) had statistically significant changes in the intervention arm, while the other two (sputum screening for tuberculosis and antibiotic prescription) did not. In addition, referral of severely ill patients to a doctor was higher in the intervention arm (OR 2.59, 95% CI: 1.06 to 6.19, $p=0.037$). The educational outreach programme in syndromic management of respiratory disease in adults achieved large improvements in the quality of tuberculosis and asthma care within the existing staff complement and without interruption to the service (Fairall, Zwarenstein *et al.* 2005).

A different approach was taken in the Western Cape Province¹⁵ when a roll-out of the WHO's Basic Antenatal Care Programme (BANC) was planned. The WHO antenatal care trial used the training-of-trainers method (Villar, Ba'aqeel *et al.* 2001). There was concern that this could result in a diluted effect depending on the confidence of the trainers. Resistance was encountered by some senior clinic staff who did not feel confident to deliver a teaching programme and who did not necessarily have mastery of the field of antenatal care (unlike their Free State Province counterparts in the Fairall study who had

¹³ A province in central South Africa with high rates of tuberculosis and HIV infection.

¹⁴ This was well above the target level of 5% off-site in-service training in the province.

¹⁵ A well resourced province in the south-west of South Africa.

specific responsibility for the TB programme and, presumably, mastery of the content in which they were required to provide training).

A restricted cluster randomised trial compared the training-of-trainers (TOT) with face-to-face (F2F) approaches (Groenewald-Neethling 2010). The content training consisted of 15 hours, was similar for both arms and was provided by a skilled trainer with mastery of the content. The TOT arm had an additional three hours of training in the approaches to teaching and had teaching material supplied. In turn, these trainers were required to provide three hours of training per week for five consecutive weeks to their colleagues. An audit of the antenatal records according to the BANC standard was done, using a 24 point checklist. Both arms demonstrated statistically significant improvements between the pre-intervention and post-intervention scores for the total score (TOT mean difference 0.84, $p=0.001$; F2F mean difference 5.08, $p<0.0001$), but the difference between the post-intervention scores of the two arms showed a mean difference of 3.32 ($p<0.0001$) in favour of the F2F arm (Groenewald-Neethling 2010).

There is a general assumption that the training-of-trainers approach is more cost-effective. Groenewald-Neethling undertook a simple costing analysis of staff time and travel costs. This determined that the F2F method (65 nurses trained directly) was almost half the cost of the TOT method (15 nurses trained directly and 52 nurses had cascaded training, i.e. 67 nurses) (F2F = R10 602.47 versus TOT = R20 903.73¹⁶) (Groenewald-Neethling 2010). Thus, the face-to-face approach was more effective in impacting practice and was more cost-effective.

In contrast to much of the available literature, both these methods have been applied successfully in a middle-income country in rural settings among nurses.

¹⁶ US\$ 1 = SA Rand 10.8 (July 2014)

2.4.4.5.4 Critical appraisal

One should look not only at learning and information transfer, which to some extent depend on outside agency, but at the ability of practitioners to decide on the value of emerging research for their own practice. A systematic review was conducted on teaching critical appraisal skills in health care settings, concentrating on professionals already in practice (Horsley, Hyde *et al.* 2011)*. The authors defined critical appraisal as “the process of assessing and interpreting evidence by systematically considering its validity, results and relevance to an individual’s work” (Horsely and Hyde *et al.* 2011:1).

An extensive search identified 148 studies of potential relevance to the review, but ultimately only three RCT studies (with a total of 160 participants) were included. None of the studies evaluated process of care or patient outcomes and only one study was explicit about using an adult learning approach. All three studies reported critical appraisal related outcomes, two of which reported significant improvements in critical appraisal skills. The authors concluded that low intensity critical appraisal teaching interventions in health care populations may result in modest gains. Horsley, Hyde *et al.* (2011) called for focussed and rigorous randomised trials to be conducted in this area, and specifically for interventions to use appropriate adult learning theories (Horsley, Hyde *et al.* 2011). Although no patient outcomes were evaluated, unless the scale of such an intervention programme is very large and sustained over a number of years, it might not be feasible to measure this directly.

2.4.4.6 Package of interventions

Practice improvement interventions usually go beyond a single limited intervention and usually include some type of system supported change. A variety of components made up the packages in the studies that were identified for this literature review.

A study undertaken in three Australian hospitals which sought to reform the way maternity care (specifically postnatal care) was organised, included an extensive set of maternity enhancement strategies as part of the intervention (Yelland, Krastev *et al.* 2007). The package of strategies included the

involvement of the intended recipients of the care as well as the midwives who were to deliver the care, capacity development of the midwives, setting up teams across intra-partum and post-partum settings to enhance continuity of care, health service change, development of evidence based practice guidelines and protocols, and consumer written information for the postnatal period. This package was supported by the state government-funded maternity services enhancement strategy (Yelland, Krastev *et al.* 2007). The before-and-after study involving 2306 respondents (1256 'before' (response rate 68.7%); 1050 'after' (58.5% response rate)) showed modest but significant improvement in the post-intervention group with respect to overall postnatal hospital care (OR 1.33, 95% CI 1.11 to 1.59), the level of advice and support received in relation to discharge and going home (OR 1.45, 95% CI 1.17 to 1.79), the sensitivity of caregivers (OR 1.26, 95% CI 1.05 to 1.52), and the proportion of women receiving domiciliary care after discharge (OR 4.09, 95% CI 3.18 to 5.27). The findings indicated that the package of care, which included a commitment to continuity and an individualised approach, improved women's experiences of care (Yelland, Krastev *et al.* 2007).

The Haxton and Fahy study introduced an intervention package consisting of developing new clinical midwifery pathways and clinical governance for midwives, and providing the necessary training and support for the advanced practice midwives responsible for implementing the new clinical pathways. The study evaluated the changes with respect to effectiveness, cost and stakeholder satisfaction levels (Haxton and Fahy 2009). A baseline and follow-up audit revealed that the average patient stay dropped by 49% ($p < 0.001$), the amount of midwifery time saved was 186 hours over the six months of the project (which equated to one shift per week) and medical time saved was 90 hours, patient satisfaction increased by 10% overall and staff satisfaction was positive.

In a cluster randomised trial conducted in 19 hospitals in Uruguay and Argentina, a multifaceted behavioural intervention (selection of opinion leaders, interactive workshops, training of manual skills, one-on-one academic detailing

visits with hospital birth attendants, reminders, and monthly feedback) was used to develop and implement guidelines on episiotomy and management of the third stage of labour, compared with no intervention (Althabe, Buekens *et al.* 2008). Primary outcomes measured after 18 months of the intervention were the rates of prophylactic use of oxytocin during the third stage of labour (increased from 2.1% to 83.6% in the intervention arm and 2.6% to 12.3% in the control arm ($p<0.01$)), and of episiotomy (decreased from 41.1% to 29.9% in the intervention arm and remained stable in the control arm ($p<0.001$)). These outcomes were sustained 12 months after the end of the intervention. Birth attendants' readiness to change increased in the hospitals receiving the intervention, with the absolute difference between the two groups for active management of the third stage being 38.4% (95% CI, 19.6 to 56.9; $p<0.001$), and for selective episiotomy an absolute difference of 34.5% between the rate changes (95% CI, 27.2 to 58.0; $p<0.001$). Secondary outcomes were postpartum haemorrhage (relative rate reduction, 45%; 95% CI: 9 to 71 for 500ml; and a relative rate reduction, 70%; 95% CI: 16 to 78 for 1000ml) and birth attendants' readiness to change their behaviour with regard to these two clinical management practices (Althabe, Buekens *et al.* 2008).

Nicholl and Cattell's study on reducing perineal trauma involved three PDSA cycles :

- ◆ cycle one focussed on allowing the woman to push when she felt the urge and not immediately after the second stage was commenced
- ◆ cycle two focussed on alternative positions for birthing specifically side-lying
- ◆ cycle three included a preference for vacuum extraction over forceps deliveries when assistance was required, so long as there was no contra-indication (Nicholl and Cattell 2006).

At the end of the six-month project timeframe, fourth degree tears had reduced by 100% (and third degree tears did not increase), intact perineae rose from 27.4% to 33.3%, and the episiotomy rate fell from 7.4% to 6.0%. This perineal

trauma study was not powered for formal statistical analysis so only trends could be identified. However, the study shows that such an exercise can be conducted with the existing resources to trial small changes within a short timeframe (Nicholl and Cattell 2006).

Sprague, Oppenheimer *et al.* (2008) conducted a pre- and post-evaluation study of clinical outcomes and knowledge associated with implementing a second stage of labour guideline in term nulliparous women with epidural analgesia, in two birthing units of a large teaching hospital. The main recommendation was to wait for up to two hours after full dilatation before pushing. A suite of strategies was employed which included recruiting five to six champions at each site, holding a project launch, a keychain reminder token¹⁷, intensive education sessions in multiple formats, documentation for second stage, displaying second stage guidelines in each labour room, and giving monthly feedback of progress. Site 1 showed a significant increase in median waiting time before pushing (33 minutes, $p=0.04$) while site 2 showed no change in median waiting time. Both sites showed significant improvement on the assessment of position of the presentation at full cervical dilatation. There were no differences in the duration of labour or operative delivery rates (Sprague, Oppenheimer *et al.* 2008).

A review of strategies for delivering interventions to reduce the global burden of stillbirths concludes *inter alia* that packages of interventions which are appropriate for the setting should include accessibility to care and health system resources as well as provider skill (Bhutta, Darmstadt *et al.* 2009).

2.4.5 Standardisation

In the context of clinical practise the concept of standardisation usually refers to clinical guidelines (based on good evidence) for clinical management, agreed procedures, and development and maintenance of an agreed level of

¹⁷ This is not clarified in the text. Other studies mention tokens or reminders used to promote the project, e.g. note pads, pens, flowcharts. It is assumed that this keychain token serves the same purpose.

competence of required skills. This concept has resonance in the risk reduction / patient safety model of thinking.

Grimshaw, Thomas *et al.* (2004) address the implementation of guidelines in their systematic review, as have other authors for single studies (Althabe, Buekens *et al.* 2008, Sprague, Oppenheimer *et al.* 2008, Haxton and Fahy 2009).

Consistency might be regarded as an aspect of standardisation. In the study by Scott, VandenBeld *et al.* (2011), uncertainty related to variability in management and leadership styles was identified. When such styles cause practitioners to be ineffective there is a need to manage this diversity constructively.

2.4.6 Evaluation

Evaluation of the process informs further improvements that may be necessary and can serve to demonstrate accountability to the system in which the service functions. This step, which is common to the Knowledge-to-Action framework and the Model for Improvement (Graham, Logan *et al.* 2006, Langley, Nolan *et al.* 2009), may become the beginning of a new cycle of development and improvement (Nicholl and Cattell 2006).

2.4.7 Health System issues

Recognising that change vested in an individual has limitations (Gülmezoglu, Villar *et al.* 2004), literature relating to system change and/or support, the nature of professional practise contexts, research utilisation by nurses, barriers to change, organisational culture, integrated care pathways and health system level interventions will be presented.

2.4.7.1 System change / support

The intervention package in the Yelland study (described in 2.4.4.6) required changes in the way the maternity care was delivered (Yelland, Krastev *et al.* 2007). This had implications for institutions within a health service network and required substantial support from managers to effect the envisaged change. The authors noted that this intervention had been a significant undertaking for

the health service and the study represented one of the “largest evaluations of system-wide change to maternity care in Australia” (Yelland, Krastev *et al.* 2007:399).

The Haxton and Fahy study (described in 2.4.4.6) was undertaken in a setting where professional resources could not be increased. By changing clinical governance, an estimated 371 hours 36 minutes was saved in waiting hours during the six months of the project, and 186 hours of midwifery time (which equates to a saving of one shift). The estimated cost saving in paid midwifery time over the six months was AUS\$ 7 020.00. In addition, approximately 90 hours of medical time was saved during the project (Haxton and Fahy 2009). The project demonstrated that it was a safe, efficient and cost-effective model of care, and was supported by senior clinicians and administrators in Midwifery and Medicine, as well as the area health service.

2.4.7.2 Nature of the professional practice context

One of the inclusion criteria in the Grimshaw, Thomas *et al.* (2004) review was that participants should be medically qualified health care personnel. However, the majority of the studies related to medical doctors alone (74%) and very few to nurses, pharmacists, dieticians, or other health professionals. The nature of the disciplines and the way they are organised could have a significantly different effect on the outcomes measured, but this is not explored. For example, medical doctors are more likely to be targeted to change behaviour for general management of a problem, drug prescribing, test ordering or other clinical regimens, whereas nurses may be monitoring a response to a particular health challenge or care intervention. The activities that flow from these different professional modalities will influence the impact of the intervention (Cheater, Baker *et al.* 2006, Kirshbaum 2008).

Petzold, Korner-Bitkensky *et al.* (2010) applied the knowledge-to-action model to address gaps in practice by rehabilitation therapists in relation to **unilateral spatial neglect** (USN) in people who had had a stroke. This included a chart audit, a national survey of rehabilitation therapists working on their use of best

practice assessment and intervention, focus group discussions on the appropriateness and effectiveness of USN assessment tools in various contexts, interviews with researchers, focus group discussion with clinicians to identify the barriers and facilitators to evidence based practice for management of post-stroke USN, and a follow-up study on the effectiveness of KT interventions. The findings suggested the need for a multimodal intervention approach, similar to the conclusions of Grimshaw, Thomas *et al.* (2004), where physicians predominated.

Petzold, Korner-Bitkensky *et al.* (2010) report on the creation of a knowledge synthesis website and the development of clinician learning tools, interactive e-learning modules, and an interactive quiz related to best practice assessment and intervention to assist the access and uptake of the available information. Further, a standardised tool, the PERFECT (Professional Evaluation and Reflection on Change Tool), which elicits information on practice change and the barriers and facilitators that incite change, was created. The use of this tool aims to foster reflection by clinical teams who wish to focus on practice change. These authors demonstrated how the knowledge-to-action framework can be used to address knowledge translation systematically. However, they do concede that this took several years and considerable funding (over \$1 million) to achieve, which would be out of reach for clinical managers (Petzold, Korner-Bitkensky *et al.* 2010).

A qualitative study in northern Vietnam identified the methods by which practitioners acquire new knowledge (Eriksson, Nga *et al.* 2011). This highlights the importance of three cornerstones for successful change of clinical practice: evidence, context, and facilitation (Eriksson, Nga *et al.* 2011). Given that the available evidence for newborn healthcare is strong, the study focussed on assessing the effectiveness of facilitation in a Vietnamese context, where traditional medicine is incorporated into the health system and where the practices might be in conflict with national evidence-based guidelines. Community health centre staff (doctors, assistant doctors, midwives and

nurses working in neonatal care in three diverse geographic areas were purposively sampled and six focus groups of seven to eight participants were formed.

Healthcare staff used a number of channels for acquiring knowledge but, despite the availability of national guidelines, these were not well utilised, owing, in part, to poor introduction. Participants preferred formal training and expressed interest in interacting with colleagues at higher levels. Support from these colleagues was regarded as necessary to implement change at a community level. Concerns were expressed about the lack of opportunity for building confidence in their skills in the absence of training when new equipment was provided, and when there were very low numbers of births (Eriksson, Nga *et al.* 2011). These perspectives were different from those described in systematic reviews.

A study, which developed an in-service education programme focussing on patient centredness, problem solving and critical reflection for primary providers (registered nurses) delivering care to patients with tuberculosis (TB) in South Africa, highlighted the need for a supportive environment where health care providers were encouraged to reflect critically on their work and plan for change (Dick, Lewin *et al.* 2004). Participants reported that the insights gained had impacted the way they interacted, not only with clients, but also with colleagues and family members. Involving the clinic managers in the education process was important for teambuilding and the implementation of change plans.

2.4.7.3 Research utilisation by nurses

A systematic review of the determinants of research utilisation by nurses included 20 studies of various research designs. Findings were presented as six themes and the results were equivocal (Estabrooks, Floyd *et al.* 2003)*. There was support for a link between beliefs and attitudes and research utilisation, but the extent to which beliefs and attitudes, involvement in research activities, information-seeking, education, professional characteristics, and other socio-economical factors influence research use remains largely unknown. The

authors suggested that this set of potential determinants was likely to be highly intercorrelated, but the study designs used typically did not allow for careful analysis of intercorrelations among variables (Estabrooks, Floyd *et al.* 2003).

Scott, VandenBeld *et al.* (2011), in their ethnographic study in a 16 bed paediatric critical care unit, found that a context of uncertainty in the work environment hindered nurse research utilisation. Sources of uncertainty related to patient acuity, unpredictability of work flow, the complexity of teamwork and the inconsistency in management and leadership styles. This feeling of uncertainty significantly decreased the nurses' ability to become active decision makers in their patient's care. In turn, because nurses did not identify themselves as clinical decision makers, they did not prioritise the accessing of evidence-based knowledge to use in their practice. These findings suggest that strategies aimed at reducing uncertainty may facilitate efforts to enhance knowledge translation by increasing practitioner receptivity to change. This reflects the approach of Prochaska and DiClemente (1982) and Rollnick *et al.* who address readiness for health behaviour change and promote motivational interviewing (Rollnick, Butler *et al.* 1997, Rollnick, Mason *et al.* 1999).

In a study of 590 primary care / general practice nurses in Australia, the main sources of evidence for practice were from in-service education and training opportunities, engagement in experiential learning and interactions with clients, peers, medical practitioners and specialist nurses (Mills, Field *et al.* 2011). Perceived lack of time to access knowledge for practice and low self-assessment of participants' skills in sourcing and translating evidence into practice were cited as barriers to using current evidence in practice. Mills, Field *et al.* (2011) constructed a model to explain the domains that influence the ability of general practice nurses to effect knowledge translation *viz.* skills in finding, reviewing and using evidence; barriers to finding and reviewing evidence; knowledge from published sources; knowledge from other sources; and barriers and facilitators to change.

These studies show that there is a capacity gap amongst nurses which prevents them from engaging with research and implementing evidence-based findings into their practise.

2.4.7.4 Barriers to change

A systematic review on interventions tailored to overcome identified barriers to change attempted to understand the dynamics that hinder (or support) change, be they at personal or health system levels (Baker, Camosso-Stefinovic *et al.* 2010)* §. RCTs were selected where at least one group received an intervention which was designed to address prospectively-designed barriers to change, using EPOC criteria *viz.* information management, clinical uncertainty, sense of competence, perceptions of liability, patient expectations, standards of practice, financial disincentives, administrative constraints, and a category named 'other' (Grimshaw, McAuley *et al.* 2003).

Twelve studies yielded a pooled Odds Ratio of 1.52 (95% CI, 1.27 to 1.82, $p < 0.001$) (classical analysis). There were wide variations between and within studies, some of which might be attributed to the variety of barriers identified and addressed, and the lack of consistency in the methods used within the tailored interventions. The authors concluded that interventions tailored to prospectively identified barriers are more likely to improve professional practice than no intervention or dissemination of guidelines (Baker, Camosso-Stefinovic *et al.* 2010).

As has been found in other similar reviews, most studies (23/26) related to physicians (and/or pharmacists). Two studies targeting nurses investigated labour support and electronic fetal monitoring in secondary and tertiary settings (Davies, Hodnett *et al.* 2002) and effectiveness of audit and feedback on Nursing practise and patient outcomes (Cheater, Baker *et al.* 2006).

Cochrane, Olson *et al.* (2007) conducted a systematic review to identify and assess barriers to optimal clinical practice. The review highlights the complex and multiple barriers between what the evidence shows and what is put into

practice. The studies included in the review were placed into seven categories which were differentiated according to the health care provider, clinical guidelines, scientific evidence, the patient, and the health system. The studies were analysed thematically and the results expand the understanding of barriers to knowledge translation, evidence implementation, and diffusion of innovation. Cochrane, Olson *et al.* (2007) highlight increasing numbers of behavioural and system barriers in the literature. In turn these barriers can impact on the provision and achievement of optimal health care. Although the review is dominated by physician-based studies, the recognition of different types of barriers across the provision of care system would enable other professional groups to apply this framework to their context.

Bingham and Main (2010) categorised potential implementation barriers as follows (using examples from studies presented elsewhere in this review):

- ◆ leader barriers, e.g. lack of knowledge / capacity, e.g. Scott, VandenBeld *et al.* (2011)(2.4.5 and 2.4.7.3)
- ◆ clinician barriers, e.g. unwilling to change, e.g. Sprague, Oppenheimer *et al.* (2008) (2.4.4.6)
- ◆ characteristics of the project, e.g. may require work flow adjustments, e.g. Yelland, Krastev *et al.* (2007) (2.4.4.6)
- ◆ implementation climate, e.g. availability of organisational support and required resources, e.g. Gülmezoglu, Langer *et al.* (2007) (2.4.4.5.1).

Once these are identified, an implementation plan can be planned systematically choosing an appropriate tactic from three broad types of interrelated strategies, e.g. discourse (communication) tactic – change champion; education (formal or informal) tactic – journal club; data tactic – chart audit (Bingham and Main 2010). Given that the quality improvement literature is often imprecise in describing interventions, this typology along with that of Michie, van Stralen *et al.* (2011) could facilitate understanding and comparison.

The methods used to identify barriers and tailor interventions to address them need further development and studies are required to compare the use of

different theories or the use of theory versus no theory (Baker, Camosso-Stefinovic *et al.* 2010).

2.4.7.5 Organisational culture

Parmelli, Flodgren *et al.* (2011)* conducted a systematic review on the effectiveness of change strategies in changing organisational culture to improve health care performance. They extended their search beyond randomised controlled trials to include well designed quasi-experimental studies, controlled clinical trials, controlled before-and-after studies and interrupted time series analysis. Studies had to be set in health care organisations where the aim was improved patient outcomes as a result of planned change in organisational culture. The main outcomes were to be objective measures of patient outcome and professional performance. Of 4239 studies identified, only 13 were potentially eligible for review and a further three were added. However, none met the Cochrane EPOC group quality criteria for inclusion in the systematic review. The authors concluded that high quality evidence on the effectiveness of strategies to change organisational culture is lacking, and that future research should concentrate on clearly defining what is meant by organisational culture and conducting evaluations with sufficiently robust designs in order to produce generalisable findings (Parmelli, Flodgren *et al.* 2011).

The work of Mercer, Sevar *et al.* (2006) in a small hospital on the remote North India / Tibet border illustrates how developing a positive organisational culture (full involvement, strong commitment and teamwork) towards audit and feedback can achieve remarkable improvements in intrapartum care (50% reduction in postpartum haemorrhage; substantial improvement in clinically useful partograph recordings) which was sustained over a seven year period despite minimal infrastructure.

While specific aspects of patient treatment can be clearly defined and measured, the broader health system based approach (in this case, organisational culture) to improvement in practice appears to be less well described. These studies have shown equivocal results, some of which may be

attributed to methodological limitations. Scott, Mannion *et al.* (2003), in a review on the correlation between organisational culture and health care performance, concluded that organisational culture may be a relevant factor in health care performance, yet describing the nature of that relationship was challenging.

Given that 'organisational culture' is a sociological or anthropological construct, there is some room for multi-method and multi-paradigmatic approaches to understanding the role that this plays in professional behaviour and practice, and how this may impact patient care. Mannion, Davies *et al.* (2005) used a multiple case study design to compare the cultural characteristics of hospitals which were regarded as high or low performers according to a star performance rating system. Mannion, Davies *et al.* (2005) found that different cultural patterns could be identified within groups with similar performance levels suggesting considerable cultural divergence. Main areas of divergence were grouped in the following areas - leadership and management orientation, accountability and information systems, human resources policies, and relationships within the local health economy. However, the authors indicated that the results should be treated with a degree of caution due to methodological considerations (Mannion, Davies *et al.* 2005).

A later study by Mannion, Konteh *et al.* (2009) Mannion, Konteh *et al.* (2009) was a national survey in 275 English National Health Service (NHS) organisations which aimed to assess organisational culture for quality and safety improvements. The focus was on improving safety / reducing risk rather than on holistic assessment of the dimensions of health care quality and performance (Mannion, Konteh *et al.* 2009), e.g. understanding the values, beliefs, meanings, relationships and processes etc., that would typify an organisation's culture.

Recognising that nurses and midwives form the major portion of the health workforce, and that the quality of care should improve by applying the best available evidence, Flodgren, Rojas-Reyes *et al.* (2012) conducted a systematic review to assess the effectiveness of organisational infrastructures in promoting

evidence-based Nursing. Organisational infrastructure was defined as the underlying foundation or basic framework through which clinical care is delivered and supported. This could take many forms, *inter alia* organisational policies, management frameworks / shared governance, skill mix, nurse development units, clinical supervision programmes. Despite expanding the net for the search by identifying alternative sources ('grey literature'), and including a wide range of research methods, only one very low quality study was included in the review. This study showed no evidence of an intervention effect three months after adopting a set of evidence-based clinical guidelines for the prevention of pressure ulcers (Shih, Aye *et al.* 2010). Of the 15 studies excluded from the systematic review, none related to midwives or reproductive health. Flodgren, Rojas-Reyes *et al.* (2012) concluded that if evidence-based Nursing is to be promoted successfully at an organisational level, policy makers and health care organisations must ensure the funding and conduct of well-designed studies to generate the evidence required to guide policy.

Penn-Kekana, McPake *et al.* (2007) highlight that programmes aimed at improving maternal health are not only technical but are also social interventions. These interventions need to be evaluated as such, using methodologies that have been developed for evaluating complex social interventions whose aim is to bring about change. Using data from research conducted in Russia, Bangladesh, Uganda and South Africa, Penn-Kekana, McPake *et al.* (2007) present a simple 'dynamic responses' model which shows that the reflexive, complex and dynamic responses of health workers and community members to policies and programmes are the key to understanding challenges in implementation. This has relevance at the health service and health system levels.

2.4.7.6 Integrated care pathways (ICP)

An integrated care pathway is a complex intervention for collective decision making and organisation of care processes for a clearly defined group of patients during a clearly defined period. A care pathway aims to enhance the quality of care across the continuum by addressing risk-adjusted patient

outcomes, patient safety and satisfaction, and optimal use of resources (European Pathway Association 2007). Such an approach might be useful in managing intrapartum care and clinical guidelines would form part of an ICP.

A systematic review by Allen, Gillen *et al.* (2009)* aimed to evaluate ICP use in child and adult populations in the full range of healthcare settings. Nine papers reporting on seven studies were included. Due to heterogeneity of the studies, neither meta-analysis nor qualitative synthesis was possible. Five studies included nurses in the implementation but it was not clear what involvement they had in the development of the ICPs. None of the studies included described explicitly if the development and implementation processes, the tool's content, its use in practice, or a combination of any of these elements was responsible for the change in behaviour of the health service providers.

Allen, Gillen *et al.* (2009) recommended that use of ICPs should be restricted to those areas where there are clearly identified deficiencies in existing care provision and / or where change is required, and be based on best practice. Best practice guidelines should be made available to staff in a form that is usable in daily practice, and staff should be supported in exercising professional judgment in those cases when adherence to the pathway is not in the individual patient's interest. Prior to ICP development, the developers should determine if the ICP is appropriate for the patient population, specify how they wish to change practice and which mechanisms are necessary to do this (Allen, Gillen *et al.* 2009).

As for many of the systematic reviews included, all studies were from high income countries.

2.4.7.7 Health system level interventions

While many innovations are conducted locally, there is benefit in strengthening the whole health system to provide improved care to women and infants. Such an innovation, the so-named **10 + 5 + 1** package, aims to reduce not only maternal and neonatal mortality but also the number of stillbirths (estimated at 2.65 million per year worldwide), which often go unreported

(Pattinson, Kerber *et al.* 2011). This package consists of :

- ◆ **10** interventions during pregnancy and childbirth (e.g. detection and management of fetal growth restriction) which could prevent 45% of stillbirths in 68 countries listed as priorities in the 'Countdown to 2015' report, and in which 92% of the world's stillbirths occurred in 2008 (Bhutta, Chopra *et al.* 2010, Bhutta, Yakoob *et al.* 2011),
- ◆ **5** interventions which would not necessarily reduce stillbirths but would result in substantial benefits for mothers and infants, e.g. active management of the third stage of labour
- ◆ **1** intervention, i.e. contraceptive cover (Pattinson, Kerber *et al.* 2011).

Interventions can be packaged and provided across the continuum of care, and can be integrated across different levels and sites of care to maximise cost-effectiveness. To function effectively the system requires skilled health workers to provide care that is accessible, and supported by protocol-based referral up to the level of comprehensive emergency obstetric care and advanced neonatal care. In addition, community based initiatives should address health-based practices and education regarding appropriate care seeking for danger signs during pregnancy (Pattinson, Kerber *et al.* 2011). It is estimated that if full coverage of care (99% coverage of the '10 + 5' interventions) was reached in 2015 up to 1.1 million (45%) third trimester stillbirths, 201 000 (54%) maternal deaths and 1.4 million (43%) neonatal deaths could be prevented per year, at an additional cost per person of US\$ 2.32 and a total package cost of \$10.9 billion in 68 Countdown countries (Pattinson, Kerber *et al.* 2011).

Strengthening the health system has also been integrated into the most recent recommendations of the Saving Mothers report which has identified five aspects **(5H)** that need to be addressed to reduce maternal deaths in South Africa, *viz.*

- ◆ **HIV and AIDS**
- ◆ **Haemorrhage**
- ◆ **Hypertension**
- ◆ **Health worker training**
- ◆ **Health system strengthening** (Department of Health 2012).

Both of these innovations highlight the need for more skillful health care providers.

2.4.8 Mentorship

While mentorship might be regarded as an implementation strategy for effecting professional behaviour change, it did not feature in literature sourced on practice change (2.4.1 – 2.4.7). This review will describe the field of mentorship, the effectiveness of mentorship for patient outcomes and the limitations of published research regarding the effectiveness of mentorship.

2.4.8.1 Defining the field

There was a multiplicity of interpretations of the terms mentorship and related concepts, e.g. clinical supervision, preceptorship, leadership and coaching, where sometimes these overlapped, were conflated or contradictory (Canadian Nurses Association 2004, Fisher and Webb 2008, McCloughen, O' Brien *et al.* 2011, Cassidy 2013, Mijares, Baxley *et al.* 2013, Chen and Lou 2014). This is exacerbated where definitions and scope of mentors may depend on particular national health and education systems, and where job descriptions do not necessarily conform to a broader understanding of the terms. The variety found in the literature will be presented highlighting the target, nature, dynamics, characteristics, expectations and settings of the mentorship relationship as well as the format of programmes for preparation for mentorship. Thereafter the interpretation used in this study will be presented.

2.4.8.1.1 Targets of mentoring

The targets of mentoring reported in European settings are undergraduates or students (Royal College of Nursing 2005, Latham, Hogan *et al.* 2008) and usually include some responsibility for assessment (Royal College of Nursing 2005, Nursing and Midwifery Council 2008)¹⁸. Studies from the United States of America (USA) or Asia which targeted registered nurses (RN) or midwives (RM) tended to focus on the first year following registration where the primary purpose of the mentoring programme is to assist the new graduate to apply theoretical knowledge to clinical practice, and to reduce attrition (Chen and Lou

¹⁸ This literature was excluded from the literature review (as indicated in the search strategy (2.2.3.1)) as this population was not relevant to the current study, but is included here to clarify the field.

2014). This probably reflects on the different educational frameworks existing in these continents and the expectations on students and newly registered nurses.

2.4.8.1.2 Nature, dynamics and characteristics of the mentorship relationship

The nature of the mentorship relationship is usually described as a one-to one relationship (Canadian Nurses Association 2004), but there are studies which refer to a one-to-many relationship or group mentoring (Brunero and Stein-Parbury 2008). The dynamics of the relationship may be voluntary (Canadian Nurses Association 2004) and initiated by the mentee (McCloughen, O' Brien *et al.* 2011). The relationship may be informal and develop organically (McCloughen, O' Brien *et al.* 2011) or formal, which may be depicted as disengaged with separate and predetermined functions (Fisher and Webb 2008, McCloughen, O' Brien *et al.* 2011, Mijares, Baxley *et al.* 2013). Given that formal mentorship relationships are usually initiated by a third party (as an institutional arrangement), careful attention to the interpersonal dynamics is required so that there is a sense of agency for the individuals concerned and an opportunity to develop an appropriate relationship (Allen, Eby *et al.* 2006).

Characteristics of mentorship relationships are variously described as mutually beneficial, facilitating personal development and self-confidence, sharing knowledge and experience, linking theoretical knowledge to clinical activity, guidance and role modelling, providing emotional support, encourage self-reflection and self-directed learning (Canadian Nurses Association 2004, Allen, Eby *et al.* 2006, Latham, Hogan *et al.* 2008, Steele 2009, McCloughen, O' Brien *et al.* 2011, Walker, Cooke *et al.* 2011, Mijares, Baxley *et al.* 2013).

2.4.8.1.3 Setting

The setting of the relationship could be on-site (Anatole, Magge *et al.* 2013) but geographic proximity alone is not critical so long as frequent interaction is feasible (Allen, Eby *et al.* 2006).

2.4.8.1.4 Preparation for mentorship

Where there are formal arrangements for mentoring, there may be formal mentor training for the mentor (Fisher and Webb 2008, Mijares, Baxley *et al.* 2013). Articles describing the preparation of mentors generally include the

following competencies, viz. communication styles and skills, managing expectations, fostering independence and promoting professional development and personal reflection (South African Qualifications Authority 2004, Pfund, House *et al.* 2014). A perceived high quality of training for mentors has shown positive results in terms of the perceived quality of the mentorship, career mentoring and role modelling (Allen, Eby *et al.* 2006). Preparation programmes of eight hours (Pfund, House *et al.* 2014), two days (Wallen, Mitchell *et al.* 2010, Anatole, Magge *et al.* 2013, Reid, Hinderer *et al.* 2013), four days (White and Winstanley 2010) and 39 hours (Reid, Hinderer *et al.* 2013, Chen and Lou 2014) have been described. Some of these programmes also have ongoing intermittent support (Latham, Hogan *et al.* 2008, Chen and Lou 2014). Preparation for mentorship also should address the organisational climate in order to support a culture of mentorship (Latham, Hogan *et al.* 2008, Steiner 2014).

2.4.8.2 Effectiveness in terms of patient outcomes

There is a general opinion that mentoring is 'a good thing' and produces positive outcomes (Canadian Nurses Association 2004, Allen, Eby *et al.* 2006, Steele 2009, Wallen, Mitchell *et al.* 2010, McCloughen, O' Brien *et al.* 2011, Mijares, Baxley *et al.* 2013). The majority of outcomes relate to the quality of, or satisfaction with, the mentorship programme from the perspectives of the mentor and/or mentee respectively, staff retention and reduction of attrition, and job satisfaction (McCloughen, O' Brien *et al.* 2011). These outcomes frequently emanate from qualitative designs making it difficult to build a consensus across studies (McCloughen, O' Brien *et al.* 2011, Cassidy 2013, Dilworth, Higgins *et al.* 2013). The literature search strategy (2.2.3) indicated a paucity of randomised controlled trials which are necessary to establish effectiveness. Although mentorship sometimes includes recognition of improved patient care or health outcomes (Anatole, Magge *et al.* 2013, Chen and Lou 2014), there was an absence of literature describing the establishment of any causal relationship between mentoring and effectiveness in terms of midwifery or nursing care, practice change, patient satisfaction or health outcomes.

Chen and Lou (2014) conducted a systematic review on mentorship amongst newly graduated nurses. Four studies were included which highlighted improved Nursing competencies, job satisfaction communication skills and interpersonal skills, other studies looked at retention and medical negligence. The evidence was not robust in that studies were rated as level III evidence, i.e. all were quasi-experimental in design, validation and reliability of instrumentation was not reported on and sample sizes tended to be low. Less than half of the studies reported any mentor training undertaken by the mentors, but one study showed a positive correlation between the seniority of the mentors and the mentees performance. However as this was the only study attempting to compare findings across a broader base, this was retained.

An outreach clinical mentoring programme to support the scale up of paediatric HIV care was instituted in Botswana (South Africa's north western neighbour) and a retrospective chart review conducted (Workneh, Scherzer *et al.* 2013). Monthly outreach visits by a physician and experienced nurse mentor focussed on supporting the national anti-retroviral therapy (ART) guidelines. During the visits side-by-side mentoring complemented didactic sessions on paediatric HIV care and treatment. The assumption was that mentoring would result in better documentation of care (process measure), which itself would be a proxy for the quality of care received. The chart review was able to analyse pre- and post-implementation dyads for six of the 15 indicators specified in the three sites where there were nurses. These indicators were pill count, correct lab monitoring, correct ART dosing, CD4 count within the last seven months, viral load testing within the last four months and disclosure. There were differences between the sites but the majority of the indicators showed broad and substantial improvements which were statistically significant (Workneh, Scherzer *et al.* 2013). While there is no detail of the quality or preparation of the mentors reported, there was a regular monthly outreach visit which reinforced the care requirements. One might argue that this is a very narrow interpretation of mentorship and it does underscore the difficulty of making comparisons or conclusions when there is no clear definition.

A nurse mentorship programme was initiated in 21 health centres in two rural districts of Rwanda in order to promote adherence to national clinical protocols in child health, women's health, adult health and HIV care (Anatole, Magge *et al.* 2013). Nurse mentors, who had a post-secondary Nursing degree and several years of clinical experience, retained a clinical attachment in their own institutions and were seconded to outreach visits to health centres in their regions, which they visited every four to six weeks. Their responsibilities were to supplement the training of the mentees¹⁹ and enhance their clinical skills, and to promote the use of national guidelines and conduct quality improvement. Mentors conducted side-by-side observation and mentoring during direct patient care, and completed checklists of both individual and system performance. The before- and-after results show statistically significant improvements in the mean percentage of assessments performed per consultation in relation to Integrated Management of Childhood Illnesses (IMCI), Integrated Management of Adolescent and Adult Illness (IMAI) and the first antenatal care visit. The correct classification of consultations for IMCI and IMAI, which had very low baseline levels before the mentor programme, showed a statistically significant improvement, whereas the first antenatal visit had been correctly classified at baseline (99.4%) and remained unchanged. The results suggested that the on-site clinical mentoring was a positive initiative to improve health care in rural Rwanda. However, as this was a before-and-after study with no randomisation or controls, it is not possible to make inferences about the intervention. In addition to the mentee performance, the study highlighted system issues that were significant barriers to high quality care, e.g. shortages in medical supplies and equipment. A long term view of creating team-based, system-focused solutions, rather than quick fixes, was regarded as essential and possible within a resource-constrained setting (Anatole, Magge *et al.* 2013).

Believing that evidence-based practice can lead to greater consistency in the care provided, greater patient satisfaction due to improved outcomes and higher

¹⁹ The nurse mentees were grade A2 nurses, i.e. had received their nursing training as part of their secondary school education. This category makes up the bulk of the nursing workforce in Rwanda.

quality of care and health care provider satisfaction, and that mentors contribute to the uptake of evidence to improve practice, Wallen, Mitchell *et al.* (2010) conducted a quasi-experimental study to measure the effectiveness of a structured multifaceted mentorship programme where nurses were targeted to lead evidence-based practice implementation. This study was situated in a well-resourced clinical research environment in the USA. The results revealed that there was a larger increase in perceptions of organisational culture and readiness, and in beliefs about evidence-based practice in the mentor group than the control group. Outcomes focussed on perceptions and 'knowledge about' rather than on patient outcomes. However, a clear need for organisational readiness and leadership support for a culture of evidence-based practice was demonstrated.

Latham, Hogan *et al.* (2008) introduced a mentoring programme in two hospitals with the intention of improving the work environment (retention, support, unit culture) and to improve selected patient outcomes, e.g. fall and pressure ulcer prevention and proper use of restraints. There was positive feedback from the mentors and most other parameters showed improvement. However, Latham, Hogan *et al.* (2008) conceded the improvement in patient care outcomes could not be attributed solely to this project. It was not possible to establish a causal relationship because the project design had no control arm and inclusion was on the basis of mutual interest.

2.4.8.3 Limitations of mentorship literature

Numerous authors, including those conducting systematic reviews, have commented on the paucity of the literature dealing with the impact or effectiveness of mentorship (Bosch-Capblanch and Garner 2008, Walker, Cooke *et al.* 2011), as well as the lack of rigour of the studies (Bosch-Capblanch and Garner 2008, Dilworth, Higgins *et al.* 2013) or empirical scrutiny (Allen, Eby *et al.* 2006).

A number of studies are survey-based and have described what mentorship could / should be, what skills mentors require, and who the beneficiaries are, e.g. health care providers, patients, services (Buus and Gonge 2009). However

these are cross-sectional studies reporting on retention, attrition and stress reduction, and it is not possible to make inferences about the causal links between mentorship and health or health care outcomes (Bosch-Capblanch and Garner 2008). Where patient outcomes have been stated, the study designs have not been robust enough to make confident inferences (White and Winstanley 2010, Anatole, Magge *et al.* 2013). Where systematic reviews have been attempted, authors have commented on the weakness of design and thus inability to make inferences (Chen and Lou 2014). It has not been possible to perform meta-analysis due to lack of conceptual clarity, standardisation of concepts and measures, thus such reviews are limited to being narrative in nature (Buus and Gonge 2009, Chen and Lou 2014).

Francke and de Graaff (2012) recognised that effect-oriented studies are necessary and these require appropriate random allocation of participants. However, the ward nature of nursing means that nurses within a team would influence one another. Therefore a cluster randomised trial design would be necessary to make inferences (White and Winstanley 2010). Further, Francke and de Graaff (2012) recommend validated measurement instruments and the use of independent researchers based on their impression that the roles of supervisor and researcher may sometimes coincide and thus increase the risk of bias. Although this relates to clinical supervision, the same caveat would apply for mentorship.

2.4.9 Instruments for evaluating partograph utilisation

Completion of a partograph is regarded as a key component of the standard of intrapartum care (World Health Organization 1989, Dujardin, de Schamphelre *et al.* 1992, Lavender, Alfrevic *et al.* 1998, 2002, Department of Health 2006a, Mathai 2009). Many studies look at the impact of the use of the partograph on maternal and neonatal outcomes (Lavender, Hart *et al.* 2008, Nyamtema, Urassa *et al.* 2008, Ith, Dawson *et al.* 2012).

The quality of intrapartum care was to be the primary outcome for this study. Assuming that the better the quality of the observations during labour, the greater the likelihood of accurate interpretation and clinical management, the

quality of the partograph record was to be used as a proxy measure of the quality of care given during labour. Thus a suitable evaluation instrument was required.

A number of studies have stated the need for clinical audit for improving quality of service (Nyamtema, Urassa *et al.* 2008). Some of these refer specifically to intrapartum care, some refer to the use of an instrument for auditing purposes, but very few give any detail about what instrumentation is used (Dujardin, de Schamphelre *et al.* 1992). A number of studies focused on particular conditions, disease processes or outcomes (sentinel events), e.g. caesarean section review or management of life-threatening conditions (Wagaarachchi, Graham *et al.* 2001), rather than auditing overall practice which may or may not be related to morbidity or mortality (Mancey-Jones and Brugha 1997, Wagaarachchi, Graham *et al.* 2001).

The following issues were of interest in establishing the suitability of a partograph utilisation instrument – the instrument used, reporting of validation and reliability testing, the availability of guidelines to inform the use of the partograph, the scoring system used, and how the evaluation was carried out. The instruments found in the literature search are presented individually highlighting these characteristics, where the information is available. Thereafter further details obtained through personal communication with relevant researchers are presented, and then commentary on the literature obtained is presented.

2.4.9.1 WHO Safe Motherhood questionnaire for assessment of partogram use

In their study based in Tanzania Nyamtema, Urassa *et al.* (2008) report that records of the parameters of labour were reviewed, as well as maternal and fetal conditions, using a WHO Safe Motherhood questionnaire developed for assessment of partogram use. This is a short survey tool where six items are checked (World Health Organization 2001). The scoring captured parameters that were ‘not recorded’, ‘substandard’ (by which was meant that the records did not meet the protocol standards) or ‘monitored / recorded to the standards’.

Standard protocols referred to specified time intervals for various clinical observations. Two other studies were found where this assessment tool was used none reported on the reliability of the scores obtained (Nyamtema, Urassa *et al.* 2008, Ogwang, Karyabakabo *et al.* 2009, Osungabade, Oginni *et al.* 2010).

2.4.9.2 Partograph checklist with five criteria

In a brief communication, Bosse, Massawe *et al.* (2002) reported on a descriptive study assessing the quality of use of partographs in southern Tanzania, the main focus being whether or not better quality use of partographs influences the caesarean section rate. A checklist consisting of five criteria developed by consensus among national experts was used, where each criterion could score a maximum of two points according to completeness, with a score of seven or less indicating unsatisfactory monitoring. Attempts to find out more from the corresponding author proved fruitless.

2.4.9.3 Partograph utilisation in Indonesia

Fahdhy and Chongsuvivatwong (2005) conducted a CRT study in Indonesia where partographs were reintroduced among a small group of midwives and the outcomes compared with a similar group of midwives who did not receive the refresher course on how to use the partograph. Evaluation was based on correctness and completeness. Correctness was based on the completeness of the partograph, i.e. graph completed, and internal consistency (with other parameters of labour progress, e.g. descent of fetal presenting part and increase in cervical dilatation). However, no further detail was reported. Communications were sent to both authors with no response.

2.4.9.4 Comparison of partograph designs in India

In a study conducted in India to compare two World Health Organization partographs, the researchers assessed user-friendliness, teachability and overall usefulness (Mathews, Rajaratnam *et al.* 2007). Although the health professionals using the partograph were limited to physicians, the study did report that the partographs were checked for completeness. However, the instrument used for this assessment was not specified, and follow-up with the corresponding author yielded no response.

2.4.9.5 Partograph utilisation in Kenya

In a study by Rotich, Maina *et al.* (2011) which evaluated partograph utilisation in two referral hospitals in Kenya, research assistants observed nurses and midwives monitoring women in labour using a structured observation checklist. The data was reflected as 'obtained and recorded correctly', 'not recorded or obtained incorrectly / incompletely' and 'not obtained therefore not recorded'. Given that the monitoring was observed could lend strength to the concept of correctness. However, it was not clear how the concept of completeness was determined when labour could last for several hours and require multiple recordings. Further, the mixing of the concepts of completeness and correctness raised queries as to the validity of the checklist and its reliability in this setting. Follow-up with the first author has not clarified this any further (Rotich personal communication 2012).

2.4.9.6 Partograph checklist in Iran

Although Simbar, Gharafi *et al.* (2009) reported on checklists developed to assess the quality of care in labour in Iran which were tested for validity and reliability, on following this up with the corresponding author it transpired that there were only four items that related to the partograph *per se*, viz. recording the examination, drawing the chart, recording fetal heart rate, and recording injections (Simbar personal communication 2012). This had limited value for the application required for this study.

2.4.9.7 Other sources identified

Given that the search parameters yielded little of value to this study, other sources explored where it was known that there had been some element of partograph evaluation, that could have contributed to measuring utilisation.

The Perinatal Education Programme (PEP) studies do not discuss how the partograph was evaluated (Theron 1997, Theron 1999). However, Theron presented his partograph scoring plan in his MD thesis, which looked at the effectiveness of the maternal care modules of the PEP programme. The scoring plan consisted of five sub-units – fetal observations including identification of risk, maternal records, contractions, progress of labour, and summary of labour.

The scoring ranged from three to minus two and did not appear to be consistently allocated for each sub-unit. There were no instructions or guidelines, therefore scoring appeared to be open to interpretation. Given that there was a single researcher using this instrument this would have reduced the possibility of different interpretations, but it would not serve the purpose of this study.

One study published in 2000 (which was referred to by several other authors) evaluated the utilisation of the partograph by Angolan midwives in a peripheral delivery unit (Pettersson, Svensson *et al.* 2000). The objective was to study the impact of an educational intervention on midwives' utilisation of the partograph. Ten variables were used to examine the partograph. This instrument is significantly short on detail. So, although this has been used in a published study, there is no information as to the validity or reliability in the study population.

Philpott, who pioneered the use of the partograph, developed a labour record review tool, along with a research associate, which has been recommended for conducting clinical audits (Philpott and Voce 2005). This is a 25 item checklist of which 11 items are directly related to the clinical observations required by the partograph. Each item has a maximum score of 1, but $\frac{1}{2}$, 0 or 'not applicable' may be assigned. The complexity of each item (which could include more than one clinical observation type, different frequencies of observation requirements depending on the phase of labour (i.e. double-barrelled (Streiner and Norman 2008)), and a loose scoring framework) suggested a high degree of variability between evaluators and even with the same evaluator at different evaluation times, thus limiting the reliability of the labour record review tool. During personal communication with one of the authors it was acknowledged that removing the $\frac{1}{2}$ mark scores could reduce the margin of error/variation so that a score either indicated completeness (1) or incompleteness (0) (Voce 2004).

A study which *inter alia* measured the inter-rater reliability of the Philpott and Voce tool found that there was a poor concordance using Lin's concordance

correlation coefficient (0.167) for the partograph-specific items (0.167), and low level of agreement (Clow 2012). Given that validity is limited by the extent of the reliability demonstrated (Streiner and Norman 2008), this indicates that both reliability of scores obtained and validity of this measure for this study population is limited.

Two other midwife researchers have separately researched the partograph and its use. Kwast did extensive testing of the WHO partograph but this related specifically to its use and its impact on obstetric outcome, not on how it was utilised (Kwast, Lennox *et al.* 1994). Although there was reference to partographs not being completed correctly but the quality of use improving with familiarity, there was no utilisation instrument reported (World Health Organization 1994). Personal communication with Kwast and access to original reports confirmed that there was no instrument which specifically measured partograph utilisation (World Health Organization 1994, Kwast personal communication 2014). Lavender has investigated the impact of different designs of partograph on the outcome of labour. Once again the focus has been on the obstetric outcome and not on the utilisation of the partograph *per se* (Lavender, Alfrevic *et al.* 1998, Lavender, Hart *et al.* 2008). Personal communication with Lavender has confirmed that there was no formally designed and tested instrument used for evaluating partograph utilisation, and neither was she aware of the existence of such an instrument (Lavender personal communication 2014).

2.4.9.8 Scoring considerations

Completeness and correctness of observations are concepts that need consideration when determining the scoring system, and these need to be well articulated. These concepts have two different meanings and cannot be scored on a single scale and simultaneously be reliable and meaningful. In the Fahdhy and Chongsuvivatwong study the definitions are problematic in that they are reflexive (circular), i.e. internal consistency is defined as 'the graph having consistency among the parameters of labour progress' and completeness was defined as 'the graph completed' (Fahdhy and Chongsuvivatwong 2005:303). A term cannot be used to define itself as this becomes meaningless.

2.4.9.8.1 *Completeness*

Given that labour takes place over a number of hours and during that time multiple observations need to be taken at multiple timepoints, a clear and unambiguous method of determining completeness should be in place. This would need to be supported by clear instructions as to how to score. This could then be assessed by a document review.

In the Nyamtema, Urassa, *et al.* study (2008), parameters were assessed according to standard protocols which were defined based on time intervals for the clinical observations. Records not meeting the protocol standard were judged 'sub-standard' or 'not recorded'. Thus categorical data was presented and an overall assessment could not be given. Meeting the standard required 100% compliance in order to achieve the 'monitored to standards' categorisation. This might be very stringent for a long labour when one set of observations is missed while all others are completed.

The Rotich, Maina *et al.* (2011) study did not include any information regarding the criteria or standards for assigning a score for incompleteness. Bosse, Massawe *et al.* (2002) did present a scoring system for completeness but the detail of how a score of zero, one or two was assigned to a criterion was absent.

The Philpott and Voce instrument had a maximum score of 1 for each item (with options of ½ and 0)(Philpott and Voce 2005). Some of the items included multiple observations, e.g. maternal pulse and blood pressure hourly, but the scoring guidelines did not indicate how to assign scores when one observation was present and the other present occasionally or absent.

2.4.9.8.2 *Correctness*

The second concept of correctness has a number of dimensions – accuracy of the observations, accuracy of recording of the observations, correct symbols used to denote various pieces of information, correct interpretation of the findings, and correct management instituted based on the recordings. A

comprehensive system is required to be able to capture these nuances and some of items would need to be observed at the time that the observations were done.

Rotich, Maina *et al.* (2011) did attempt to determine if the observations were done accurately by having research assistants doing direct observation of the nurses and midwives who were monitoring the labours (Rotich, Maina *et al.* 2011). However, this cannot be a sustainable plan. Philpott and Voce in their work in Limpopo Province instituted a system that all comprehensive clinical assessments (usually done four hourly, or more frequently if indicated) should be checked by an advanced (or senior) midwife or doctor as a matter of routine (Voce 2005). If staffing levels permit, this could be a sustainable strategy for addressing most of the issues raised surrounding correctness. While the issue of staffing levels may be regarded as a smokescreen by some, if there is only one registered nurse-midwife on duty in the labour ward / women's ward and there are no resident doctors, then such a strategy is not feasible.

Fahdhy and Chongsuvivatwong (2005) attempted to bridge the issue of completeness and correctness by measuring internal consistency. This assumes that progress of labour is predictable. The example given by them is that of increasing cervical dilatation matching the descent of the fetal presenting part. However, descent might be arrested or slow despite evidence of a dilating cervix due to cephalo-pelvic disproportion, a full bladder or other reasons. The recordings could be quite accurate / correct, but the application of internal consistency would deem them to be incorrect.

Rotich, Maina *et al.* (2011) did not describe how correctness was determined. Bosse, Massawe *et al.* (2002) and Nyamtema, Urassa *et al.* (2008) did not report on assessing correctness.

2.4.9.9 Length of a partograph utilisation checklist

Some of the checklists identified in the literature consisted of five or six items, The partograph consists of up to 35 items in five clinical observation and clinical management domains with numerous observations within each, and at multiple

timepoints (the extent depending on the duration of labour). Given that a labour could last over several hours during which time there could be ten or more timepoints when observations are required, this should be accommodated in any instrument designed. Streiner and Norman explain that due to there being some associated error of measurement for every response / observation, by averaging this over a series of questions one can reduce the error. Thus there is good reason for long tests as brevity could result in reduced accuracy (Streiner and Norman 2008) The poor concordance and agreement shown for the Philpott and Voce labour record review instrument, particularly in respect of partograph scores suggest that brevity did not serve this instrument well. This may well apply to other partograph utilisation evaluation instruments which are brief.

2.4.10 Conclusions

2.4.10.1 Practice change

The EPOC group has described the methodological and practice issues that need to be considered when conducting a systematic review in the area of practice improvement (Grimshaw, McAuley *et al.* 2003). In so doing they highlight aspects that need attention in designing of studies in this field so that a larger and more scientifically robust database of studies can be used for future systematic reviews, or updates of existing systematic reviews. They present some of the challenges encountered in this field where there is so much scope for variation (Grol, Baker *et al.* 2002). While a number of strategies have received some attention, e.g. audit and feedback, reminders, and continuing medical education, it is suggested that randomised controlled trials will not be able to explain some of the basic questions about the critical success factors in change processes, and that RCTs will need to be complemented by observational and qualitative studies (Grol, Baker *et al.* 2002).

There is a similar degree of change across a number of interventions and combinations. When using interventions such as educational outreach visits, audit and feedback or local opinion leaders as quality improvement strategies, one could expect absolute improvements in practise of five to ten percent. While

multifaceted interventions using a combination of successful single interventions are likely to yield greater effects than single interventions, the context in terms of health care provider, health status, health service, and the setting should be considered in making an appropriate choice.

Further, any intervention needs also to address system and / or environmental factors including organisational culture (Latham, Hogan *et al.* 2008, White and Winstanley 2010). Bingham and Main (2010) suggest that implementation practices of knowledgeable, tenacious and creative frontline physicians and nurse leaders may have the greatest impact on quality improvement implementation effectiveness, since they are the individuals who decide how the strategies and tactics are tailored. This highlights the need for an actively and intentionally managed programme with institutional pressure for change to occur. The use of a Practice Improvement model with its rapid cycle approach to allow for multiple incremental changes and/or the Knowledge-to-Action framework can guide participants in such endeavours.

2.4.10.2 Mentorship

In the field of mentorship in Midwifery and/or Nursing there appears to be a lack of conceptual clarity on the terminology, roles, and expectations of mentors, which can traverse various national, educational and health system contexts. While there are a number of articles which focus on mentoring and students, there is very little literature available where the focus is on the registered midwife or nurse in practice, and where these do exist the focus often appears to be on the first year after registration, stress management or reduction, retention strategies, rather than on patient outcomes (Allen, Eby *et al.* 2006, Chen and Lou 2014). In many of the studies described it seems that mentorship is an end in itself. Articles tend to be descriptive rather than intervention studies, thus research questions remain regarding the effectiveness of mentorship (for registered midwives or nurses) on patient and health outcomes.

2.4.10.3 Instrumentation to evaluate partograph utilisation

The literature reveals inadequate instrumentation to evaluate partograph utilisation.

2.5 What gaps exist

2.5.1 Practice change

This literature review shows that most of the studies which were included in the major systematic reviews were undertaken in developed countries, in urban areas, amongst the physician population. Some of the studies included patient outcomes but most concentrated on professional practice. While various reviewers thought that many of the strategies should (theoretically) be translatable into developing countries, the lack of infrastructure (including internet access and library material), lack of professional / clinical opinion leaders, staff shortage, and high turnover of staff (mitigating against sustainable system change), made such translation less certain.

One might ask if studies from developing countries or Nursing populations do not exist, if the inclusion criteria for the systematic reviews are not sensitive to what is being researched in these contexts, or if the scientific quality causes them to be excluded. One noteworthy systematic review concentrated on low and middle-income countries but the studies included were all before-and-after studies and the quality was described as low. However the authors indicated that the consistency of findings suggested that audit maybe a useful tool for decreasing perinatal mortality rates and improving the standard of care (Pattinson, Kerber *et al.* 2009).

Systematic reviews which listed references to excluded studies and those studies awaiting assessment did not show higher proportions of these studies being from developing countries nor of Nursing or Midwifery populations (Cheater, Baker *et al.* 2005, Doumit, Gattellari *et al.* 2007, Jamtvedt, Young *et al.* 2007, O'Brien, Rogers *et al.* 2007, Forsetlund, Bjorndal *et al.* 2009, Baker, Camosso-Stefinovic *et al.* 2010). This suggests that the bias in the literature reflects the resource-rich sectors of professionals rather than a disproportionate exclusion of Nursing and Midwifery literature. One cannot extrapolate the findings from one professional group to another as the ability to access and adopt new practices are influenced by the educational and practice contexts

which have different preparation, programmes, expectations, professional values, power, status, which may be important effect modifiers (Cheater, Baker *et al.* 2006, Thompson, Estabrooks *et al.* 2007). However, when nurses and Nursing practice were targeted, the quality of studies did not meet the quality criteria set, despite extensive searching of published and unpublished research (Flodgren, Rojas-Reyes *et al.* 2012). This highlights the need for more researchers to be encouraged and developed in low and middle income countries and amongst non-physicians, with the necessary support to conduct well designed studies.

Given the complexity of professional behaviour change and how difficult this makes comparison between studies, there may be a place for introducing a format similar to the CONSORT statement which requires the type, frequency and number of strategies used, as well as a clear description of these, to be reported for each study. This should also indicate the type of population and the estimated cover of the population targeted. Forsetlund, Bjorndal *et al.* (2009) made a similar suggestion. The taxonomy of interventions defined by the EPOC group may require adaptation for Nursing (Thompson, Estabrooks *et al.* 2007).

While RCTs or CRTs may yield clear outcomes, where a complex intervention is introduced it is not possible to determine the contribution of each aspect (Milne, Scotland *et al.* 2004). It may be necessary to accept this limitation when complex interventions are utilised. Another way of understanding the contribution of the components of a complex intervention would be to employ mixed methods or more qualitatively orientated research as was done in the Ethiopian study (Pitchforth, Lilford *et al.* 2010). Pitchforth, Lilford *et al.* (2010) suggest that in resource-poor settings and where information systems are weak, this approach could be a useful extension of more traditional clinical data collection. This approach is particularly valuable when including client / patient or health care provider views and experiences especially when needing to engage with, and understand professional politics, norms of working and broader organisational factors which influence provision of care (Pitchforth, Lilford *et al.* 2010).

There is a growing realisation of the need to conduct an economic evaluation or at least a costing analysis of the various interventions. Generally the literature sourced has not included this but a number have called for this to be included in the future (Fitzpatrick, Davey *et al.* 1998, Grimshaw, Thomas *et al.* 2004, Horsley, Hyde *et al.* 2011) and some have answered the call (Lawn and Kerber 2006, Groenewald-Neethling 2010, Pattinson, Kerber *et al.* 2011).

Thus there is a need for good quality studies of interventions and processes with well designed implementation strategies (based on an appropriate theoretical framework) to be undertaken in low- and middle-income countries, which take account of the contribution of midwives and nurses, and the context of their work.

2.5.2 Mentorship

There is lack of conceptual clarity in the field of mentorship, clinical supervision and preceptorship making it difficult to build a body of knowledge in a well-defined area of study. There is very little literature on the effectiveness of mentorship in terms of patient or health outcomes – both from research design and discipline-specific, i.e. Midwifery, perspectives.

2.5.3 Instrumentation to evaluate partograph utilisation

Disclosure about the evaluation tools used for evaluating partograph utilisation is minimal. Where the partograph has been evaluated in terms of the completeness of use, there has been very little formal validation or reliability testing of instruments. The partograph evaluation instruments presented in 2.4.9 (Theron 1997, Pettersson, Svensson *et al.* 2000, World Health Organization 2001, Rotich, Maina *et al.* 2011) do not exhibit the rigour necessary for a research quality instrument.

Designers of quality assurance or improvement programmes should be concerned with reliability and validity of measuring tools. Marshall, Lockwood *et al.* (2000) found that in 300 randomised controlled trials in schizophrenia, studies were almost 40% more likely to report that treatment was effective when unpublished tools were used than when validated ones were used. Therefore,

an instrument to be used in a quality assurance programme should be evaluated for its appropriateness.

2.6 Implications for this study

It has been seen across the literature review on implementation strategies and professional behaviour change that the lack of consistency and clarity of terminology and approaches hampers comparisons between studies. In this study careful attention has been paid to international standardised reporting requirements where these exist (Ogrinc, Mooney *et al.* 2008, Zwarenstein, Treweek *et al.* 2008, Schulz, Altman *et al.* 2010, Campbell, Piaggio *et al.* 2012), and where these do not to ensure that clear definitions are made and descriptions given.

Originally, the literature review concentrated on the effectiveness of strategies and programmes that aim to change professional behaviours and improve practice. Given that context and attributes of roleplayers influence the success of implementation strategies, it appeared that certain interventions tended to have positive outcomes. These were enhanced when used in combination but no single intervention package provided the 'magic bullet'. However, it did appear that certain combinations working at both individual and institutional levels, appropriately packaged and delivered were likely to yield a positive result.

To this end an intervention package consisting of focussed intrapartum in-service training programme (educational outreach and educational materials) combined with the development of clinical facilitation (safe climate, regular feedback based on audit-type data, support) by an experienced midwife and supported by the nurse manager who had responsibility for quality assurance and audit, was developed.

Originally the package of interventions was to include a 'mentorship' component to support the practise change and system level support required. However, given the divergent interpretations of mentorship it was decided to name the

person undertaking this role a 'clinical facilitator', and the related activity termed 'clinical facilitation'. These terms are not widely used in the literature. For the purposes of this study a **clinical facilitator** is an experienced registered midwife who has responsibility for intrapartum clinical care and who works alongside less experienced registered midwives. Her (his) role is to facilitate a positive learning climate in the labour ward, to assist colleagues to improve clinical insight, clinical judgment and clinical skills in order to improve the standard of patient care. The clinical facilitator may also be the nurse manager who is responsible for quality assurance and clinical audit of the labour ward, and who supports and complements the experienced midwife in the clinical facilitation role. **Clinical facilitation** relates to the activities and behaviours of the clinical facilitator to address the clinical development needs of the registered midwives with whom s/he works.

Lessons from the literature in the broad field of mentorship were incorporated into the training of clinical facilitators. Details of the intervention package are presented in 3.7.4.

The study was aimed at the institutional level and the study design adopted was that of a cluster randomised trial, where outcomes measured both patient outcomes (partograph utilisation as a proxy measure) and provider knowledge and skills. Given that an appropriate partograph utilisation instrument for this study was not found it was necessary to develop a valid instrument and test the reliability of the test scores. A summary of this process is presented in 3.7.6.

2.6.1 Aim

The aim was to evaluate the effectiveness of a multifaceted intervention package of training and clinical facilitation on the quality of clinical management in labour by registered midwives in primary level public sector health facilities in rural Western Cape, South Africa. As this intervention targeted all involved in intrapartum care in an institution, and in order to reduce the possibility of contamination, the unit of randomisation was the institutional site.

2.6.2 Objectives

1. Conduct a document audit of individual partographs and compare control and intervention sites/clusters over a 12 month period following the introduction of the package of interventions with respect to
 - ◆ completeness of the documenting of intrapartum observations
 - ◆ evidence of adherence to the management of labour guidelines

This primary outcomes was measured at the individual record level (sites were clusters).

2. Evaluate and compare the performance of the individual registered midwives in control and intervention arms with respect to
 - ◆ knowledge
 - ◆ clinical judgment
 - ◆ clinical assessment
 - ◆ the ability to use and interpret the partograph, and
 - ◆ adherence to the management of labour guidelines

This secondary outcome was measured at the individual participant level (sites were clusters).

3. Evaluate and compare the performance of individual registered midwives at months one, three and 12 within the intervention arm with respect to
 - ◆ knowledge
 - ◆ clinical judgment
 - ◆ clinical assessment
 - ◆ the ability to use and interpret the partograph, and
 - ◆ adherence to the management of labour guidelines.
4. Compare the effectiveness of the current unplanned *ad hoc* approach to training, to full coverage training in combination with the clinical facilitation intervention, measuring outcomes at the individual level.

2.6.3 Hypotheses

Hypothesis 1 : A well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators improves the standard of practice of registered midwives.

Null hypothesis 1 : A well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators does not improve the standard of practice of registered midwives

Hypothesis 2 : A well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators mentoring and monitoring by trained clinical mentors improves the knowledge and skills of registered midwives.

Null hypothesis 2 : A well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators does not improve the knowledge and skills of registered midwives.

Chapter 3 Methods

3.1 Introduction

This chapter sets out the research methods considered for this study (3.2), discussing the various options and implications for design, sampling, analysis and reporting, and the rationale for the research approach and design chosen. Thereafter a detailed description of cluster randomised trial design considerations through all research phases is presented (3.3), followed by considerations in the design of the intervention (3.4), indicators to be used (3.5) and the instrumentation (3.6). The process undertaken to validate and test the instrument for evaluating partograph utilisation (to measure the primary outcome) is summarised in 3.6.

Having laid the foundation, the methods employed and their implementation, the instruments used, and the process of data management, analysis and reporting are presented (3.7), followed by the ethical considerations for a pragmatic cluster randomised trial (3.8) and the chapter's conclusion (3.9).

Figure 3.1 illustrates the methodological considerations (discussed in 3.2 and 3.3), and design and analysis decisions made (discussed in 3.7). The underlined text indicates aspects that were considered but not incorporated into this study, and the green font indicates the decision making process.

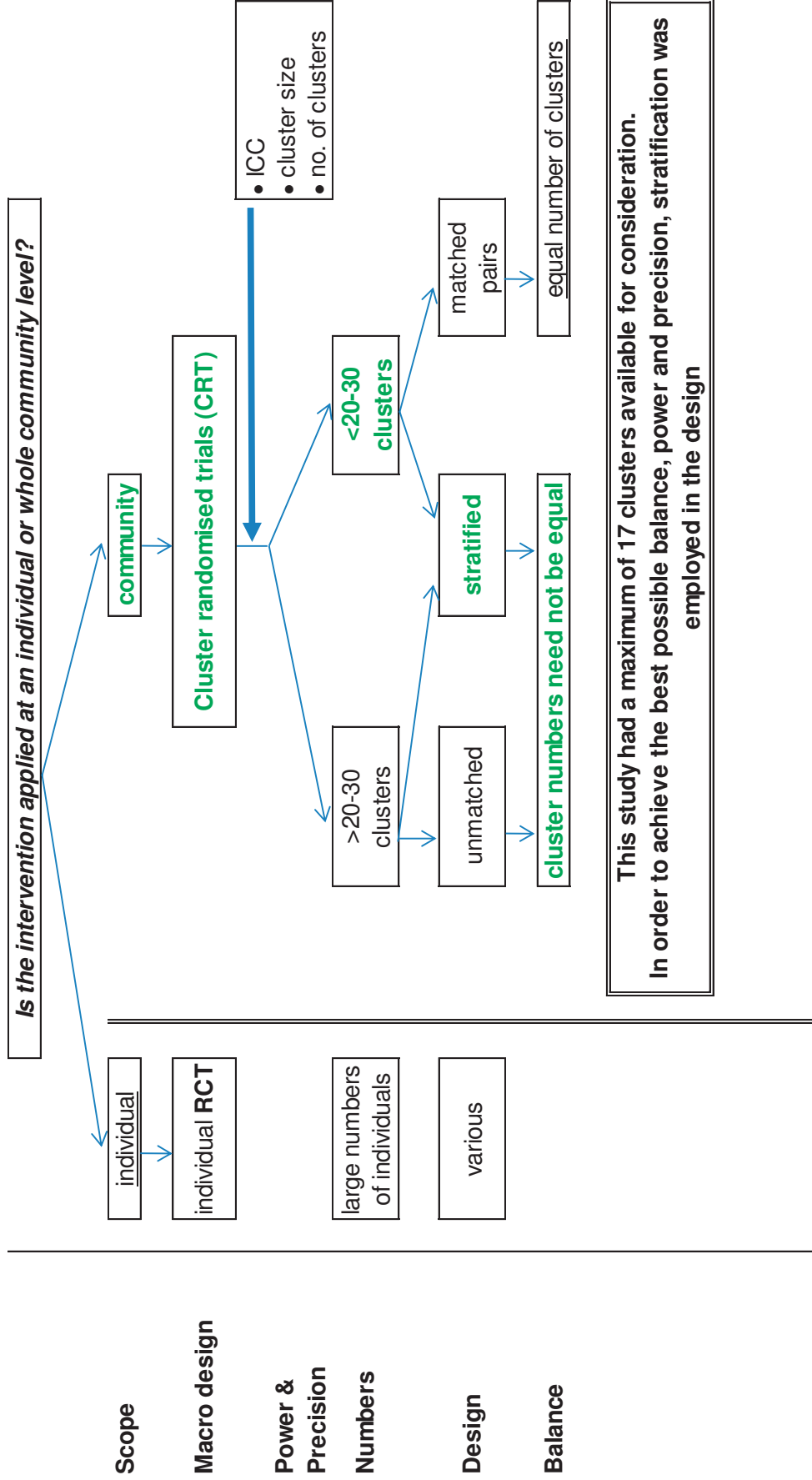
3.2 Considerations for appropriate study design

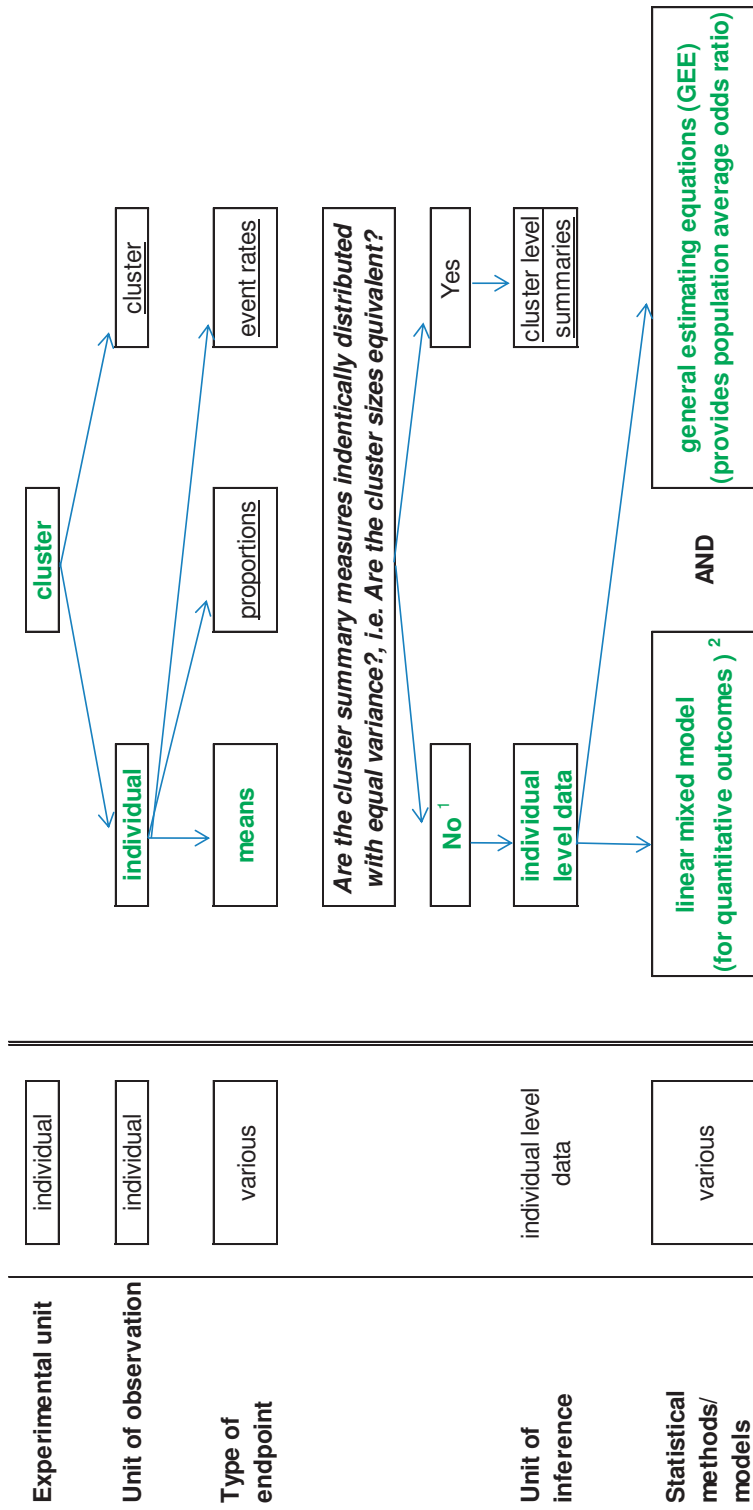
3.2.1 Research method

3.2.1.1 Randomised Controlled Trial

Explanatory trials aim to measure the efficacy of an intervention when applied under optimal conditions. Such studies are usually designed as randomised controlled trials (RCT) as they are regarded as the 'gold standard' (Katzenellenbogen, Joubert *et al.* 1997, Moher, Hopewell *et al.* 2010), and thus

Figure 3.1 Algorithm to determine the appropriate components of the study design and analysis





Notes :

Green type represents the choices made in this study; underlined type represents inappropriate options for this study

1. Two sets of observations are made. The cluster sizes in terms of the primary endpoint (partograph audit) are equal; the cluster sizes in terms of the secondary endpoint (midwife tests) vary considerably

2. Application of random effects model for means, taking account of between-cluster variation. This model assumes that the observed value of x_{ijk} is given by : $x_{ijk} = m_{ijk} + \eta_{ijk} = a + b_i + S_j + g_l + z_{ijk} + u_{ij} + \eta_{ijk}$, where v_{ijk} represents individual-level departure from the mean, with variance s^2_w , and u_{ij} represents the cluster-level variation

Explanation of terms : x = quantitative variable (endpoint of interest); m_{ij} represents the true mean of that variable among individuals in the j th cluster in the i th treatment arm ($i=1$ - intervention; $i=0$ - control); k represents the individual.

Based on information obtained from Hayes & Moulton 2009; Katzenellenbogen, Joubert *et al* 1997

dominate the evidence based practice movement (Sackett, Rosenberg *et al.* 1996, Higgins and Green 2011). Randomised trials estimate the impact of an intervention applied to a randomly allocated group which is directly compared to a control group which receives no intervention or an alternate intervention.

The RCT design carries the assumption that observations on individuals are statistically independent of each other. This assumption has implications for randomisation, data collection and analysis. In RCTs the individual is the unit of randomisation, and data are collected and analysed at an individual level. The key features of this design are :

1. the interventions / treatment conditions are allocated to the participants by the investigator
2. there is a control arm that is followed up in parallel to the intervention group using similar methods so that the outcomes can be compared between the groups over the same time period
3. randomisation is used to allocate the participants to the treatment conditions being compared (Hayes and Moulton 2009)

The randomisation process ensures that, everything else being equal, both known and unknown biases are distributed equally between the trial arms (Campbell, Steen *et al.* 1999). These assumptions do not hold when researching groups whose characteristics are not statistically independent.

While the RCT approach is appropriate for many intervention studies, some authors have acknowledged that the appellation 'gold standard' is something of an exaggeration and the evidence from RCTs and its appropriateness for application need to be assessed (Hotopf, Churchill *et al.* 1999). Further, the strict conditions required limits the application of this design in more complex settings (Miller, Sloan *et al.* 2003). Given the limitations in design of RCTs for group settings/dynamics and where behaviour change is being measured, this approach was not chosen and is not discussed further.

3.2.1.2 Cluster randomised trial

Cluster randomised trials (CRTs) are used to measure the population-level effects of interventions delivered to groups of individuals where defined groups are referred to as clusters, and where there is an assumption that observations on individuals in the same cluster are correlated (Hayes and Moulton 2009). This research design is regarded as a key tool for evaluating interventions in health services research particularly where there is an educational intervention amongst the health professionals, or where clusters are complete communities (Campbell, Donner *et al.* 2007). Clusters may be geographic e.g. village; institutional, e.g. a health unit or workplace; or smaller e.g. household. CRTs have many of the key features of individually randomised trials but there are additional logistical, ethical and statistical issues to consider (Elbourne and Campbell 2001).

3.2.1.2.1 Rationale for using a CRT design

3.2.1.2.1.1 Setting

When investigating the effect of an intervention at a public health level, e.g. a geographic area, political region or health service, there are dynamics due to the group nature of these settings where some degree of correlation might be expected, and thus the assumption of independence which is made for randomised controlled trials (RCT) is usually invalid. Not only is there a degree of correlation within such a grouping (cluster), but within-cluster correlation depends on the existence of other clusters, i.e. it has no meaning if there is only one population being studied. In such studies it is therefore more appropriate to consider a CRT which takes account of these dynamics.

The main reasons for considering a cluster randomised design are that the intervention has to be applied to a whole community or grouping; that contamination (that might result if individuals within the same community were to be allocated to different arms of the study) be avoided; and to capture the population-level effects of an intervention applied to a large proportion of a population (Hayes and Moulton 2009). If such conditions apply, a CRT is more appropriate than a RCT.

3.2.1.2.1.2 *Type of intervention*

Educational interventions tend to be evaluated at the individual or cohort level but have not taken into account the effect of correlation within sites or differences between sites (Murray 2001). Further, when evaluating behavioural change strategies which are likely to involve complex interventions amongst groups, simple randomised trials which evaluate the effect of a single intervention may be less robust (Campbell, Steen *et al.* 1999, Campbell, Fitzpatrick *et al.* 2000). While RCTs are common, there is increasing interest in the areas of health promotion and knowledge translation where it is more appropriate to randomise according to a practice setting or other coherent grouping. In the field of safe motherhood, the importance of CRTs for programme evaluation has been emphasised (Miller, Sloan *et al.* 2003).

3.2.2 **Pragmatic approach**

Pragmatic trials aim to measure the effectiveness of an intervention when applied routinely in a population, whereas RCTs have rigidly controlled processes thus replicating as far as possible laboratory-like (optimal) conditions. Given that the health service is a dynamic organisation a pragmatic trial was considered more appropriate for this study, although it is recognised that due to anticipated imperfect compliance in the delivery of the intervention, effectiveness is generally assumed to be lower than efficacy (Hayes and Moulton 2009). While pragmatic evaluation has the advantage of its focus being aligned to that of policy makers and thus likely to have action taken on its findings, some authors have suggested that this could result in compromising methodological rigour to accommodate the needs / demands of the policy makers (Pawson and Tilley 1997).

3.3 **CRT design considerations**

The type of intervention, the logistics associated with implementing it and the specific scientific question influence the choice of design (Hayes and Moulton 2009). This has implications for the study in terms of design, cluster determination, sample size, randomisation, analysis and statistical modelling, and reporting.

3.3.1 Characteristics of study design

When planning the design and analysis of CRTs the following characteristics are key :

1. the unit of assignment is an identifiable group rather than an individual, and should have some form of association.
2. groups are assigned to different study conditions creating a nested or hierarchical structure for the design and for the data.
3. units of observation are the members of the clusters so they are doubly nested within the groups and conditions of the study.
4. group randomised trials often involve small numbers of groups in each condition, usually less than ten (Murray 2001).
5. because cluster sizes are often large it is usually not possible to randomise a large number of clusters to each treatment arm (Hayes and Moulton 2009).

3.3.2 Design strategies

Bias is more of a concern in group randomised trials than in clinical trials (where the individual is the unit of assignment) because there are fewer units (clusters) to randomise, therefore increased attention needs to be paid to design strategies that will reduce possibilities of bias or evenly distribute all potential sources of bias (Murray 2001). Because one cannot rely on randomisation to ensure adequate balance between arms, and because cluster sizes are often large, making it difficult to randomise a large number of clusters to each treatment arm, certain design strategies may be used to address this issue, e.g. matched pairs, restriction and stratification.

3.3.2.1 Matched pairs

By grouping clusters into similar pairs (where these are presumed to be more similar in respect of the primary outcome), a matched analysis can be done where comparisons are made between the matched pairs within the treatment arms. Examples of the use of this strategy are found in studies by Grosskurth, Mosha *et al* (1995), Bhandari, Bahl *et al* (2003), Manandhar, Osrin *et al* (2004), and Stiell, Clement *et al* (2009). The implication of this strategy is that the overall coefficient of variation, k , across all communities is replaced with the

coefficient of variation within the matched pairs (which is likely to be smaller than k) thus resulting in increased power and precision (Hayes and Moulton 2009). However, the most serious disadvantage of the pair-matched design is the loss of power and precision due to the loss of degrees of freedom resulting from the lack of replication within each matched pair (Hayes and Moulton 2009). Other disadvantages are the potential drop-out of clusters (which would affect the paired cluster of the dropped cluster) and the limitations in statistical inference for matched trials. The latter would require adjustment for covariates because regular regression methods do not work. Further, testing for variation in the intervention effect cannot be done due to lack of replication, and estimation of intra-cluster correlation coefficient and coefficient of variation cannot be calculated. It was thus decided that this would not be a suitable choice for this study.

3.3.2.2 Restriction

This is done by restricting clusters to allocations that satisfy certain pre-determined criteria. One allocation is then selected randomly from this restricted subset (Hayes and Moulton 2009). This may be used to ensure an acceptable level of overall balance using baseline or pre-existing data on each cluster, particularly if there are limited number of clusters. Examples of this strategy are found in studies by Hayes, Chantaluchab *et al* (2005), and Groenewald-Neethling (2010).

3.3.2.3 Stratification

Stratification involves the grouping of clusters into two or more strata that are expected to be similar with respect to the outcome of interest. Clusters within each stratum are then randomly allocated between treatment arms so that, at least in some strata, more than one cluster is allocated to a treatment arm (Hayes and Moulton 2009). A stratified design lies between the two extremes of unmatched and pair-matched designs, and has the advantage of replication of clusters within each treatment arm within a stratum. The aim of stratification is to reduce the variance of the estimated treatment effect. Examples of this strategy are found in studies by Gülmezoglu, Villar *et al* (2004), More, Bapat *et al* (2008), and Jackson, Cheater *et al* (2011).

A stratified design helps to improve the balance between the study arms and to reduce the between-cluster variability. This design has an advantage over the pair-matched design in that it does not fall prey to the same reduction in degrees of freedom available when estimating intervention effects (Hayes and Moulton 2009). The specific advantage of the stratified design is that there is replication within strata which allows for a direct estimate of between-cluster variance (or equivalently the intraclass correlation coefficient) within strata. Replication also means that it is possible to examine if there is a variation in the intervention effect between strata which is not possible in a pair-matched design. Pair-wise differences reflect between-cluster variation and any variation in the intervention effect, and it is not possible to separate these out (Hayes and Moulton 2009).

A further consideration is how many strata to use. If there are few clusters and many strata there is loss of degrees of freedom and a further reduction in study power will occur. Where there are no more than six clusters per arm, strata should be limited to two. Three strata can be considered where there are seven to ten clusters per arm (Hayes and Moulton 2009).

3.3.3 Sample considerations

A CRT requires an appropriately constructed sample that pays attention to the relationships within and between the clusters, as well as the number and size of the clusters. If there is not sufficient replication designed with an appropriate balance / relationship between these various concepts, it will not be possible to make any statistical inferences (Hayes and Moulton 2009).

3.3.3.1 Design effect

The design effect (the ratio of total number of subjects required using cluster randomisation to the number required using simple randomisation) can be calculated from the intraclass correlation co-efficient and the cluster sizes, where the higher the ICC and corresponding design effect, the greater the loss in power when using a CRT than an individually randomised trial design (Elbourne and Campbell 2001).

The formula for the design effect [$DEff = 1 + (m-1)\rho$] shows that if the intra-cluster correlation coefficient (ρ) remains constant, the design effect increases with cluster size (m). Similarly, even if the intra-cluster correlation is small the design effect may be substantial when there are large clusters. Thus a large number of small clusters is statistically more efficient than a small number of large clusters. Since the design effect for a CRT is generally greater than one, the required sample size when clusters are randomised is greater than for an individually randomised trial addressing the same study question (Hayes and Moulton 2009). Ignoring the effect of clustering in the design stage of a trial can lead to an elevated type two error²⁰, whereas ignoring this at the analysis stage inevitably results in an elevated type one error²¹ (Campbell, Donner *et al.* 2007).

3.3.3.2 Relationships within clusters

A cluster design assumes a degree of similarity or dependence within clusters. The intra-cluster correlation co-efficient (ICC) (represented as ρ) is a measure of the extent of the reduction in independence. Sources of within-cluster correlation may be based on correlation due to interaction between individuals, population characteristics, or variations in responses to interventions (Hayes and Moulton 2009). This has implications for study size and power as there will be some loss of power due to randomising by cluster rather than individual, and this should be reflected in the sample size calculation (Kerry and Bland 1998).

A ρ value of zero implies that there is no clustering so that individuals within the same cluster are not more similar than individuals from different clusters, i.e. there is no between-cluster variability. A ρ value of one implies that individuals in the same cluster are perfectly correlated, i.e. they all have the same outcome (Hayes and Moulton 2009). With quantitative data, ρ can be determined by using estimates of the variance between clusters and within clusters which may be based on empirical data available on the outcome of interest, by one way analysis of variance of between- and within-cluster variation, or by informed judgment (Hayes and Moulton 2009).

²⁰ Type two error – accepting an hypothesis when it should be rejected.

²¹ Type one error – rejecting an hypothesis when it should be accepted.

3.3.3.3 Variation between clusters

Inferences made from a CRT depend on the between-cluster variation and the outcome of interest. The structure of the data and the correlations between observations should be illustrated. If these are ignored by using only standard methods of analysis, invalid conclusions would be made (Hayes and Moulton 2009). Estimates of the coefficient of variation (k) may exist where prior data on between-cluster variation have been obtained from the proposed study clusters or from comparable clusters in a different but similar population (Hayes and Moulton 2009). However, empirical data on between-cluster variation of the outcome of interest may not be available when a CRT is designed. If it is not feasible to carry out preliminary research to obtain such data, the required sample size should be examined for various plausible values of k based on expert judgment (Hayes and Moulton 2009) or computer simulation based on previous research findings (Christie, O'Halloran *et al.* 2009).

3.3.3.4 Number and size of clusters

When determining the construction and size of the sample it is essential to ensure that there is sufficient replication. In all situations the total number of subjects required is greater than if a simple random allocation is used. A cluster randomised trial which has a large design effect requires many more subjects than a trial of the same intervention which randomises individuals (Kerry and Bland 1998).

If there is a high intra-cluster correlation then one can reduce the cluster size and increase the number of clusters to increase the power of the trial. (Indeed, where $\rho=1$ it would not be necessary to have a cluster size of more than one as all values within the cluster would be the same).

The size of cluster is influenced by statistical and logistical considerations, as well as by the need to minimise the risk of contamination. One can consider reducing the number of clusters but increasing the size of the clusters in order to maintain the same level of power for a study (Kerry and Bland 1998). The feasibility of the various options should be considered during the design.

Unlike a RCT where one can increase the power by increasing the size of the sample, careful attention should be paid to how the sample size will be altered. This depends on the within-cluster correlation or between-cluster variability so that either the size of the cluster or the number of clusters is altered (Hayes and Moulton 2009).

3.3.4 Randomisation

Randomisation of clusters is determined by the study design strategy employed, e.g. matched pairs, restriction, stratification. Thereafter, random sampling within clusters may be specified as simple, systematic or stratified depending on the nature of the population and the outcome to be measured.

3.3.5 Analysis

As individuals within a cluster are regarded as sharing a measure of common variance, i.e. not statistically independent, application of standard t-tests, analysis of variance (ANOVA) or analysis of covariance (ANCOVA) without considering the clustered nature of the data are likely to be inaccurate as standard errors will be underestimated (Christie, O'Halloran *et al.* 2009).

3.3.5.1 Approach to analysis

There are two main approaches to analysis in CRTs, *viz*

1. analysis based on cluster level summary measures (Horbar, Carpenter *et al.* 2004, Baqui, El-Arifeen *et al.* 2008), and
2. regression analysis based on individual level data using regression methods that allow for clustering (Bhandari, Bahl *et al.* 2003, Koethe, Westfall *et al.* 2010).

Some studies may use a combination (Manandhar, Osrin *et al.* 2004, More, Bapat *et al.* 2008, Lewycka, Mwansambo *et al.* 2010, Zhang, Deneux-Tharaux *et al.* 2010).

The cluster level summary approach is not statistically efficient as each cluster of individuals produces only one data value and, if there is wide variation within clusters, each summary value would need to be weighted (Christie, O'Halloran *et al.* 2009). The advantage of performing analysis based on individual level data is statistical efficiency as this is a one stage processes where the endpoint

of interest and the covariates are analysed together in the same model, and interaction terms can be incorporated in the model to examine the evidence for effect modification. The main disadvantage is that this does not perform reliably for CRTs with fewer than 15 clusters per arm, with reference to significance tests and confidence intervals which may not have the correct size and coverage (Hayes and Moulton 2009).

3.3.5.2 Statistical modelling

A number of regression models can be used depending on the parameter of interest. These are extensions of models used in epidemiological research modified to take account of intracluster correlation. Two specific models which have advantages over other methods are :

1. Random effects models - regarded as performing well for quantitative outcomes and event rate data, but not as well for analysing proportions
2. Generalised estimating equations (GEE) – these take account of correlations between observations within the same cluster, but do not explicitly model for variation between clusters. This is regarded as a more robust approach than the random effects regression model as fewer assumptions have to be made. Where GEE is fitted for quantitative outcomes (or events-rate data), an exchangeable correlational matrix is assumed and robust standard errors obtained. The Wald test is used to carry out significance tests (Hayes and Moulton 2009).

Regression analysis based on individual-level data uses the same GEE and random effect models as presented above, but additional fixed effect parameters are included in the strata.

3.3.5.3 Intention-to-treat versus per-protocol analysis

Intention-to-treat (ITT) is regarded as the ‘gold standard’ when determining the sample of participants to be analysed. This means that study groups are analysed based on the study arm to which they were originally allocated regardless of what treatment they actually received, or any protocol deviations that might have occurred thus reducing sample bias. ITT is acknowledged as

giving a conservative estimate of the treatment effect, compared with what would have occurred if there had been full compliance (Heritier, Gebski *et al.* 2003).

An alternative strategy to ITT is 'per protocol' analysis. This means that only those participants who are deemed to have sufficiently complied with the trial protocol are included in the analysis. This means that a number of participants could be excluded for any number of protocol violations, and in doing so there might be a certain amount of bias introduced.

3.3.5.4 Addressing bias

Bias is more of a concern in CRTs than in clinical trials (where the individual is the unit of assignment) because there are fewer units (clusters) to randomise, therefore increased attention needs to be paid to design strategies that will reduce possibilities of bias (Murray 2001).

Four primary sources of bias that can affect the validity of the design of a group randomised trial are :

1. selection bias
2. bias due to differential history
3. bias due to differential maturation, and
4. bias due to contamination (Murray 2001).

The first three are more likely when there is non-random assignment or insufficient numbers of groups per condition / arm. These sources of bias can be avoided by randomisation of a sufficient number of groups for each arm / condition of the study (as this increases the likelihood that potential sources of bias are distributed evenly among the conditions being studied) and careful matching / stratification which increases the effectiveness of the randomisation especially when the number of groups is small. Murray (2001) cautions that where the number of groups is less than 20 per arm, careful matching or stratification should be done before randomisation.

Contamination occurs when intervention-like activities appear in the comparison groups thus favouring the null hypothesis. This cannot be addressed through randomisation.

3.3.6 Reporting requirements

In order to address concerns regarding biased estimates of treatment effects associated with inadequate reporting and design, particular requirements for analysis and reporting of RCTs are contained in the standard CONSORT (CONsolidated Standards Of Reporting Trials) statement (Moher, Hopewell *et al.* 2010, Schulz, Altman *et al.* 2010). The purpose of this reporting standard is to facilitate critical appraisal and interpreting of RCTs. The statement covers each stage of the design and reporting process, indicating what requirements are necessary for each type of study. CONSORT has been extended for CRTs (Campbell, Piaggio *et al.* 2012) and pragmatic trials (Zwarenstein, Treweek *et al.* 2008). Specific requirements for reporting CRTs are the inclusion of a rationale for adopting a CRT design, describing how the effects of clustering were incorporated into sample size calculations and analysis, and describing the flow of both clusters and participants through the trial from assignment to analysis (Campbell, Donner *et al.* 2007). Specific requirements for the reporting of pragmatic trials are a clear description of the setting, population, providers of health care, the health service and its financing, where applicable; description of resources added to or removed from the usual settings; details of the intervention and the comparator; relevance of the outcomes chosen and follow-up period; an explanation and motivation for the use of blinding or not; and key aspects of the setting, population or intervention that might have determined the results (Zwarenstein, Treweek *et al.* 2008).

3.4 Considerations for the design of the intervention

This section will outline the factors considered in designing the intervention, i.e. context and understanding the problem, participatory process, behaviour change, and implementation strategy. The details of the design and implementation, and choices made will be presented in 3.7.4.

3.4.1 Context and understanding the problem

This study was located in a low-to-middle income country where there were resource constraints and high demands for health care. The location of the sites were a considerable distance from any urban centre. This factor had an impact both on accessibility to formal or informal educational opportunities for the midwives, and on clinical decisions that needed to be made based on anticipated transfer times. The perceived shortage of staff meant that continuing education was regarded as a luxury. The lack of regulatory requirement in South Africa for continuing professional development for midwives and nurses meant that an important driver for midwives and health service managers to actively seek or supply opportunities for ongoing learning was absent. In some settings the registered midwife was the only qualified midwife on duty for any shift and felt isolated. Added to this, the practice of rotating staff through all departments including maternity meant that there was a continuous change, uncertainty, lack of consolidation of skills and lack of capacity building.

From the various national reports and audits on maternal and neonatal mortality presented in 1.3.2 it was clear what problems were common in the intrapartum setting – lack of appropriate monitoring, failure to act timeously on abnormal findings, low use of the partograph, fetal hypoxia – and what some of the proposed solutions were – all centres offering labour care to use a partograph and to institute an audit of its utilisation (Department of Health 2006a).

3.4.2 Participatory process

It was clear that any intervention could not be aimed at an individual but needed to strengthen the institutional system. However, if an intervention was imposed it could have met with resistance at various levels. The key role players that were identified included regional MCWH managers (who had invited the study to be located in their regions), nurse managers of all the sites in the two regions identified for the intervention, the provincial maternity guidelines working group

(which included the regional obstetricians of these two regions) and the regional paediatricians / neonatologists.

The nurse managers of all the sites were invited to meet to hear a presentation about the proposed study and to share their concerns, indicate their willingness to participate / co-operate (or not). From these meetings input into the design and implementation were made, including the criteria for identifying appropriate clinical facilitators, how this would be recognised as part of their operational responsibilities rather than as an added burden, as well as agreement to support the project and each other during the project.

3.4.3 Behaviour change

The theory of behaviour change (Rollnick, Mason *et al.* 1999) identifies the stages of readiness for change that exist. Generally there was concern amongst the nursing managers at site level about the standard of care in labour, and they were glad that this could be addressed. Given that their role was to be supportive rather than having to arrange, organise and manage an education programme relieved them of that responsibility and resulted in their being amenable to play the role of clinical facilitator / supporter. Using the 'Stages of Change' model developed by Prochaska and DiClemente (1982), they all arrived at the preparatory stage and were ready for the action stage.

3.4.4 Implementation strategy

From the literature already presented, it appeared that any intervention would need to be a package rather than a single intervention, and that this intervention should be aimed at individuals and the institution. The intervention should include a multi-method educational programme with various components of educational outreach supported by some form of audit and feedback, and should involve multiple interactions.

3.5 Considerations regarding choice of Indicators

Focussing on the impact of intrapartum care, there are numerous factors that could result in poor clinical management, and many potential consequences of poor clinical decision making. The 'hard' indicators that are used and available usually relate to mortality. Morbidity or 'near-miss' data are less specific and open to a wide range of interpretation. The potential for an effective audit is limited if the quality of record keeping is poor. Process indicators can assist in determining what will create the necessary improvements which will ultimately result in decreased mortality and / or morbidity.

Unless one is able to be a participant-observer it is not likely that one would be able to objectively measure the actual quality of care being given or accuracy of the observations, without influencing the data itself as midwives and patients would be constantly reminded of this observation process. One is therefore dependent on proxy indicators for standard and quality of care.

3.5.1 End points

Appropriate endpoints for measuring the effectiveness of an intervention on intrapartum care would normally be maternal and / or perinatal or neonatal mortality. However, maternal deaths are less common than perinatal deaths, thus the maternal mortality ratio (MMR) is less responsive than the perinatal mortality rate (PNMR) as an indicator of differences and trends in quality of care (Graham, Fillippi *et al.* 1996). Further it is suggested that there are problems with validity, specificity and reliability of PNMR as an indicator of quality of care at an institutional level, especially in developing countries where there may be a low level of institutional births (Mancey-Jones and Brugha 1997).

Such endpoints are historic measures and would require large numbers of deaths to occur before one would be able to determine a statistically significant change. This was beyond the capacity of this project, as 140 000 – 250 000 births would have been required in order to have sufficient power to determine a significant change in terms of neonatal indicators.

3.5.2 Proxy indicators

3.5.2.1 Intrapartum care

Some neonatal indicators reflect the quality of intrapartum care. Where such indicators are not feasible or the data not credible, process indicators of intrapartum care can be considered.

3.5.2.1.1 Neonatal indicators

Several neonatal indicators were considered, viz. the incidence of intrapartum hypoxia or the Apgar scores at 1 and 5 minutes post birth, and the need for neonatal resuscitation.

Hypoxia in labour resulting in ischaemia has been presumed to be the main cause of neonatal encephalopathy, particularly in full term non-dysmorphic infants. However, there is often an absence of evidence of severe intrapartum asphyxia in infants with neonatal encephalopathy, and conversely many infants who do have signs of fetal distress and asphyxia do not develop neurological sequelae (Cowan, Rutherford *et al.* 2003). This has opened up considerable discussion in the literature as to the cause of this condition and questions the view that most risk factors for neonatal encephalopathy lie in the intrapartum period. Some of the pathways are antenatal in origin with infection and hyperthermia playing a role (Badawi, Kurinczuk *et al.* 1998a), some relate to postdates and prolonged second stage of labour (Fawcus, van Zyl *et al.* 2004), other significant associations are intrapartum maternal pyrexia, a persistent occipito-posterior position and an acute intrapartum event (Badawi, Kurinczuk *et al.* 1998b), non-cephalic presentations and prolonged rupture of membranes (Ellis, Manandhar *et al.* 2000). Neonatal encephalopathy can be regarded as a valid indicator for intrapartum hypoxia and, by extension, intrapartum monitoring and care. Although intrapartum hypoxia is one of the leading causes of death in the perinatal period in the Western Cape Province, it has a relatively low incidence and therefore an alternative indicator to neonatal encephalopathy would be helpful for monitoring the quality of care.

Apgar scores could be used as an indicator (Ellis, Manandhar *et al.* 2000, Young, Hamilton *et al.* 2001, Smith and Kirsten 2004). However, there was

doubt as to the accuracy and reliability of the scores given to newborns in the services studied and therefore this could not be utilised as a stand-alone indicator for this study.

The need for neonatal resuscitation was considered as a possible indicator. However, the wide range of understanding within the context of the study sites of what constitutes neonatal resuscitation could give a high false positive record.

3.5.2.1.2 *Format and utilisation of the partograph*

The partograph is a graph used to record the monitoring of the progress of a labour, where the key features of labour (maternal condition, fetal condition, progress of labour, treatment given, management decisions) are entered. The graphical representation enables the relationship between the various aspects of the labour process and the impact of treatment on these processes to be monitored. The original intention for introducing the partograph was to identify and prevent prolonged labour by use of alert and action lines applied to the cervicograph. The alert line is plotted at a rate of 1cm per hour (representing the slowest 10% of cervical dilatation of primigravid women), and the action line was originally plotted at four hours after the alert line (Philpott and Castle 1972a, Philpott and Castle 1972b). The four-hour action line was evaluated when the WHO carried out a large multicentre trial of 35 484 women in South East Asia, and the widespread use of a partograph with four-hour action lines was recommended (Kwast, Lennox *et al.* 1994). Various modifications in different settings have set the action line at two or three hours after the alert line (Lavender, Alfievic *et al.* 1998). More recently there has been a suggestion to remove the latent phase of labour from the partograph (Kwast, Poovan *et al.* 2008).

The partograph is merely a recording instrument. Its value is dependent on accurate observations, completion in full, and effective application of knowledge and skills in order to make the necessary conclusions which affect clinical decision making, and appropriate guidelines for intrapartum clinical management. The correct utilisation of the partograph is recognised by the

World Health Organization as a key strategy for effective management of labour and timely referral of developing complications (World Health Organization 1989, World Health Organization 2003, Fesseha, Getachew *et al.* 2011), and as such has been regarded as the standard of care for over 30 years.

In a Cochrane review on the effect of partogram use on outcomes for women in spontaneous labour at term, Lavender, Hart *et al.* (2008) identified five suitable studies – two of which enabled comparison of outcomes in terms of introducing the use of a partogram versus routine care without a partogram, and the other three which reported on the impact of different designs of the partogram. Comparing the use of the partogram versus routine care without a partogram, the evidence was inconclusive as this portion of the review was limited to two studies, described as differing in methodological quality. Lavender, Hart *et al.* (2008) conceded that using partograms in settings with limited access to healthcare resources may be beneficial, as there was some reduction in caesarean section rates, and early intervention for delayed progress in labour where the partogram was used. However, as the use of the partograph has been regarded as the standard for intrapartum care, the burden of proof is to show that not using the partograph results in better outcomes than using it. The evidence does not show this.

Health workers must be motivated and able to utilise the partograph. Studies report various levels of utilisation of the partograph in settings where it is meant to be part of routine use, from 25% (Ith, Dawson *et al.* 2012) to 70% (Ogwang, Karyabakabo *et al.* 2009). In a study which took place in Pikine, Senegal over a four month period involving 1022 labours recorded on a partograph, the authors showed that regular, constant supervision is needed even when the partograph has been used for years (Dujardin, de Schampheleire *et al.* 1992).

3.5.2.1.3 *Suitability as a proxy indicator*

The utilisation of the partograph serves as a proxy indicator for the standard of care with the assumption that the correct and complete record of observations on a partograph will facilitate appropriate clinical decisions, which should result

in a better standard of care and better health outcomes. The partograph is recognised as a useful instrument for monitoring and assisting decisions in managing labour. Crossing the alert line is significant as an indicator for hypoxia, the need for neonatal resuscitation, and fresh stillbirth, and as a proxy indicator for the standard of care in labour. In the Pikine (Senegal) study, Dujardin, de Schampheleire *et al.* (1992) found that among the women who crossed the alert line only, neonatal resuscitation was four times higher than in normal labour. For those who crossed both the alert and action lines the fresh stillbirth rate was 10 times higher than for women in the normal labour group. The results show the usefulness and efficacy of the partogram and underscore the value of medical intervention as soon as the alert line is crossed (Dujardin, de Schampheleire *et al.* 1992). The Pikine study showed that alert and action lines display sensitivity, specificity and positive predictive value to the need for neonatal resuscitation (fetal / respiratory distress) and fresh stillbirths respectively (Dujardin, de Schampheleire *et al.* 1992). This underlines the link between intrapartum care, prolonged labour and neonatal outcome and thus is a suitable proxy measure.

3.5.2.1.4 *Some debates about the use of the partograph*

Some concerns have been raised as to the applicability and transferability of the partograph across populations and its relevance in high resource settings (Groeschel and Glover 2001).

There is evidence that there are ethnic differences in the progress of labour. The differences in pelvic size and capacity may account for the higher rates of cephalo-pelvic disproportion found amongst Black African women (Steyn, Evans *et al.* 1994, Farrell 2005). It was in this population that Philpott conducted his work on the partograph and thus it is suitable for use in a southern African population.

Concern has been expressed that partograph use results in higher levels of unnecessary intervention, and that this may restrict the autonomy of clinical practitioners (Lavender and Malcolmson 1999). Lavender, Hart *et al.* (2008)

found no difference between the rate of caesarean sections and instrumental deliveries between the two groups (monitoring labour with or without a partograph) thus refuting the concern that unnecessary interventions flow from the use of the partograph. A great concern is that insufficient numbers of women have access to the necessary interventions where maternal and neonatal mortality is high (World Health Organization 2009).

The utilisation level of the partograph could serve as a process indicator. A key recommendation in the 'Saving Mothers' reports was the correct use of the partogram should become the norm in each institution conducting births and quality assurance programmes should be implemented using an appropriate tool (Department of Health 1999, Department of Health 2002, Department of Health 2006a). In the 2002-2005 triennium report the following indicators and targets were recommended :

- ◆ Percentage of institutions conducting births that use the partogram – 100%
- ◆ Percentage of institutions having a quality assurance programme – 80%
- ◆ Percentage²² of institutions scoring satisfactory [for use of the partogram] on the quality assurance programme (Department of Health 2006a).

What is not stated is what constitutes 'satisfactory'. If the utilisation of the partograph is to become the norm then one could assume that this level would be set in the upper quartile range.

3.5.2.2 Midwifery knowledge and skills

For the secondary outcome, it was assumed that knowledge and skills are required for good clinical observation and assessment in order to enhance clinical decision making and management. Thus the legal framework governing midwifery practice in South Africa was used as the point of reference for this study for what could be expected from the registered midwife (1.6.1). However, it is acknowledged that an increase in knowledge does not necessarily result in a change of behaviour, i.e. clinical management.

²² No target specified.

3.6 Instrumentation – partograph utilisation checklist

Having determined that there was no suitable instrument available to conduct an evaluation of partograph utilisation (see 2.4.9), an instrument was developed, validated and tested for reliability amongst the study population.

3.6.1 Developing and validating the instrument

The validation process was guided by that described by (Lynn 1986) for instrument development and consists of a developmental stage and a judgment stage. The developmental stage, underpinned by a literature search and consultation with experts, consists of :

- ◆ domain identification. As the checklist was to evaluate the utilisation of an existing clinical tool (partograph) which itself has been extensively researched, many constructs were predetermined. The domains were bound by the content of the partograph itself. Two domains were added, viz. ‘medico-legal considerations’ and ‘adherence to guidelines’ as well as a project administration section.
- ◆ sampling and item generation where a checklist was compiled. Using the Philpott and Voce labour record review tool as a starting point (Philpott and Voce 2005), the researcher compiled a checklist in a logical sequence following the ordering on the partograph as closely as possible. The checklist was sent to expert users for suggestions and recommendations.
- ◆ instrument formation - where the checklist was arranged in a usable form. This was completed by the researcher.

The judgment stage consisted of establishing a credible jury (expert panel) and the judgment process. The jury should be involved in content validation to enable the determination of the content validity index (CVI) (Lynn 1986). The jury should have a diversity of professional backgrounds and relevant expertise in order to get as rich a set of responses as possible. Lynn indicates that if there are five or fewer judges there should be consensus whereas if there are six or more, one may disagree with the content validity assessment. Ten people were appointed to undertake the judging process so that the threshold of six could be

guaranteed. Seven judges participated in all three rounds, and an eighth judge participated in two rounds (having missed the first). The jury included clinicians, researchers and health service managers in intrapartum settings across the urban/peri-urban/rural spectrum and across provinces²³.

The judgement process required a structured procedure for evaluation for content validity of the instrument to be given to the experts (Kerlinger 1973, cited in Lynn, 1986), and this included appropriateness, accuracy and representativeness of the specifications (Yaghmaie 2003, Grove, Burns *et al.* 2013). Using the provided guidelines the jury rated their agreement on the Content Validity Index (CVI) instrument using a four-point Likert scale from 'not relevant / least agreement' (1) to 'very relevant and succinct / very strong agreement' (4). The current version of the partograph evaluation checklist²⁴ together with the validation instrument, and response form (**Appendix 4**) were sent to the identified experts for content judgment. There was an opportunity for more detailed comments and suggestions. The judges commented on the instrument's contents until consensus was reached. After each round the researcher compiled a consensus document and revised partograph evaluation checklist, and highlighted areas that needed further consideration, e.g. ambiguity, correctness and completeness, and scoring, as well as setting out the different / divergent opinions that had been expressed in the previous round. A detailed set of explanatory notes were developed to accompany the instrument. In addition to the item scoring, the instrument as a whole was scored (quantity judgement). If deletions or additions were proposed, then the instrument as a whole was rescored. This ensured that any changes proposed would retain clinical coherence / relevance of the checklist.

²³ Four advanced midwives, two of whom have extensive teaching experience and the other two manage large programmes; three obstetricians, one of whom leads an academic research unit; two medical officers with responsibility for maternal and child health; one public health practitioner/researcher.

²⁴ This is not included in the appendices. Each round of judging had a modification of the partograph utilisation checklist from the previous round. Including all versions could cause some confusion between the various versions 'in process' and the final version used as the testing instrument.

Unlike other forms of validation, which are based on the scores from a scale, differences in performance between people or changes as a result of an intervention, content validation is based on the judgment of experts regarding the content of items. Lynn recommends that where there are more than five judges, consensus is not required, but no more than one should score less than three in order for the CVI to be statistically significant at the 0.05 level (Lynn 1986, Streiner and Norman 2008).

After round three, seven of the 13 items had all judges scoring three or above, five of the 13 items had only one judge scoring below three; and one of the 13 items received four out of eight scores above three, which did not meet the standard. This was the item relating to evidence of previous studies. A précis of the sparse evidence had been summarised in covering letters to the judges. (Some judges stated that they were not aware of other studies addressing this specific issue.) Thus the required level for agreement amongst the judges was met for face and content validity (Lynn 1986).

The partograph utilisation checklist consisted of the following sections :

◆ Identification, history and risk assessment	8 items
◆ Fetal condition	6 items
◆ Labour progress	6 items
◆ Maternal condition	8 items
◆ Comprehensive assessment and management	5 items
◆ Alignment with clinical management guidelines (including frequency of observations)	15 items

Each item had only one characteristic. Each item was scored as 1, ½ , 0 or N/A (not applicable) using the detailed guidelines for scoring (see scoring notes and explanatory notes attached to the partograph utilisation checklist in **Appendix 5**)²⁵.

²⁵ This appendix contains the finalised version of the partograph utilisation checklist after testing which differs slightly from the version used to test inter-reader reliability.

3.6.2 Testing of the instrument in this population

3.6.2.1 Population, sample size and sampling

3.6.2.1.1 Population

The population consisted of partograph records of singleton pregnancies greater than 36 weeks gestation where cervical dilatation was ≤ 6 cm on admission to the labour ward.

3.6.2.1.2 Sample size determination

A categorical scale was used and the sample size was based on using the kappa statistic to assess reliability. Using NQuery (Elashoff 2000), the sample size of 324 partographs was based on the following : the proportion of successes (partograph completion) = 70%; level of significance = 5%; power = 80%; 2-sided test. A null hypothesis for the intraclass kappa test with kappa = 0.55, against an alternative kappa of 0.7, led to the indicated sample size. The sample size was rounded to 330 and stratified by categorisation of quality of completeness.

3.6.2.1.3 Sampling

In order to have a spread of quality of completion of partographs, a recognised expert (obstetrician) in the field of intrapartum care categorised the partographs into one of three categories based on his assessment of the quality of completion (poor, average and above average), without reference to any instrument. As the partographs were coded, he was blinded to their origin, and the order in which he received them for categorisation was not related to site or study arm. Once 110 partographs were selected for each category the categorisation stopped. This resulted in a stratified sample in terms of completeness. These were then reshuffled prior to analysis so that there was no obvious grouping.

3.6.2.2 Evaluation by two midwives (inter-reader reliability)

The two midwife evaluators for this process were similar in terms of their educational preparation, years of clinical experience and management responsibility in primary or secondary maternity units. Twenty records were used for the pilot study, the purpose of which was to clarify any uncertainties

about interpretation of the instrument, the instructions and the application to actual clinical records. These records were not used for the inter-reader reliability testing. Thereafter all 330 partographs were evaluated.

3.6.3 Analysis to determine concordance

Four processes were used to determine concordance – Lin's concordance correlation coefficient (Lin 1989, Lin 2000), Bland and Altman's limits of agreement (Bland and Altman 1986), correlation between the difference and the mean (Pitman 1939), and the F test of equality of means and variances (Bradley and Blackwood 1989).

3.6.3.1 Lin's concordance correlation coefficient

Concord computes Lin's concordance correlation coefficient (CCC) for agreement on a continuous measure obtained by two persons or methods (Lin 1989, Lin 2000). Lin's CCC increases in value as a function of the nearness of the data's reduced major axis to the line of perfect concordance (the accuracy of the data) and of the tightness of the data about its reduced major axis (the precision of the data).

3.6.3.2 Bland and Altman limits of agreement

Bland and Altman's limits-of-agreement (LOA) procedure (Bland and Altman 1986) is a data-scale assessment of the degree of agreement and is a complementary approach to the relationship scale approach of Lin.

3.6.3.3 Correlation between the difference and the mean

The correlation between difference and mean. As an explorative diagnostic, a value near zero implies concordance (Pitman 1939). Using NQuery, based on the sample size of 330, a correlation ≥ 0.14 is significantly different from 0 with 80% power using a 1-sided test. Thus for interpretation of this test a correlation ≤ 0.1 is regarded as no correlation (Elashoff 2000).

3.6.3.4 F test of equality of means and variances

This test assumes bivariate normality (Bradley and Blackwood 1989). Non-significance implies concordance.

3.6.3.5 Results and interpretation (inter-reader reliability)

The data revealed a Lin's CCC of 0.817 indicating good concordance. Bland and Altman's limits of agreement showed that there was small mean difference (-1.52) reflecting that one reader read consistently higher than the other but the

limits of agreement at the 95% confidence interval was wide. The correlation between the difference and the mean was $-0.24(p<0.001)$ indicating weak correlation. When plotted graphically the differences were found to be at the extreme upper and lower scores. The F test of equality of means and variances gave an F-value of 34.63 ($p<0.001$) representing non-concordance.

In summary, the partograph scores of the two readers using the partograph utilisation checklist showed a high Lin's concordance correlation coefficient (0.817). However there is some evidence that concordance is not perfect due to differences in extreme upper and lower scores.

3.6.4 Evaluation by multiple midwives (inter-reader reliability)

A third midwife (the researcher) evaluated the new instrument twice using the same 330 partographs to test for intra-reader reliability (see 3.7.6.1.2). This provided two more sets of readings and an opportunistic analysis was performed to determine to what extent the intra-rater readings matched the readings from this portion of the study.

3.6.4.1 Analysis - Estimating difference and evaluating agreement

3.6.4.1.1 Estimating difference

The General Estimation Equation (GEE) model of the partograph score on the reader was done. The analysis was done by clustering the readers within each partograph resulting in 328 clusters with a size of 4 records, and two clusters with a size of 2 records (due to the fact that there were two incomplete pairs in the instrument testing.) (The GEE model is sufficiently flexible to be able to accommodate such differences.) For the purpose of this analysis the two readings by the same reader (SC) were combined and the mean score used. (These scores had a high degree of correlation – Lin's CCC = 0.957)

3.6.4.1.2 Evaluating agreement

A mixed effects linear regression model with reader and partograph as crossed variance components (random effects) was fitted. The variance components were estimated and the intraclass correlation coefficients (ICC) for reader and partograph were calculated.

3.6.4.1.3 Results and interpretation

Differences between readers were small, but are statistically significant for each pair of readers (Table 3.1). (This is achieved due to large number of clusters.) However, the range of bias between the three pairs of readers is small (0.6 - 2.1 of a unit score out of a possible total score of 48, i.e. 1.3% - 4.4%) and the confidence intervals are narrow. The largest difference between readers (2.1 units) would result in a 4.4% difference. As this tool is not used for diagnostic purposes, this range of difference (<5%) is regarded as acceptable for research purposes as well as for a clinical audit in this setting.

Reader	Contrast	Std error	95% confidence interval		P> z
LK vs CA	-1.5	0.2	-2.0	-1.1	<0.001
SC vs LK	-0.6	0.1	-0.9	-0.4	<0.001
SC vs CA	-2.1	0.2	-2.6	-1.7	<0.001

Table 3.1 Differences between three readers

The ICCs were determined showing the ICC for the partograph was 0.93 and for the readers 0.014. This means that 93% of the overall variance of the partograph scores can be attributed to between-partograph variability, and only 1.4% of the variance is between readers, which is very small (ICC=0.014). Thus, although readers have significant but small biases between them, the intraclass correlation shows that they have good agreement overall.

3.6.5 Conclusion

In conclusion, the partograph utilisation checklist has a good concordance. The analysis confirms that for this population, and for this level of evaluator, the modified partograph evaluation checklist will deliver reliable measures.

3.7 Research methods used for this study

The preceding sections of this chapter have presented the various options that were considered when designing this study as well as the validation and testing of the instrument to be used to measure the primary outcome. The section that follows presents the decisions made (see Figure 3.1). The background explanations for each aspect have been presented in 3.3 above.

3.7.1 Design rationale and description

A pragmatic cluster randomised trial was used (where the unit of randomisation was the site / cluster) to evaluate the impact of an intervention package (an on-site mentorship programme and the training of registered midwives) on the management of labour. The points of interest for this study were sites where intrapartum care was offered and the groups of professionals linked to these sites. It was assumed that it was highly likely that there would be higher degrees of similarity within sites than between sites. The intention was to evaluate effect of an intervention on the standard of intrapartum care (which would be delivered by a site rather than an individual). In such a situation a CRT was an appropriate design.

The unit of assignment was the institutional site in which births take place. These were stratified into two strata based on geo-political region in which the sites were located and size of service defined by the number of births *per annum*. After the strata were assigned, clusters were randomly allocated to either the intervention or control arms of the study. The detail is described in 3.7.3.3. The units of observation were the partograph records and the midwives in the institutional groups. An individual level approach to analysis for primary and secondary outcomes was selected as it offered a more efficient approach (see 3.3.5.1).

3.7.2 Setting, sites and population

The population included all primary level public health sector facilities in two rural regions of the Western Cape Province, viz. Boland / Overberg and West Coast / Winelands, which offered intrapartum care (**Appendix 1**). Clusters /sites were based on district boundaries. In this study, institutional sites were the identifiable groups with a degree of association assumed due to institutional norms; assignment of clusters was to the control or intervention arms; the units of observations were doubly nested - the record of observation and management of labour, and the midwives' knowledge and skills within the institutional cluster, and across the intervention and control arms; and there were 17 clusters – eight in the control arm and nine in the intervention arm.

The sites consisted of district hospitals, small hospitals, and midwife obstetric units (MOU), each of which constituted a cluster. There were 214 registered nurses²⁶ employed in the study sites. However, a number of them would not be allocated to intrapartum settings (e.g. not registered as midwives, preference for other clinical areas like operating theatre), and therefore were not considered for the study. There were 19 327 births in these two regions during the data collection period (2006-2007), of which 13 497 occurred in the study sites (Provincial Government of the Western Cape Department of Health Directorate of Information Management personal communication 2012).

The motivation for using these two regions was threefold :

1. The aim of the study was to evaluate the performance of registered midwives. The population of pregnant women using primary level facilities are regarded as being low risk. There is an expectation that any complications that occur in labour will be referred to a secondary or tertiary level facility as required. Therefore, the practice of midwives is more likely to be exercised more fully at this level than at any other level of the health service.
2. Registered midwives working in rural regions are appointed to registered nurse posts and are expected to work in a variety of clinical settings, whereas in the metropolitan region, registered midwives are appointed to specified registered midwife posts. It is therefore necessary to ensure that the practice being evaluated is similar in both regions identified for this study. Further, in a preliminary study it had been shown that the knowledge and skills of registered midwives in the rural regions tended to be weaker than those of their metropolitan region colleagues (Clow 2006).
3. There are 3 rural regions in the Western Cape (described in detail in 1.4). For the purpose of this study the Southern Cape / Karoo region was not considered due to the limited capacity to conduct the intensive

²⁶ Despite numerous attempts via numerous channels to identify the number of posts requiring registration as a midwife, this information could not be provided by the Provincial Government of the Western Cape.

interventions required. (The ethical implication of this decision is discussed in 3.8.6).

All sites received the relevant clinical guidelines which were developed by the provincial maternal, child and women's health sub-directorate. *Ad hoc* workshops were arranged by the regional maternal child and women's health co-ordinator. All sites were invited to send participants but this was dependent on the capacity of the service to release someone. This activity predated this study by a year.

Outreach activities are carried out by the regional obstetrician. There was no difference in terms of availability of policies, procedures, guidelines, and outreach activities between control and intervention sites. Neither of the obstetricians was aware of what the allocation of sites was in this study.

During the study period, the co-operation of the regional co-ordinators for maternal and child health was obtained so that there was no in-service training specific to intrapartum care directed to any of the sites. Any in-service training programmes were directed to entire regions.

The presence of the researcher did not reveal the allocation of the site as data was being collected for both arms, and her visits did not coincide with those of the regional obstetricians. There was no line function between these clinicians, the MCWH co-ordinator or the researcher, which could reveal the allocation. There was no documentation in the sites that would identify them as intervention or control sites.

3.7.3 Sample size, sampling and randomisation

The unit of randomisation was the site / cluster. The size of the sample was co-determined by the two units of analysis, *viz.* standard of clinical practice, and registered midwife knowledge, analysis and skill.

3.7.3.1 Clinical practice (Partograph utilisation audit)

In a preliminary study it was shown that the baseline score on the use of the partograph in the rural regions was 63.5% (Clow 2006) with a standard deviation of 10.8.

Using ACluster software (World Health Organization 2000) it was established that in order to detect a change of 5% (approximately half the standard deviation in the preliminary study), with a cluster size of 60, an interclass correlation (ICC) factor of 0.05 and a power of 0.9, it would be necessary to have a total of 406 records in each arm of the study which amounted to seven sites per arm. Using all 17 sites would thus be appropriate and provide insurance for the inference of the study.

The following were considered in setting the ICC at 0.05. This level is regarded as a modest correlation. The constitution of the clusters was population-based and therefore with the larger population one would anticipate a greater degree of variation. Generally the ICC adopted in health services research is 0.01 – 0.05 (Adams, Gulliford *et al.* 2004), and so this level was used in the calculations for the sample size. Campbell, Fayers and Grimshaw (2005), in their study of determinants of the intra-cluster correlation coefficient in cluster randomised trials, found that ICCs for process indicators were higher than for outcome variables, i.e. median value of 0.063 for process variables compared to median value of 0.03 for outcome variables, and that this was significant $p < 0.001$. The value chosen for this study was similar to the ICC used in another South African study in a primary level clinic setting assessing the effect of educational outreach to nurses on case detection for lung disease (Fairall, Zwarenstein *et al.* 2005).

Inclusion criteria : All labours at term (≥ 36 weeks) where monitoring was commenced at any of the sites where the cervical dilatation on admission in labour was ≤ 6 cm. Given that a woman in preterm labour should be referred to a higher level of the service and preterm labour might involve other pathologies

which could effect the management of labour, preterm labours were not included in the study population. Records of the labours of women who were transferred during labour were followed up at the secondary institution to ensure that a birth took place from that admission, but only the monitoring done at the primary site was evaluated.

Exclusion criteria : Records of women who had an elective caesarean section.

Stopping rule : This was not a treatment study and no adverse events were anticipated. There was no interim analysis planned or done. There was no stopping rule.

Sampling : One thousand and twenty partographs were systematically sampled by the researcher for the 12 month period following the intervention. The sequence of partographs at each site was established based on the time of admission to the labour ward. The ratio n was calculated as the total number of deliveries for each month in that site divided by the number of partographs required for that site for the month. In each site, the first record of the sequence was randomly selected from 1 to n using a computer-generated table of random sequencing, and thereafter every n th record was selected.

3.7.3.2 Registered midwives' knowledge and skills

In a preliminary study, the overall baseline score obtained by the registered midwives in the rural regions was 49.5% (Clow 2006) with a standard deviation of 11. Using ACluster software (World Health Organization 2000) it was established that in order to detect a change of 10% it would be necessary to include 100 registered midwives [$\alpha = 0.5$; power 0.9; interclass correlation coefficient 0.1].

The following were considered in setting the ICC at 0.1. This level is regarded as a reasonably substantial correlation and implies that 10% of the overall variance is due to the between-cluster dynamics. In this study there were generally small groups of midwives constituting a cluster. Within these clusters

one assumed a high level of agreement, similar approaches to clinical management where there is peer pressure to conform to institutional norms (which may be different between institutions), and a common management style. Generally the ICC adopted in this type of study is 0.01 – 0.05 (Adams, Gulliford *et al.* 2004). However, health care providers tend to be more correlated than the general population due to their smaller population size.

Given that the average number of registered nurse / midwives per site was estimated as 13, eight sites (i.e. four control and four intervention) would be required to achieve a sample of 100 midwives. Given that there was a large variation in the numbers of registered midwives per site (6-26), and this margin would be fragile if one site should drop out, it was decided to use all sites available (17).

Inclusion criteria : All registered nurse / midwives who could be expected to work in an intrapartum setting. Staff who joined after the training intervention were also invited to join the study and participate in the post-intervention evaluation, as this would also reflect the impact of the mentorship intervention.

Exclusion criteria : Any registered nurse / midwife who left the service during the post-intervention monitoring period was not included in any evaluation that took place after leaving the service, but existing evaluations were retained in the study. The registered nurse / midwives who were mentors in the study were excluded from the evaluation as some of them were no longer working in the clinical setting.

Stopping rule : This was not a treatment study and no adverse events were anticipated. There was no interim analysis planned or done. Analysis was planned to be done only once all tests were completed to facilitate a standardised approach to the evaluations. There was no stopping rule.

Sampling : All eligible registered nurse-midwives were invited to participate in the study. Their allocation to study arm was based on the allocation of the site in which they worked.

3.7.3.3 Sites / clusters

Stratification was the strategy chosen for this study. It has a number of advantages over the matched pair design in that fewer degrees of freedom are lost, resulting in a higher power and precision; it is possible to test for variations in the intervention effect because there is replication of clusters within each stratum; it is possible to use regression methods to adjust for and analyse covariates; and if one cluster is lost it does not require loss of data from the entire stratum (Hayes and Moulton 2009). Hayes and Moulton (2009) recommend that this design be used even for trials with five clusters per arm, provided that potential matching factors that are likely to show some degree of correlation with the endpoint of interest are present.

In this study two strata were identified based on the geo-political region in which the sites / clusters were located and the number of births (deliveries) per annum (Table 3.2).

Size of site	Number of sites located in Region 1	Number of sites located in Region 2
Small number of births*	4	4
Moderate to large number of births*	5	4

*Births per annum: Small <500 Moderate 500 – 1000 Large >1000

Table 3.2 : Stratification framework

A number was assigned to each site / cluster within a stratum to conceal its identity. The intervention sites were randomly selected 'drawing from a hat' (Katzenellenbogen, Joubert *et al.* 1997). The additional 'moderate-to-large' site in Region 1 was purposively allocated to the intervention group in case there was attrition from the intervention sites. This process took place in the presence of the regional managers for Maternal, Child and Women's Health. Once the

process was complete the allocation was communicated to the managers of the various institutions who would facilitate the attendance of their staff at the training days. The nature of the intervention required the presence and active participation of registered midwives and their informed consent, and so the allocation would be obvious.

3.7.3.4 Comparator

Control sites continued to receive information, circulars, and guidelines from the region or provincial offices through the normal communication routes, and outreach activities and mortality and morbidity (M&M) meetings by the regional obstetrician continued as usual. Midwives working in the control sites received no new training. There were no limits placed on midwives having the opportunity of in-service training as this would have been discriminatory against the midwives, their services and their patients. However the control sites were requested to not hold any intrapartum-specific in-service training during the data collection period, and this was honoured. MCWH co-ordinators have responsibility for in-service training which focused on antenatal and neonatal aspects during the data collection period.

3.7.4 Intervention package

Against the backdrop of the local and national context outlined in 1.3, 1.4 and 3.4.1, and with the support of the nurse managers (who had agreed to facilitate the attendance of ALL midwives at the training days that would be held in various sites in their regions so that as far as possible saturation would be achieved), an intervention package consisting of two interventions were developed for this study. These were :

- 1) a clinical facilitator development programme for selected experienced midwife clinicians / managers. The intention was that the clinical facilitation in the sites flowing from this programme would provide a sustainable support structure to enhance clinical decision making and encourage a higher standard of clinical practice.
- 2) an educational update programme for registered midwives. This was included in the mentor programme.

This intervention package therefore took place at both cluster (1) and individual (1 and 2) levels.

This multi-method intervention was offered close to the location in which the midwives and clinical facilitators work. The midwives' programme included educational materials in the form of clinical guidelines²⁷, audiovisual aids and clinical models. The clinical facilitation programme also included audit and feedback, clinical guidance, adherence to clinical guidelines, facilitation of clinical decision-making, development of a learning environment.

The programme drew on a number of theoretical models – adult learning (Rogers 1969), emancipatory learning (Freire 1970, Shor and Freire 1987), scaffolding from the Zone of Proximal Development (Vygostky 1987) and motivational interviewing (MI) for behaviour change (Rollnick, Mason *et al.* 1999). While MI is usually applied to health behaviour change, it has been applied to health professional behaviour change in this study. This approach has relevance for both clinical facilitator and midwives' training. It was particularly relevant for the clinical facilitators as they continued the learning process that was started in the midwife training sessions. In order to increase motivation and decrease resistance, specific attention needed to be paid to the character and quality of communication. This is consistent with adult education approaches where one aims to have active participation in learning rather than passive top-down communication.

The educational intention was to ensure maximum integration of learning - content, skills, application and attitudes - by using various methods and strategies. These included working from familiar to less familiar, using the content as a vehicle for the method, using the method as a vehicle for the

²⁷ Guidelines for the use of the partograph developed by the provincial maternity guidelines reference group of the Western Cape Province were available as a printed handout and had been distributed through all the regions of the province. Clinical management guidelines developed by the provincial and national departments of health were available and had been distributed to all primary level facilities by circular as was the usual practice.

content, an outcome for each activity, participation and activity, clinical simulation, case scenarios, roleplay, reflection, use of triggers, and the development and repetition of key messages (Soumerai and Avorn 1990, Fullerton, Thompson *et al.* 2013). Knowledge, attitudes and skills were challenged and developed throughout using a variety of methods to accommodate the different learning needs in each group. Active participation was encouraged *inter alia* through group activities, problem-solving exercises, case-based learning, role play and quizzes. Peer learning and teaching was facilitated and where areas of disagreement or lack of knowledge were identified within the groups this formed part of the wider discussion.

The programme was facilitated by the researcher (an experienced midwife clinician and educator who had been working in this field for over 20 years) from a credible institution, who could be regarded as being 'educationally influential'. Delivery of the intervention was 'standardised' insofar as the same person facilitated all training sessions and the same materials and resources were used. However, the programme was responsive to the learners' expressed needs, and consistent with an adult learning approach.

This part of the intervention package was the only part that was added to the existing resources. No resources were removed from the services. The sites used were well prepared so that there was a state of readiness for the study (Rollnick, Mason *et al.* 1999). The interventions were intentional with the aim of establishing sustainable support mechanism practices to encourage good quality intrapartum care.

3.7.4.1 Clinical facilitator development programme

The scope of this programme was to develop capacity in identified individuals from each site to support, guide and strengthen the clinical practice of the colleagues in the intrapartum setting, and to develop skills for continuous quality assurance e.g. use of clinical audit. This role was already part of their job performance requirement, but they had not received training in this regard. Thus this study did not place demands on them that were additional to the

expectations of their posts. It was planned that this programme be offered to at least two clinical facilitators per site in order to foster support between the two. The regional MCWH co-ordinator for each of the two regions also participated, so that they could support and encourage the clinical facilitators through regular follow-up. This was part of the strategy for sustainability and ongoing support.

3.7.4.1.1 *Objectives*

The objectives for this programme were that, at the end of the programme, the clinical facilitators should be able to establish and maintain a learning environment, understand the purpose and role definition of a clinical facilitator, become familiar with the guidelines for intrapartum care, be able to supervise and monitor intrapartum care, be able to conduct a clinical audit of intrapartum care using an appropriate tool, be able to manage individual and group feedback, be able to conduct a clinical review meeting and develop jointly an acceptable strategy for ongoing support of the mentor group.

3.7.4.1.2 *Target group and selection*

It was planned that there would be at least two individuals for each of the intervention sites. These should be the nursing head of the site and a suitably experienced registered nurse-midwife clinician. Where there were only two registered nurse midwife clinicians at a site then both should attend to ensure that there was no sense of one being inferior or less worthy than the other. Operational managers selected clinical facilitators based on the following characteristics : willingness, regarded as a role-model, clinically skilful and experienced, a good communicator, passionate, reliable and stable, self-motivated, e.g. has attended perinatal updates, and had insight into the need for change.

3.7.4.1.2.1 *Preparation of clinical facilitators prior to development programme*

Suitable registered midwives / nursing managers who had responsibility for quality assurance were invited to attend the clinical facilitator development programme. Although this function is an operational requirement, the readiness of such a person to undertake this responsibility was regarded as key to the successful implementation of the programme. Once appropriate potential

clinical facilitators were identified for each site, and prior to the commencement of the programme, they were given the information sheet regarding the study so that they would be familiar with the expectations. (**Appendix 6** - Information sheet and consent form for mentors²⁸.) They were free to refuse to participate without penalty. No-one refused. However, if anyone had refused, an alternative person would have been identified for that site so that at the start of the training programme there would be a group that was willing to participate.

3.7.4.1.3 *Time period*

This was a four-day residential programme and preceded the registered midwife educational update training. There was one training programme in each region. It was necessary to offer a third clinical facilitator training programme halfway through the study period in order to reach the target capacity. This is discussed in 5.4.2.2.1.

3.7.4.1.4 *Programme outline*

This was developed co-operatively with regional co-ordinators, managers of maternity services and clinical experts. Cognisance was taken of the criteria and standards set by the respective standards generating bodies for two unit standards, viz. 'Mentor a colleague to enhance the individual's knowledge, skills, values and attitudes in a selected career path' and 'Conduct on-the-job coaching' (South African Qualifications Authority 2003, South African Qualifications Authority 2004).

The Philpott and Voce labour record review instrument (**Appendix 7**) (Philpott and Voce 2005) was used as a standard monitoring instrument for the clinical auditing function of the mentors²⁹.

²⁸ This was the term used at the time and this is reflected in the documentation. The term has since been amended to 'clinical facilitator' but the documentation is included using the original term.

²⁹ For the purposes of the audit and feedback aspect of clinical facilitation, the existing Philpott and Voce tool for auditing care in labour was used. Although this has been used nationally, it was recognised that this had not been validated nor tested for reliability, and there were aspects of this tool that could be strengthened, but it was adequate for this purpose. It was anticipated that once the development and testing of the partograph utilisation tool was completed, it would not be difficult to introduce the revised tool for future auditing and quality assurance.

The course planning and an outline of the programme with its multi-layered learning and teaching are included in **Appendix 8a**. The programme included an implementation planning exercise which included both group- and site-specific strategies. These were not pre-designed by the researcher but were a response of the clinical facilitators. These are reported in **Appendix 8a**.

3.7.4.2 Registered midwife educational update

The scope of this educational update was to develop the knowledge, critical thinking and clinical skills necessary to improve the management of the intrapartum period. The nature of the intervention meant that the participants could not be blinded to their allocation in terms of the arm of the study. However, after the initial training period, the researcher had minimal contact with the participants thus not creating a constant reminder that a study was being conducted.

3.7.4.2.1 Objectives

The objectives for this educational update were that the registered midwife should be able to make accurate maternal and fetal observations pertinent to labour, be able to record clinical observations correctly and completely on the partograph, demonstrate understanding of the process and progress of labour and of the clinical guidelines relating to the use of the partograph and managing labour, be able to apply knowledge to a variety of clinical situations, be able to interpret clinical data from the partograph and recognise when referral is necessary, demonstrate understanding of risk and developing risk in labour, and demonstrate mastery of clinical skills necessary for labour management, e.g. diagnosis of labour, abdominal and vaginal assessment, fetal monitoring and identification of fetal distress³⁰.

3.7.4.2.2 Time period

The educational update took place on two separate days approximately one month apart. Multiple interactions have been found to be more effective than single visits (Soumerai and Avorn 1990, O'Brien, Rogers *et al.* 2007, Ivers, Jamtvedt *et al.* 2012). There were two to three training sessions for each

³⁰ One might expect any registered midwife to be competent in these areas. A preliminary study conducted a year previously had identified these as areas of weakness.

training day in each region in order to facilitate full coverage of training of registered midwives who work on different day shifts as well as night shifts. These training days took place at the different intervention sites and the training days were open to all registered midwives in the training sites. Training saturation was achieved as far as possible, as the nurse managers for each intervention site rostered the attendance at the training days.

3.7.4.2.3 Programme Outline

The programme was a portion of the clinical facilitator training with a specific focus on intrapartum management. An outline is included in **Appendix 8b**.

3.7.4.3 Timing of intervention and data collection

The study was implemented in two phases to manage the logistics within the resources available. This data was collected over a 12 month period. The researcher was interested in establishing how robust and stable this intervention would be in a real-world situation. Such a time period would extend beyond a 'honeymoon period' of enthusiasm and would absorb the staff rotations across units and day/night shift allocations (within sites), and changes in the staff complement (including newly qualified registered midwives into the system) which would not be achieved by a shorter study period. A longer period would have become logistically difficult with changes in the health service and researcher capacity.

The timelines for this are reflected in Table 3.3.

- ◆ Region 1 commenced with baseline testing and the intervention in July 2006 with records for the partograph utilisation audit collected from August 2006 until July 2007.
- ◆ The corresponding dates for Region 2 were September 2006, October 2006 and September 2007.
- ◆ In the intervention arm a midwife test was done in month 3 for each region, viz. October 2006 and December 2006 / January 2007 respectively. This data was collected over a 12 month period.

	2006					2007											
	July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	March	April	May	June	July	Aug	Sept		
Region 1 / Month of study	0	1	2	3	4	5	6	7	8	9	10	11	12				
Mentor training	X																
Midwife baseline test – Test 1	X																
Midwife training commenced	X	X															
Partograph collection		X	X	X	X	X	X	X	X	X	X	X	X				
Midwife Test 2 (intervention sites only)				X													
Midwife Test 3													X				
Region 2 / Month of study			0	1	2	3	4	5	6	7	8	9	10	11	12		
Mentor training			X														
Midwife baseline test Test 1			X														
Midwife training commenced			X	X													
Partograph collection				X	X	X	X	X	X	X	X	X	X	X	X		
Midwife Test 2 (intervention sites only)						X											
Supplementary mentor training										X							
Midwife Test 3															X		

Table 3.3 Timeline for intervention and data collection

3.7.5 Outcomes

The **primary outcome** was the clinical practice measured by the audit of partograph utilisation as the proxy indicator, analysed at the individual level.

The **secondary outcome** was the midwives' knowledge and skills measured by the midwife tests, and analysed at the individual level.

3.7.6 Instruments and data collection

3.7.6.1 Clinical practice (Partograph utilisation audit)

As indicated in the literature review, it appeared that there was no suitable instrument to evaluate partograph utilisation. The development and testing of the partograph utilisation checklist are fully described in 3.6. This checklist (**Appendix 5**) used for data collection consisted of the following variables (Table 3.4):

Score	No. items	
Partograph	47	Combined as 'observations' (19 items)
Medico-legal (including identification, history and risk assessment)	8	
Maternal	8	
Fetal	5	
Progress of labour	6	
Alignment with clinical management guidelines	15	
Frequency of observations	10	
Risk status	3	
Comprehensive assessment	4	

Table 3.4 : Variables for the partograph utilisation audit

Some items appear in more than one score, but none is duplicated in the partograph score.

3.7.6.1.1 *Validity*

The validation process used was that described by Lynn (1986). Ten judges were used and the Content Validity Index (CVI) was established as above three as required by Lynn (Lynn 1986). See paragraph 3.6.1 for details.

3.7.6.1.2 *Reliability*

There was one evaluator for the partograph utilisation audit. A sub-group of 330 partographs was re-evaluated after an interval of at least two weeks to determine the intra-rater reliability. (The sampling strategy is described in 3.6.2.1.3.) The tests applied were the same as those described in 3.6.2 and obtained the following results :

- ◆ Lin's concordance coefficient (Lin 1989, Lin 2000) of 0.957 ($p < 0.001$), indicating very good concordance.
- ◆ Bland and Altman's limits of agreement procedure (Bland and Altman 1986) indicated a small mean difference (0.13) with fairly narrow limits of agreement at the 95% confidence interval. A graphic representation of these limits showed them to be narrow with no systematic variation over the range of the measurement, thus confirming a high level of concordance.
- ◆ the correlation between difference and mean (Elashoff 2000), was -0.03 ($p = 0.387$). As this is an absolute value near zero (i.e. < 0.1), this implies concordance.
- ◆ the F-test of equality of means and variances (Bradley and Blackwood 1989) gave an F value of 0.95 ($p = 0.387$), where non-significance (as in this result) implies concordance.

These results show a high degree of concordance and agreement for all measures. See paragraph 3.6.3 for details of the tests used. Inter-rater reliability was established and is described in 3.6.2 and 3.6.3.

3.7.6.1.3 *Pilot study*

A pilot study was done on 100 partographs. Issues of clarity, sequencing, formatting, instructions, interpretation and scoring were identified and amended as required. This data was not used in the trial study and was destroyed.

3.7.6.1.4 Data collection

Patient folders were identified according to the sampling procedure described in 3.7.3.1. Once the inclusion and exclusion criteria were applied and the records were deemed eligible (from reading admission records or free text record following admission), the partograph was extracted. If no partograph was found, the free text records were read to determine at what point the partograph should have been commenced and this was recorded against the patient identifier in the research database. This ensured that the non-use of the partograph could be determined.

Photocopies were made of the partographs, as well as the admission record and summary of labour record so that there was evidence of required study data such as time of admission, labour status on admission, type of birth and birth weight of the baby. Where there was no partograph in the patient folder the free text record was photocopied. This ensured that data could be cross-checked (e.g. when diagnosis of labour could have been made) even though it would not be used in the partograph utilisation analysis. All photocopied data had all identifiers removed. A code was assigned immediately and then entered into the research database. This was used later to assign records for analysis. Photocopies were made at the institution so that records remained intact.

3.7.6.2 Midwives' knowledge and skills

Baseline data were collected on all consenting registered midwives. Each test consisted of short questions, a case study and interpretation of a partograph³¹, a cardiotocograph tracing and a clinical assessment of intrapartum assessment. Tests administered in months zero and 12 measured knowledge, analysis and clinical skills covering all aspects of intrapartum care, i.e. risk assessment, maternal condition, progress of labour, fetal condition, management decisions. All tests were matched for content areas and cognitive levels (**Appendix 9**). The nature of the pattern of health service usage and the unpredictability of labour made it unfeasible for clinical tests to be conducted on real patients

³¹ The partographs were simulated case notes and names were included as part of the 'record' as requirement for such a document, as well as to ensure that partographs were allocated correctly to the respective tests.

Therefore standardised testing was done using clinical models. These models were used in the training programme, so were familiar to the midwives.

In addition, testing was done in month three in the intervention group. This test had the same structure and scope as the other tests and was matched for content areas and cognitive levels. This data, along with that collected in month zero and month 12, provided information on short term changes. More frequent testing in the control group would have been logistically difficult and could have had the effect of an intervention by highlighting their knowledge and skills performance.

3.7.6.2.1 *Validity*

The scope of the test instruments were based on standard norms recognised by the International Confederation of Midwives and the South African Nursing Council. The specific content and structure of the tests were developed from the material presented four times at a national workshop in December 2003, and a preliminary study conducted in 2004 (Clow 2005). The case studies used were extracted from undergraduate midwifery examination papers which had previously been moderated by an external examiner (**Appendix 10**). Thus the tests demonstrated face validity.

3.7.6.2.2 *Reliability*

The matching of the tests for content areas and cognitive levels (according to Bloom's taxonomy) (Bloom 1956) enhanced inter-test reliability. The marking of these tests was done by one person (the researcher) therefore there was no need to establish inter-rater reliability. Marking was commenced only after all tests were completed for all timepoints in both regions so that she would not be influenced when administering the tests, especially as she was doing the clinical testing directly with individuals. Non-sequential codes were assigned to the midwives' evaluation forms. The same codes were used for the subsequent evaluations. The clinical evaluations were set up in an OSCE format. Evaluations were coded and collected by the researcher at the time of the evaluation. Once the records were entered into the database, they were filed in numerical order (which was unrelated to the site) so that at the time of marking

the tests, the researcher did not know the source of the data. Only after all the evaluations were marked were the codes linked to the sites for data analysis. The robust process followed in compiling and evaluating the tests contributed to the internal reliability.

3.7.6.2.3 *Pilot study*

A pilot study for each of the three tests was conducted amongst five registered midwives in the metro region who had no contact with the midwives in the trial. These were checked for language, translation, interpretation, ambiguity, timing, and utility of answer sheets. The data from the pilot study was not used in the study and was destroyed.

3.7.6.2.4 *Data collection*

Test one was conducted on the first day of the training intervention in the intervention groups (month zero). For the control groups Test one was administered at the various control sites and the clinical testing was conducted by the researcher at each site during month zero.

Test two which was only for the intervention sites was administered in month three after the study commenced. The researcher visited all sites to conduct the clinical testing.

Test three which took place at the end of the study was conducted at the end of month 12. The researcher visited all sites to conduct the clinical testing.

Due to the extensive area in which the study took place (400km in a northerly direction and 250km in an easterly direction) the researcher made every attempt to test all registered midwives. The theoretical test was sent to the co-ordinator at each site (either the mentor or the nursing manager) to facilitate its completion at a time suitable to the service. The clinical test however, was examined by one person (the researcher) in order to ensure consistency of administration and evaluation. All test answers were recorded on prepared answer sheets. The data collection tools can be found in **Appendix 11**.

Due to the clinical assessment being evaluated by the researcher, it was not possible for the researcher / evaluator to be blinded to the participants during

this test. However, all findings were recorded onto the relevant participant's answer sheet which was coded. Once all the written tests were received the clinical answer sheets were attached to the written test by the research assistant. All tests were collected and stored until all were available. Thereafter the evaluation was done. The nature of the coding was unsystematic so it was not possible to assume the site identity of any of the tests.

3.7.7 Data management

The researcher was the evaluator so she needed to ensure that there was no question about having access to the record or participant codes, or data sheets of the records which had to be evaluated a second time for the intra-reader reliability testing. Thus, the record and coding management was done by the research assistant who worked under the direction of the researcher. Different types of coding systems were used for the trial data and the reliability data for the partograph audit, and for the midwife tests. (See also 3.8.7) The assistant kept the data in a filing cabinet in a secure place to which no one else had access.

At the time of evaluation the partographs were shuffled so that the order in which they were evaluated was not related to sites. Records were divided into batches of 60 records in order to manage the quantity of records. Each batch had at least two control sites and two intervention sites' records included. This was done by the research assistant in order to ensure that the researcher remained blinded to the source of the data in terms of study arm or site/cluster.

The pilot study evaluations were all recorded on yellow paper to ensure that they did not get mixed with the records for the rest of the study.

At the end of evaluation of the partographs and midwife tests, the data was double-entered at the Medical Research Council in Cape Town. Datasets were compared. Incompatibilities which were identified were checked against the source data and corrected. A back-up copy of the entered raw data was kept by the researcher.

3.7.8 Data analysis

Despite this study having fewer than 15 clusters in each treatment arm, the disadvantages for the cluster level summaries made a cluster level approach to analysis less attractive. This was particularly in relation to the cluster sizes for the secondary endpoint of midwife knowledge and skills where all midwives were followed up during the study period, and personnel movements were dynamic. Thus, the individual-level approach offered a more efficient approach.

3.7.8.1 Clinical practice - primary outcome

The unit of analysis was the individual record, i.e. the partograph record. All analyses of partograph records were done on the basis of intention to treat (ITT) (described in 3.3.5.3). The partograph scores post-intervention were compared between the two arms of the study using a linear model with adjustment for clustering within centres (GEE). The model used had an exchangeable correlation matrix which is usually the most appropriate correlation structure for individuals living in the same community or patients in the same medical practice (Hayes and Moulton 2009). Sub-scales of the labour record were coded by defined categories and analysed in the same way.

The partograph scores were analysed as follows:

- ◆ descriptive statistics showing frequencies by arm
- ◆ graphics – boxplots and histograms by arm
- ◆ inference for the trial – comparison of arms was done by design and randomisation of clusters was taken into account. Generalised linear models for continuous ordinal and discrete outcomes were used. These are equivalent to the t-test, the chi-square test and the test for trend, but adjusted for clustering.
- ◆ use of quantile regression for exploratory analysis to find the cutpoint for the partograph score
- ◆ comparison of scores was done only for those births where a partograph was used.
- ◆ sub-scales of the labour record were coded by defined categories and analysed in the same way.

- ◆ *Post hoc* secondary analysis – a linear mixed effects model of the partograph score with main effects of group (intervention, control), month (of partograph collection), region (Boland / Overberg, West Coast / Winelands) and size of service (small, moderate–large in terms of number of births *per annum*) with site as the random effect. The three main effects were introduced as possible confounders in the model and the interactions were investigated. Only the significant interaction terms (*viz.* group and size of service) were retained in the model.
- ◆ The profile of the partograph score over the year of prospective data collection was investigated within each group by means of a Lowess non-parametric regression model. This analysis and graphical representation was exploratory.
- ◆ The intracluster correlation coefficient (ICC) is reported for the primary outcome and subsector results. There were 17 clusters throughout and the size of each cluster was 60. The ICC was calculated by ANOVA, using Stata version 11 (Statacorp 2009), with data pooled over control and intervention arms, post intervention.

3.7.8.2 Midwife tests – secondary outcome

Analysis was done at the individual level (midwife).

The midwife test score post intervention was compared between the two arms of the study using a linear regression model with adjustments for the clustering of midwives within centres (General Estimation Equations – GEE) using an exchangeable correlation matrix. The model also included an adjustment for the stratification by regions and size of the sites. Specific sectors of the test (e.g. question type and cognitive levels) and practice aspects (e.g. risk assessment, diagnosis of labour, fetal assessment) were analysed further. Because GEE is generally accepted for studies with 15 or more clusters per arm, the analysis on the total score was done using both GEE and mixed effects models. The findings were shown to be consistent and so the remaining subsector analyses were done using GEE.

All sites remained in the study. However, staff movement as well as some logistical challenges resulted in some data not being complete. Hence the decision was made to analyse the data in two ways (described in 3.3.5.3).

- ◆ For those who were not present throughout, the data was treated on the basis of 'intention to treat'. This was regarded as the primary analysis.
- ◆ For those who were present throughout, the data was treated 'per protocol'. This represented the core group. The scores for this analysis were adjusted for baseline, where the same measure at baseline is taken as a covariate. This strategy is used to improve the precision in this smaller subgroup.

Tests were divided into theoretical and clinical portions, and analysed accordingly. There were some non-completions due to operational and logistical reasons. These are presented in the results chapter (4.4.1) and the implications discussed in the limitations section (5.4.2.4.2).

The midwives' test total score post intervention was compared between the two arms of the study and analysed as follows :

- ◆ descriptive statistics – mean, standard deviation, minimum, and maximum, and frequencies per arm
- ◆ graphics – box plots per arm and scatterplots
- ◆ inference for the trial – comparison of arms was done by design and randomisation of clusters was taken into account. Generalised linear models (adjusted for clustering) for continuous ordinal and discrete outcomes were used. Due to the relatively small number of clusters, the GEE may be limited in its applicability so the mixed effects model was also used and these are compared for the main findings. The rationale for using GEE is presented in 3.3.5.2.
- ◆ full analysis was done on those who were present throughout the study and had completed tests at all timepoints ('per protocol analysis' - PPA). Adjustment was made for nominated baseline variables by using analysis of covariance.

- ◆ tests were divided into theoretical and clinical portions, and analysed accordingly ('Intention to treat' (ITT) or PPA). Where theory and clinical components were undertaken and complete, this analysis was also presented.
- ◆ specific sectors of the tests (cognitive levels, types of questions and practice aspects) were coded by defined categories and analysed in the same way.
- ◆ the intraclass correlation coefficient (ICC) is reported for the secondary outcome and subsector results. There were 17 clusters with a mean size of cluster of 9.1 and a range of 4-14. The ICC was calculated by ANOVA, using Stata version 11.0 (Statacorp 2009), with data pooled over control and intervention arms, post intervention.
- ◆ The changes within the intervention arm at three timepoints over the 12 month period were described by converting the total and sector scores to percentages and plotting these for each timepoint.

3.7.9 Managing potential bias

Careful attention was paid to minimise the occurrence of the four primary sources of bias that can affect the validity of the design of a cluster randomised trial (Murray 2001).

Selection bias, bias due to differential history and bias due to differential maturation were minimised by use of a randomised trial design where there was random assignment of sufficient numbers of groups per condition / arm. A careful sample construction and size determination was done taking into account the capacity of the field sites. This is described in 3.7.3.1 for the partograph audit, and 3.7.3.2 for the midwives' tests. Further, as this study had fewer than 20 clusters per arm, stratification was done by geo-political region, and for the sites based on the number of deliveries per year. The stratification was done prior to randomisation and is described in 3.7.3.3.

Contamination cannot be addressed through randomisation. The pragmatic nature of this trial recognised the real-world setting of this study. To this end, all

region-wide activities continued as usual, e.g. distribution of policies and clinical guidelines, staff development programmes, perinatal review meetings with the regional obstetrician (see 3.7.3.4). However, an embargo was placed on intrapartum in-service training during the data collection period. What was different in the intervention arm was the added activity, i.e. the clinical facilitation programme and focussed in-service training.

Given the nature of the intervention, participants knew to which arm of the study they were allocated. This was unavoidable. By using whole sites as clusters (and where the unit of randomisation was the clinical site), and the fact that these sites were the only health services offering intrapartum care in a large district, the likelihood of a person moving from one site to another in another district was highly unlikely. In a sense this was similar to the 'fried egg' strategy (Lewycka, Mwansambo *et al.* 2010) where the intervention and data collection occurred in sites that were located reasonably centrally within a district, and these were not contiguous with other sites. This did not have the potential disadvantage for representivity across the cluster, as the primary and secondary endpoints related to in-hospital dynamics and not community-wide dynamics.

The evaluation personnel were limited to the single researcher. However, there was a rigorous coding and recoding process to reduce the possibility of identification of the source data, and this was administered by a research assistant under direction of the researcher. Records were not evaluated on a per-site basis, thus further reducing the possibility of identifying relationships. This is described in more detail in 3.7.6.1.4 (data collection partograph utilisation audit), 3.7.6.2.4 (data collection midwife tests) and 3.7.7 (data management).

Given the pragmatic nature of the study there were different analysis groups (i.e. ITT or PPA) and theory and clinical portions of the tests were analysed separately due to logistical challenges (presented in 5.4.2.4.2.1). In order to

determine if this had resulted in any bias, comparisons were made to determine the sensitivity of the results across these various groupings.

3.7.10 Reporting the study

As indicated in 3.3.6 there are standard reporting requirements for various trial designs. These requirements are incorporated in the text where applicable, and a table identifying the location of each is presented in **Appendix 12**.

3.8 Ethical considerations

Fundamental principles underpinning ethical conduct are respect for persons, beneficence (and non-maleficence), and justice (Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO) 2002)³² Translating these to the conduct of research requires a study to be relevant with an appropriate study design, investigator competence, ethical review, appropriate information for potential participants in order for them to make an informed decision regarding participation, patient / participant privacy, and a balance of risks and benefits for individuals and communities involved in the research (Department of Health 2006b, Weijer, Grimshaw *et al.* 2012, World Medical Association 2013).

Typically in randomised controlled trials the research subject simultaneously is the unit of randomisation, the unit of the intervention and the unit of observation (Weijer, Grimshaw *et al.* 2011). However, the unique design of CRTs means that in a single CRT the units of allocation, intervention and outcome measurement may differ, thus posing distinct ethical challenges, particularly for informed consent procedures. Randomisation of clusters usually takes place before it is possible to identify and recruit individuals. Where there are cluster-level interventions it may be difficult for individuals to avoid participating, thus limiting their right to refuse consent. In certain social or organisational groupings it might be unclear who has the status or authority to speak for the group (Weijer, Grimshaw *et al.* 2012). In this study the unit of allocation was the hospital or midwife obstetric unit, the intervention was delivered to professional staff, and data were collected from patient records and midwives.

³² This is abbreviated to “CIOMS in collaboration with the WHO” in subsequent references

Concern has been expressed at the low level of reporting of research ethics practices for cluster randomised trials (Taljaard, McRae *et al.* 2011). As this study was an intervention study, the principles contained in the World Medical Association's Declaration of Helsinki and the Council for International Organizations of Medical Sciences (CIOMS) which set out the ethical principles for medical research involving human subjects, and the guidelines for scientific and ethical standards required to conduct health research in South Africa were applied³³ (CIOMS in collaboration with the WHO 2002, Department of Health 2004, Department of Health 2006b, World Medical Association 2013). The International Council of Nurses' (ICN) Code of Practice for Nurses is complementary to these documents (International Council of Nurses 2012). Specific considerations relating to CRTs were guided by the Ottawa Statement on the ethical design and conduct of cluster randomised trials (Weijer, Grimshaw *et al.* 2012). The ethical requirements for this study were met as follows :

3.8.1 Scientific soundness of a study

Scientific soundness is a pre-requisite for it to be ethical, and the study should address relevant health and development needs (Department of Health 2004, Department of Health 2006b, World Medical Association 2013). This study addressed one of the four focus areas for South African health, viz. maternal and child health. A study design appropriate to accommodate health service dynamics was used, and the protocol was reviewed through an academic departmental scientific review process.

3.8.2 Investigator preparedness

The investigator had the necessary research training, clinical expertise and education experience to conduct this study. A module on Experimental Epidemiology, Clinical and Field Trials and a Good Clinical Practice course were successfully completed. Appropriate consultation with a statistician during design and analysis was done (Department of Health 2004, Department of Health 2006b, World Medical Association 2013).

³³ Many of the CIOMS in collaboration with WHO (2002) guidelines are incorporated in the Department of Health (2004) guidelines. Only where there are additional criteria relevant to this study or greater clarification available, is the CIOMS reference cited.

3.8.3 Ethics approval

Ethics approval was obtained from the University of Cape Town Faculty of Health Sciences Human Research Ethics committee (HREC) (Study Reference: REC 311/2004 and REC 244/2007) (**Appendix 13**) (Department of Health 2004, Weijer, Grimshaw *et al.* 2012) A copy of the full protocol is available from the HREC office - <http://www.health.uct.ac.za/research/humanethics>.

3.8.4 Permission to conduct the study

Permission to conduct the study in public sector health facilities was obtained from the Provincial Government of the Western Cape Health Department, and the relevant regional directors. This included permission to access relevant clinical records for the purposes specified, as well as staff (**Appendix 14**). (Department of Health 2004) As some of the records were lodged at referral sites which were outside the study sites or regions, further access was obtained to source these records (Study Reference : 19/18/RP244/2007). Consent for access was sought before randomisation. Although such permission has a gatekeeping function, this permission did not assume proxy consent on behalf of individuals in the clusters (Gallo, Weijer *et al.* 2012). See 3.8.5.

3.8.5 Informed consent

Managers, clinical facilitators, and registered midwives (intervention and control sites) were deemed to be human research subjects for a CRT as they met at least one of the conditions described by Weijer, Grimshaw *et al.* (2012). All received an information sheet (available in English, Afrikaans and Xhosa), indicating the scope of the project, their role and expectations of them, their rights as research participants and their protection during this study (**Appendix 6** Information to mentors and **Appendix 15** Information to registered midwives). Given that potential participants were approached after cluster randomisation, the information provided a detailed description of the intervention in the arm to which they were assigned (McRae, Weijer *et al.* 2011). All potential participants were given an opportunity to ask questions and obtain clarification about the study, whereafter they signed their consent to participation witnessed by a third party (CIOMS in collaboration with the WHO 2002, Department of Health 2004, World Medical Association 2013).

The patient records did not meet the criteria for human research subjects for CRTs as there were no patient level interventions, the researcher had no interaction with individual patients and there was no use of identifiable private information (Weijer, Grimshaw *et al.* 2012). Thus there was no need to obtain informed consent from patients.

3.8.6 Distributive justice

Distributive justice required that the midwives of each of the three rural regions of the province should have had an equal opportunity to be exposed to this project, as should the mothers and babies (CIOMS in collaboration with the WHO 2002, Department of Health 2004, World Medical Association 2013). The project was conducted in only two of the three rural regions of the Western Cape Province due to resource constraints of a study undertaken by a single researcher across a large geographic area. All regions had received guidelines regarding management of labour, and had received training in the use of the partograph as part of the provincial roll-out in 2005. An undertaking was made that should this study's intervention be shown to be beneficial, it would be made available to the control sites of the participating regions as well as the remaining rural region and the metropolitan region, but would not form part of the study. Thus, no region was disadvantaged. This decision was supported by the provincial maternity guidelines group.

Although there are three main languages used in the Western Cape province, training was offered in English and Afrikaans only. The documentation of all hospitals and clinics in the province is done in either of these two languages and proficiency is expected. However, due to the interactive nature of the training programme, there was opportunity for midwives to communicate in Xhosa and get the necessary feedback or clarification, when this was necessary. All training materials were available in English and Afrikaans.

Information and consent documents were available in Xhosa (see above). No participants were prevented from participating due to language constraints.

3.8.7 Privacy, confidentiality and data protection

The right to privacy and confidentiality must be protected. Any personal information about participants or patient records must be collected, stored and destroyed in a way that will protect these rights (CIOMS in collaboration with the WHO 2002, Department of Health 2004).

Partographs or relevant clinical data were photocopied. Identifiers were removed from the partographs and a study code assigned. When the study is complete this coding system will be destroyed. Clinical records did not leave the site while data collection was undertaken. Photocopied data was kept in a secure cabinet in the researcher's office or with the research assistant during the period that the researcher needed to be blinded.

Registered midwife participants were assigned a randomly generated study code which was used on the test answer sheets. This was only available to the researcher and was used only to assign test scores to the appropriate site and arm for analysis. When the study is complete this coding system will be destroyed.

3.8.8 Risk benefit analysis

Research with human subjects should only be conducted if the importance of the study outweighs the risk and burden (World Medical Association 2013) .

3.8.8.1 Benefits

The analysis of benefits of this study suggested that individuals, health service sites and ultimately the standard of intrapartum care of women were likely to benefit from this study (CIOMS in collaboration with the WHO 2002, Department of Health 2004, World Medical Association 2013).

The midwives benefitted from the learning opportunity which addressed an identified need in the province. This learning was applied directly in their working situation.

It was anticipated that this training programme would result in better clinical decision making during the management of labour thus accruing benefit to the pregnant population, fetal and newborn wellbeing, and to the improved performance of the clinical services. If inappropriate clinical management practices were identified, these could be addressed directly in the training session, or else appropriate channels for addressing these were identified or made available.

An undertaking was made that if the intervention was found to be beneficial, training and training material would be given to the control sites and regions which were not included in the trial once the project was complete (see 3.8.6.1) to facilitate the running of similar workshops on an ongoing basis.

Mentorship³⁴ and audit and feedback forms part of the responsibility for quality assurance which is vested in unit managers. The programme of clinical facilitation equipped them to meet this responsibility. Regional managers agreed that anyone undertaking the clinical facilitation development programme would have it recognised as part of the performance evaluation system thus adding personal professional value.

3.8.8.2 Potential risk

The analysis of potential risks associated with this study anticipated that there would be negligible risk to individuals or inconvenience to the health service sites (CIOMS in collaboration with the WHO 2002, Gallo, Weijer *et al.* 2012).

³⁴ Mentorship is the term reflected in the role descriptions. This includes clinical facilitation. For the purposes of this study the terms clinical facilitator / facilitation have been chosen.

A potential risk was that of exposing inadequate or poor knowledge and skills among registered midwives. All tests were identified by a study code (3.8.7) and an undertaking was made that individual scores would not be made available to any supervisor / line manager, and could not be used for any operational processes. Thus any risk of exposure of a midwife's poor performance was eliminated. In a situation where incompetence might have been identified which could cause potential harm in clinical practice, the researcher would have been obliged to confidentially identify the specific midwife and arrange a confidential counselling session with him/her in an effort to commence a programme of capacity building. Where a general level of incompetence might have been identified in a particular site, this would have been addressed through the training programme which formed part of this study.

It was possible that research activities could have negatively affected service operations and inconvenienced individuals. The clinical facilitation training for managers and clinical supervisors, and the intrapartum training for registered midwives, were regarded by the regions as appropriate in-service training and supportive of the operational requirements. The time allocated to this training was not additional to the time that would be set aside for relevant in-service training, so there was no detrimental effect on the services. This was regarded as on-duty time and transport and meals / refreshments were provided according to the policy of the Provincial Government of the Western Cape Department of Health Human Resource Development department, as for any other training programme. There was therefore no inconvenience or additional cost to the individuals who participated.

3.8.8.3 Risk benefit analysis conclusion

The risk benefit analysis indicated that the benefits would outweigh any potential risk, and there was no reason to prevent this study from proceeding.

3.8.9 Dissemination of results

Participants and communities are entitled to be informed of the findings of research in which they have participated (Department of Health 2004, World Medical Association 2013). Final reports will be submitted to the Human

Research Ethics Committee, the funding agencies, to the provincial health department, and a summary report to all sites involved in the study will be provided. The study will be presented at relevant national and international conferences, and articles will be submitted to appropriate peer-reviewed journals.

3.9 Conclusion

This study was a pragmatic cluster randomised trial, with due attention to methodological rigour. There was a package of multifaceted interventions which was offered close to the work location of the participants. The training was facilitated by an experienced midwife from a credible institution, who could be regarded as being 'educationally influential'. The sites used were well prepared so that there was a state of readiness for change

Outcomes were measured using a proxy indicator for standard of care and testing of midwives using requirements recognised by the International Confederation of Midwives and the South African Nursing Council. Analysis was done at the individual level taking the clustering into account, and the necessary sample size and power considerations were taken within the constraints of the available sites.

Therefore this project which was designed as a cluster randomised trial aimed to bring a new dimension to the understanding of clinical improvement in intrapartum care in a rural setting.

Chapter 4 Results

4.1 Introduction

This chapter presents profiles of the sites that were used in this study (4.2.1), the profile of mothers and neonates whose intrapartum records were used for the partograph utilisation audit (4.2.2), as well as the profile of the registered midwives in terms of their professional education and experience (4.2.3), illustrating the effectiveness of the randomisation process. Thereafter the results of the primary (4.3) and secondary outcomes (4.4) are presented along with the subsidiary analyses, followed by the findings in respect of the intervention arm over time (4.4.3).

4.2 Study Profiles

4.2.1 Profile of Sites

All eligible units in the two regions were approached during the first half of 2006 and all agreed to participate in the study. In total there were 17 sites, representing all primary level services in the two regions which offer intrapartum care. Each site formed a cluster. These were stratified on the basis of geo-political region (Boland / Overberg (region 1) and West Coast / Winelands (region 2)) and the number of births that occurred in the site, and were randomly allocated to intervention (n=9) or control (n=8) groups. This is reflected in Table 4.2.1. The stratification process is fully described in 3.7.3.3. All 17 sites remained in the study and there were no sampling violations.

Intervention sites / clusters				Control sites / clusters			
Site ID	Type of facility	No. births p.a.	Region	Site ID	Type of facility	No. births p.a.	Region
1	Hospital	<500	1	3	District Hospital	>1000	1
2	District Hospital	500-1000	1	5	MOU	<500	1
4	MOU	>1000	1	7	Hospital	<500	1
6	District Hospital	500-1000	1	8	Hospital	500-1000	1
9	District Hospital	<500	1	10	Hospital	<500	2
13	Hospital	<500	2	11	Hospital	<500	2
14	Hospital	<500	2	12	District Hospital	>1000	2
15	District Hospital	>1000	2	17	District Hospital	500-1000	2
16	District Hospital	>1000	2				

Table 4.2.1 Description of the sites (clusters) and their study allocation

The distribution of the registered midwives in relation to the sites is shown in Table 4.2.2.

Characteristic	Intervention arm (n = 79)	Control arm (n = 75)
Type of service		
• District Hospital	49	27
• Hospital	22	40
• Midwife Obstetric Unit	8	8
No. births per year³⁵		
• <500	34	34
• 500 – 1000	20}	24}
• >1000	25} 45	17} 41
Region		
• 1	54	43
• 2	25	32

Table 4.2.2 Randomisation of the registered midwives in relation to the sites

Conclusion : The stratification yielded reasonably comparable arms.

³⁵ For purposes of analysis the 'moderate' (500-1000) and 'large' (>1000) numbers of births per annum groups were treated as one group

4.2.2 Profile of partograph records

The profile of the pregnant women included the age, gravidity, parity, cervical dilatation and type of birth (method of delivery). The profile of the babies included gestational age and birth weight (Table 4.2.3). The birthweights in both arms are regarded as being appropriate for term neonates in this population.

Characteristic	Intervention 540 records in 9 clusters	Control 480 records in 8 clusters
Mean age of mother ($\pm$SD), y	24.6 (6.2)	24.3 (6.6)
Gravidity n/N, (%) • 1 • 2-4 • >4	246 / 540 (45.6) 273 / 540 (50.6) 21 / 540 (3.9)	241 / 480 (50.2) 212 / 480 (44.2) 27 / 480 (5.6)
Parity n/N, (%) • 0 • 1-3 • >3	266 / 540 (49.3) 262 / 540 (48.5) 12 / 540 (2.2)	258 / 480 (53.8) 203 / 480 (42.3) 19 / 480 (3.9)
Cervical dilatation ($\pm$SD), cm • on admission • on commencement of partograph	2.7 (1.4) 4.2 (1.6)	2.5 (1.4) 4.2 (1.7)
Type of birth n/N, (%) • Normal vaginal delivery • Forceps delivery • Caesarean section • Vacuum delivery • Other • Not recorded	462 / 540 (85.6) 1 / 540 (0.2) 56 / 540 (10.4) 15 / 540 (2.8) 5 / 540 (0.9) 0 / 540 (0)	420 / 480 (87.5) 3 / 480 (0.6) 40 / 480 (8.3) 12 / 480 (2.5) 3 / 480 (0.6) 2 / 480 (0.4)
Gestational age ($\pm$SD), weeks	38.8 (1.9)	38.7 (1.9)
Birth weight ($\pm$SD), g	3100 (507)	2985.9 (500)

Table 4.2.3 Profile of the records of women and their babies included in the study

Conclusion : Overall in terms of the labour records describing women and their babies, the study arms appear to be matched and thus the randomisation was regarded as effective.

4.2.3 Profile of registered midwives

A total of 154 registered nurse-midwives enrolled in the study.

4.2.3.1 Professional education and experience

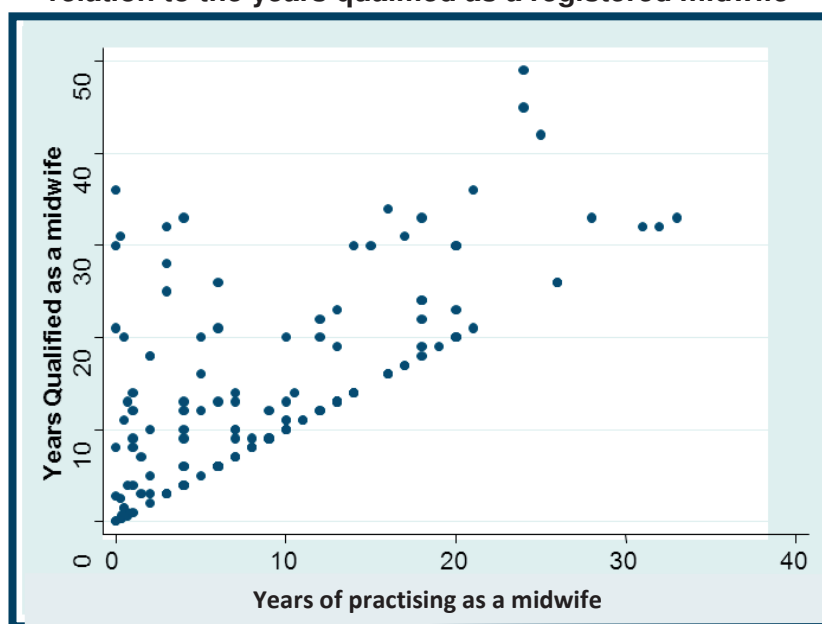
The midwives are described in terms of their midwifery education (either integrated into the basic four year training as a nurse and midwife, or as a one year diploma in Midwifery after becoming registered as a general nurse), and the number of years they had practised as a midwife. This profile is reflected in Table 4.2.4. There was a small difference between arms in respect of the type of educational programme undertaken. The mean number of years qualified and practising as a midwife were evenly matched between the two arms.

Characteristic	Intervention arm 79 midwives in 9 clusters	Control arm 75 midwives in 8 clusters
Midwifery education, n (%) • Integrated 4 year programme • 1 year post registration diploma	42 (47.2) 37 (58.7)	47 (52.8) 26 (41.3)
Total number of years qualified, y • Mean $\pm$SD • Median • Range	15.4 $\pm$ 11.1 13 0-45	13.7 $\pm$ 9.3 12 1-49
Total number of years practising as a midwife, y • Mean $\pm$SD • Median • Range	9.0 $\pm$ 7.4 7 0-31	9.7 $\pm$ 7.7 7 0-33

Table 4.2.4 Profile of the registered midwives in each arm in terms of educational preparation and experience

The scatter plot (Figure 4.2.1) illustrates that many registered midwives had been qualified for many more years as a midwife than the number of years' experience they claimed. Many of them had one to two years' experience, not all of which would have been consecutive (to consolidate knowledge and skills, and to build confidence) or current.

Figure 4.2.1 Scatter plot illustrating the years of experience (practice) in relation to the years qualified as a registered midwife



4.2.3.2 Knowledge and clinical skills

The baseline scores for the control and intervention arms are presented to demonstrate the effectiveness of the randomisation process. Due to the different composition of midwives at both timepoints (presented in 3.7.8.2 and 4.4.1), baseline data in respect of the test scores are presented

- ◆ for all who enrolled at baseline (but who might not have completed both theory and clinical components)³⁶, and
- ◆ for two subgroups
 - ◇ those who had complete tests at baseline, and
 - ◇ those who had complete tests at baseline and also completed test three.

One hundred and twenty-three midwives from 17 clusters enrolled at baseline. The pattern of theory and clinical component completions for test one for each arm is reflected in Tables 4.2.5a and b.

³⁶ The theory and clinical scores are exclusive sections making up the total score. These scores are presented in order to avoid the problem of either theory or clinical portion being incomplete and thus reducing the number of test scores that could be reported on.

		Complete	Incomplete ³⁷	Not undertaken ³⁸	Total
Intervention arm	n	58	0	21	79
	%	73.4	0.0	26.6	51.3
Control arm	n	60	5	10	75
	%	80.0	6.7	13.3	48.7
Total	n	118	5	31	154
	%	76.6	3.3	20.1	100

Table 4.2.5a Pattern of completions for Test one - theory component

		Complete	Incomplete ³	Not undertaken ⁴	Total
Intervention arm	n	58	0	21	79
	%	73.4	0.0	26.6	51.3
Control arm	n	61	3	11	75
	%	81.3	4.0	14.7	48.7
Total	n	119	3	32	154
	%	77.3	1.9	20.8	100

Table 4.2.5b Pattern of completions for Test one - clinical component

This data is reflected to give a comprehensive picture of each arm. These tables illustrate that similar proportions were completed within arms for theory and clinical components.

As all midwives did not complete both portions of the tests, the theory and clinical scores are presented for 119 and 118 midwives respectively (Table 4.2.6). This shows that for those enrolled at baseline the scores were similar in both arms.

³⁷ It is appropriate to have no 'incompletes' for the intervention arm for test one as all were tested on day one prior to the commencement of the intervention.

³⁸ The 'not undertaken' column refers to those who enrolled in the study any time in the 12 months that followed the baseline tests, i.e. they were not part of the study at this timepoint.

Test component (max. possible score)	Obs ³⁹	N	Arm	mean	SD	min	max
Theory (93)	58	58	Intervention	36.0	7.0	19	56
	61	61	Control	35.9	9.5	13	54
Clinical (36)	58	57	Intervention	13.9	4.8	8	35
	61	61	Control	13.3	3.7	6	22

Table 4.2.6 Descriptive statistics for baseline theory and clinical scores for all midwives enrolled at baseline

Given that there were considerable changes in staff during the year in which the data was collected, analysis was done on the basis of ‘intention to treat’ as well as ‘per protocol’. In order to facilitate comparison across the total enrolment at baseline and these two subgroups of midwives, the distinction between theory and clinical scores is maintained for this presentation. The baseline scores in respect of these two subgroups of midwives are presented in Tables 4.2.7 and 4.2.8.

There were 96 participants who had complete tests at baseline, i.e. both theory and clinical components were complete (Table 4.2.7).

Test component (max. possible score)	Obs	N	Arm	mean	SD	min	max
Theory (93)	48	34	Intervention	35.5	6.2	21	47
	48	39	Control	35.1	9.4	13	54
Clinical (36)	48	33	Intervention	12.7	3.4	8	19
	48	39	Control	12.7	3.6	6	20

Table 4.2.7 Descriptive statistics for baseline theory and clinical scores where the tests were complete (‘intention to treat’)⁴⁰

There were complete data at both timepoints (baseline and month 12) for a subgroup of 70 midwives (‘per protocol’) (Table 4.2.8).

³⁹ Obs refers to the number of people enrolled at that time point.

⁴⁰ While this is not a statistical analysis, this has been labelled as the ‘intention to treat’ group to assist the reader to identify these distinctions later on.

Test component (max. possible score)	Obs	N	Arm	mean	SD	min	max
Theory (93)	34	34	Intervention	35.8	6.2	21	47
	36	36	Control	36.2	9.4	13	54
Clinical (36)	34	33	Intervention	12.7	3.4	8	19
	36	36	Control	12.4	3.6	6	19

Table 4.2.8 Descriptive statistics for baseline theory and clinical scores ('per protocol group')⁴¹

These two tables (4.2.7 and 4.2.8) show that the two arms were similar in the complete group at baseline and the subgroup at baseline that completed the tests at both timepoints.

Conclusion : From the profiles presented in respect of the registered midwives, viz. education, years of experience, and tested knowledge and skills, one can conclude that the randomisation in relation to the registered midwives was successful as the two arms were well matched.

4.3 Clinical practice (Partograph utilisation audit) (Objective 1)

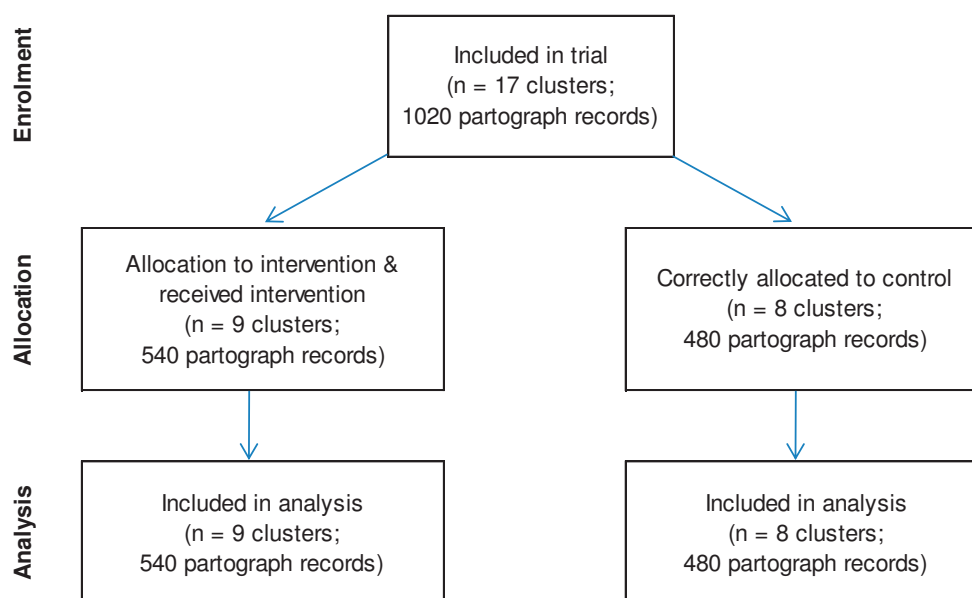
Comparison of the two arms in respect of the partograph score is presented. A detailed analysis is also presented to show where the differences are most marked, as well as three potential sources for an interaction effect. The unit of analysis was the individual partograph record. In order to identify where the strengths and weaknesses lay in completing the partograph, analysis of various sectors of the partograph was performed and is presented. Further, the results specific to innovations introduced during this study are presented. In order to understand the reasons for non-completion of partographs, contextual information was captured at the time of the data collection. This information is presented after the results.

⁴¹ While this is not a statistical analysis, this has been labelled as the 'per protocol' group to assist the reader to identify these distinctions later on.

4.3.1 Distribution of partograph records

Figure 4.3.1 illustrates the participating sites (clusters), their allocation and the numbers of records in each arm. There were no protocol deviations.

Figure 4.3.1 Flow chart indicating the sites / clusters and partograph records enrolled in the study ⁴²



Of the 1020 records that were selected, no partograph was used for 306 (30%). Comparing the arms in terms of the utilisation of partographs (adjusted for 17 clusters) revealed that, although there was an absolute difference of 4% in favour of the intervention group, this was not statistically significant ($p=0.734$) (Table 4.3.1). Given that the difference in proportions of partographs used was not statistically significant, further analysis of the subgroup where partographs were used was permissible.

⁴² 'Correctly allocated to control' means that the allocation to the control arm was performed as specified in the protocol, i.e. there were no protocol violations

Variable	Intervention	Control	Coefficient (intervention effect)	95% Confidence Interval		ρ
				Low	High	
Utilisation of partographs	0.72	0.68	0.04	-0.19	0.27	0.734

Table 4.3.1 Comparison of the utilisation of partographs between the two arms

4.3.2 Primary outcome : Clinical practice (partograph utilisation)

The partograph score as a proxy measure of clinical practice was the primary outcome. A map to navigate this section of results is provided in Figure 4.3.2.

Scores for the completion of the partograph were based on the records where there was evidence of the utilisation of the partograph (n=714). The maximum possible score was 47 (Table 4.3.2).

Arm	N	Mean score	SD	min	p25 ⁴³	p50	p75	max
Intervention	388	22.3	6.7	8	17	22	27	39
Control	326	20.8	5.2	3	17	21	25	38

Table 4.3.2 Descriptive statistics for the partograph score

The descriptive statistics in Table 4.3.2 are reflected graphically in Figure 4.3.3. Although the median for the partograph score was calculated at 22.3 and 20.8 in the intervention and control arms respectively, there was an indication of difference between the distribution of scores for the two arms at the higher percentiles illustrated in Figure 4.3.3. In effect, the shift in intervention effect only commenced above the 50th percentile.

⁴³ p25 refers to 25th percentile; p50 refers to the 50th percentile; p75 refers to the 75th percentile. This also applies to similar tables of descriptive statistics

Figure 4.3.2 : Map of presentation of results and findings - partograph utilisation audit

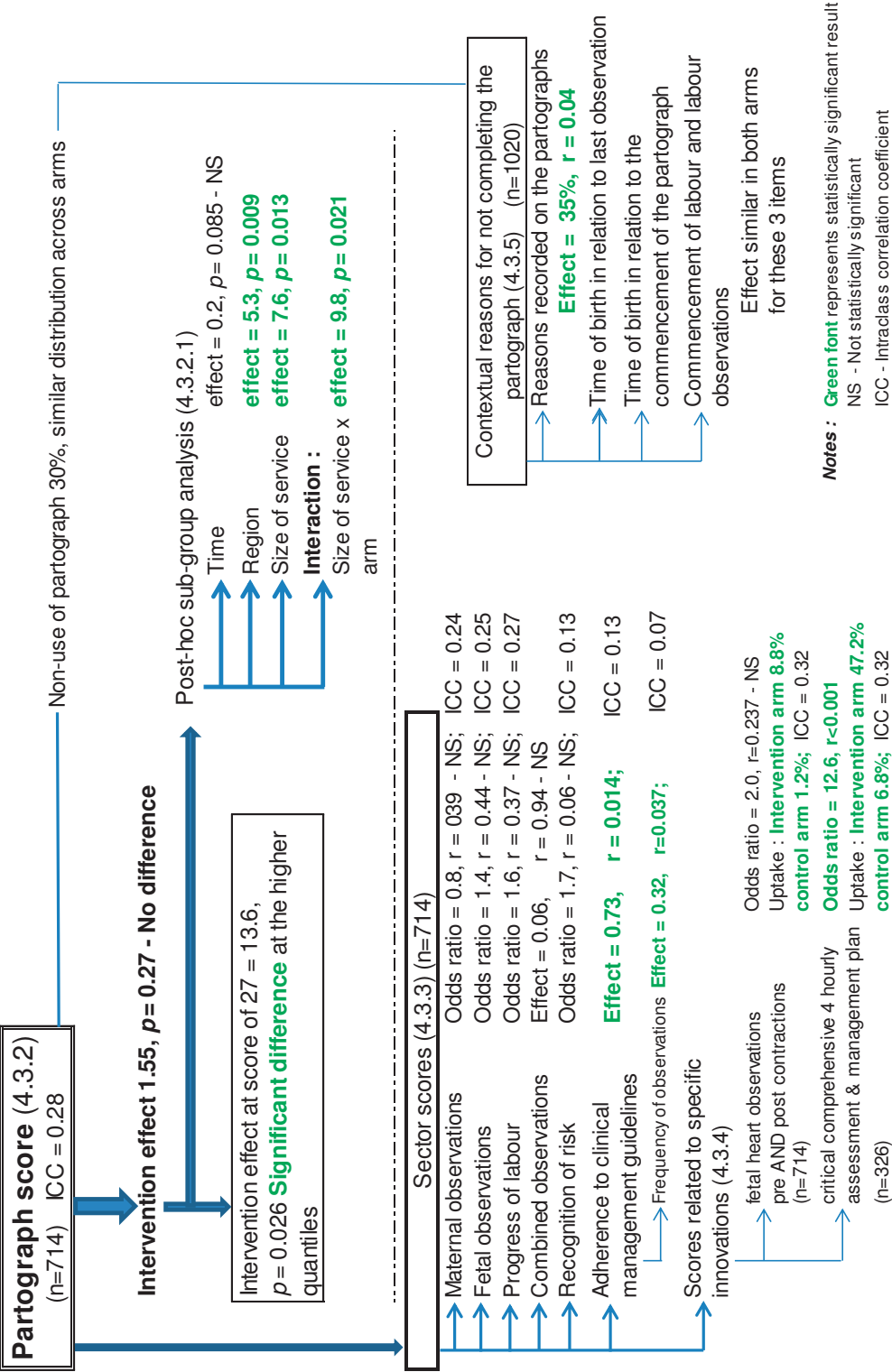
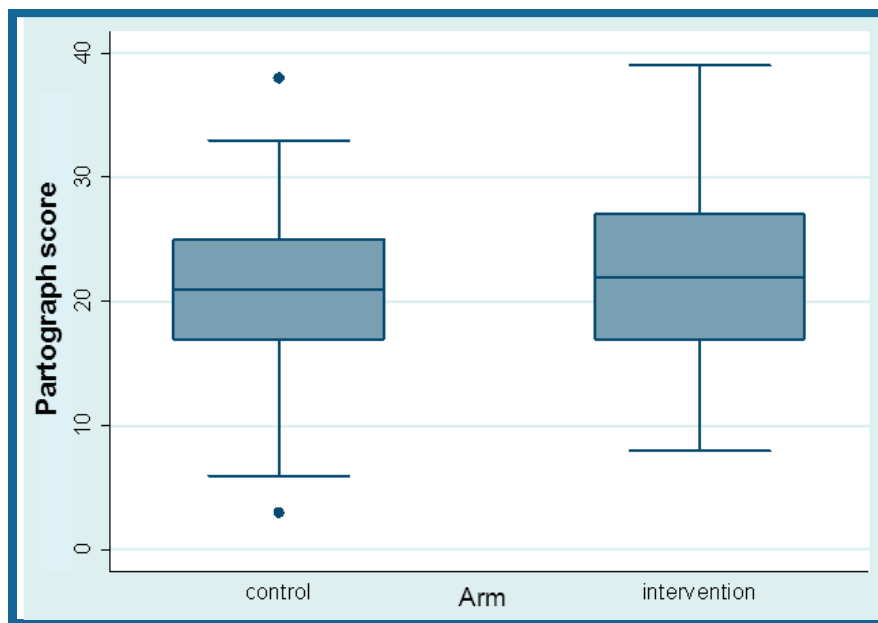
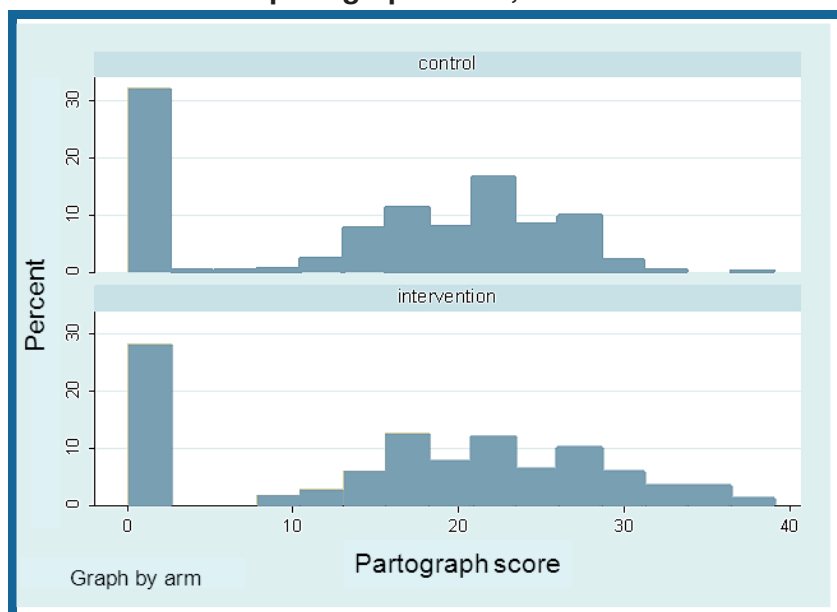


Figure 4.3.3 Boxplot of total partograph scores



A closer look at the profiles of the scores illustrates where the differences were to be found. In addition to indicating the records where a partograph was not used, Figure 4.3.4 illustrates a different pattern between the two arms in that there was a higher minimum score and a heavier tail in the upper score levels in the intervention arm. This pattern was confirmed by the distribution in the box plot above (Figure 4.3.3).

Figure 4.3.4 Distribution of the partograph score, reflected for both arms



The comparison of the mean scores between the two arms shows that the intervention effect was 1.55 marks (95% CI: -1.18 to 4.28) and the p -value was 0.267, thus this difference was not statistically significant (Table 4.3.3). The difference relative to the total score using the overall scale is 3.3%. The intraclass correlation coefficient (ICC)⁴⁴ for this score was 0.28 which is regarded as high, as it means that 28% of the variability in the partograph score could be explained by variability between sites.

Variable (max possible score)	Intervention	Control	Coefficient (intervention effect)	95% Confidence Interval		p	ICC
				Low	High		
Partograph score (47)	22.4	20.8	1.6	-1.2	4.3	0.267	0.28

Table 4.3.3 Comparison of mean partograph scores between the two arms

In order to determine at what level the intervention effect was observed (as suggested by the box plot and the histogram (Figures 4.3.3 and 4.3.4 respectively)), simultaneous quantile regression was performed at the deciles 10 – 90. Bootstrapping is a type of regression analysis which uses simulation to estimate the variability in the true population, where more information is required about the properties of estimators for unknown populations (SPSS Inc.). This regression analysis was exploratory as it did not have the capacity to take clustering into account. Table 4.3.4 indicates that the intervention effect was greater at the higher quantiles. The cutpoint was tested at various levels taking account of clustering. In effect, a score of 27 in the control arm represents the 90th percentile, but in the intervention group the same score represents the 80th percentile.

⁴⁴ ICC : Intraclass Correlation Co-efficient. See 3.3.3.2 for description of this

Quantile	Control Score	Coefficient (intervention effect)	Bootstrap Std Error	t	P>t	95% Confidence Interval	
						Low	High
q10	14	0	0.77	0.00	1.00	-1.51	1.51
q20	16	0	0.70	0.00	1.00	-1.36	1.36
q30	18	0	0.85	0.00	1.00	-1.67	1.67
q40	20	0	0.79	0.00	1.00	-1.54	1.54
q50	21	1	0.60	1.65	0.10	-0.19	2.19
q60	22	2	0.87	2.30	0.02	0.30	3.70
q70	24	2	0.79	2.52	0.01	0.44	3.55
q80	25	3	0.91	3.28	0.00	1.20	4.79
q90	27	5	0.80	6.23	0.00	3.43	6.57

Table 4.3.4 Simultaneous quantile regression at deciles 10 – 90 for the total partograph score

In those records where a partograph was used, the estimated difference between arms was 13.6% (95% CI : 0.16 to 0.25), and the p -value was 0.026 (95% CI : 0.16 to 0.25) (Table 4.3.5). This is a statistically significant difference between the two arms at this level (q90).

Control	Risk difference	Semi-robust Std Error	95% Confidence interval		p
			Low	High	
0.131	0.136	0.613	0.16	0.25	0.026

Table 4.3.5 The risk difference at a score of 27 on the partograph

Table 4.3.6a illustrates that the intervention arm had 19.3% of partographs obtaining scores of 27 or higher compared to 9% in the control arm.

Partograph score ≥ 27	Arm		Percentage of overall total
	Intervention n (% of arm)	Control n (% of arm)	
No	436 (80.7%)	437 (91.0%)	873 (85.6%)
Yes	104 (19.3%)	43 (9.0%)	147 (14.4%)
Percentage of overall total	540 (52.9%)	480 (47.1%)	1020 (100.0%)

Table 4.3.6a Comparison of the partograph scores for the total trial population at the cut point score of 27

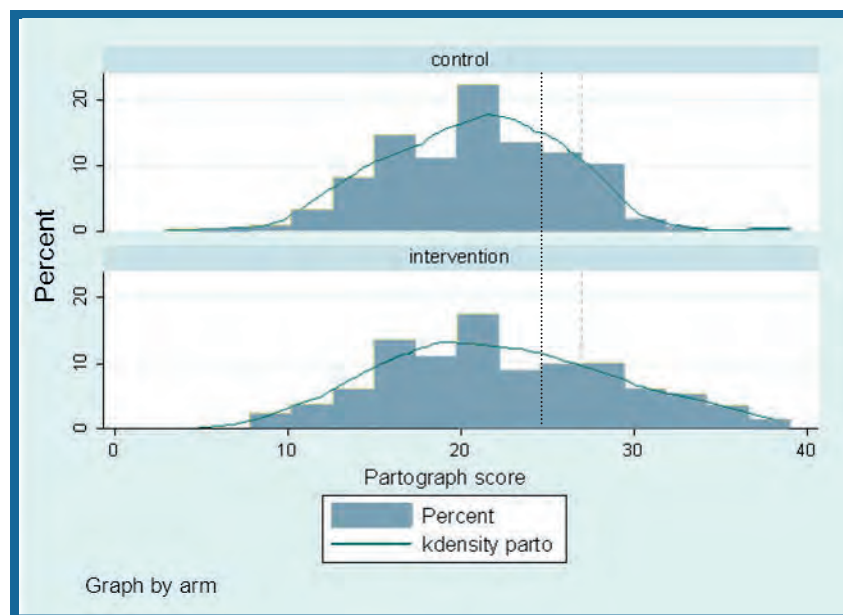
Having determined the cutpoint of 27 where the intervention effect was expressed, a comparison of the partograph for the total trial population was done. The distribution between arms (adjusted for clustering) and different cutpoints is reflected in table 4.3.6b (an expansion of Table 4.3.6a) which illustrates that the proportions were similar for the levels up to the cutpoint of 27.

Partograph score	Arm		Percentage of overall total
	Intervention	Control	
Not used	152 (28.1%)	154 (32.1%)	306 (30.0%)
1 – 17	101 (18.7%)	91 (18.9%)	192 (18.8%)
18 – 26	183 (33.9%)	192 (40.0 %)	375 (36.8%)
27 – max	104 (19.3%)	43 (9.0%)	147 (14.4%)
Percentage of overall total	540 (52.94%)	480 (47.06%)	1020 (100.00%)

Table 4.3.6b Comparison of the partograph scores for the total trial population

In figure 4.3.5, the cutpoint at a score of 27 is indicated and the difference in distribution in the upper scores between the two arms is highlighted.

Figure 4.3.5 Distribution of the partograph score with estimated data-dependent distribution⁴⁵



⁴⁵ The 'k density parto' represents the Kernel density estimation method for the estimated distribution of scores.

In conclusion, there was parity for the partograph score in the lower levels of completion, but a significant difference in those of better quality completions.

4.3.2.1 *Post hoc secondary analysis*

Post hoc secondary analysis was done to determine the effect of time, the regional stratum, and the size of the service on the scores. Testing for interaction⁴⁶ aims to determine the pattern of homogeneity of the intervention effect. This model investigated various interactions between the covariates. The interaction between size and group/arm was significant and was retained in the model. Each of the coefficients is presented (Table 4.3.7).

Partograph score in relation to	Coefficient	95 % Confidence Interval		ρ
		Low	High	
Group / control arm	-3.4	-9.4	2.6	0.269
Time : continuous in months	-0.2	-0.3	0.0	0.085
Region : 2	-5.5	-9.7	-1.4	0.009⁴⁷
Size : moderate – large	-7.6	-13.6	-1.6	0.013
Size by group : moderate-large x intervention	9.8	1.5	18.1	0.021
Intercept	21.6	16.8	26.5	0.000

Table 4.3.7 Linear mixed effects model of partograph score (adjusted for 17 clusters)

4.3.2.1.1 *The effect of size of the service*

The estimated intervention effect between the two sizes of service was due to significant interaction. When analysed further, a statistically significant difference was found between arms in the moderate-to-large sites where the intervention arm scored 6.4 units more than the control arm, ($\rho=0.029$ (95% CI : 0.7 to 12.1)) (Table 4.3.8). With the maximum possible partograph score being 47, this equates to a 13.6% difference. Although small services in the intervention arm scored 3.4 units less (95% CI 2.6 to 9.4), this was not statistically significant ($\rho=0.269$).

⁴⁶ Interaction (in a statistical model) – where the effect of two or more variables is not simply additive.

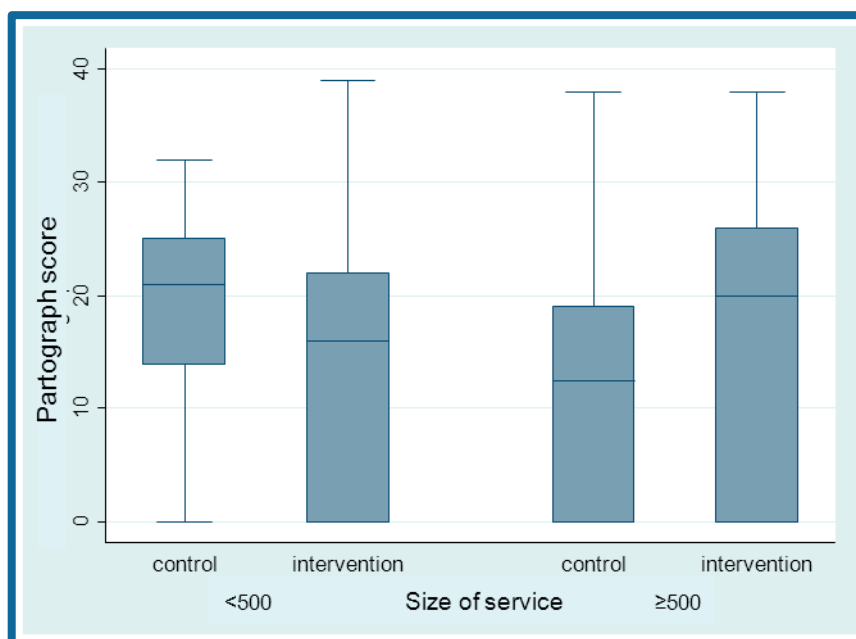
⁴⁷ Highlighted figures indicate statistical significance. This also applies to other tables.

Partograph score per arm in relation to	Coefficient	95% Confidence Interval		ρ
		Low	High	
Small number (<500) births per annum	-3.4	-9.4	2.6	0.269
Moderate – large number (500 - >1000) births per annum	6.4	0.7	12.1	0.029

Table 4.3.8 Partograph scores per arm and per size of service

Figure 4.3.6 illustrates the pattern of scores where the small services in the control arm are on a par with the scores in the larger services in the intervention arm. This may be explained in part by the fact that the small control sites had a good distribution of scores whereas the other three categories had a large number of zero scores.

Figure 4.3.6 Box plot of partograph scores per arm and per size of service

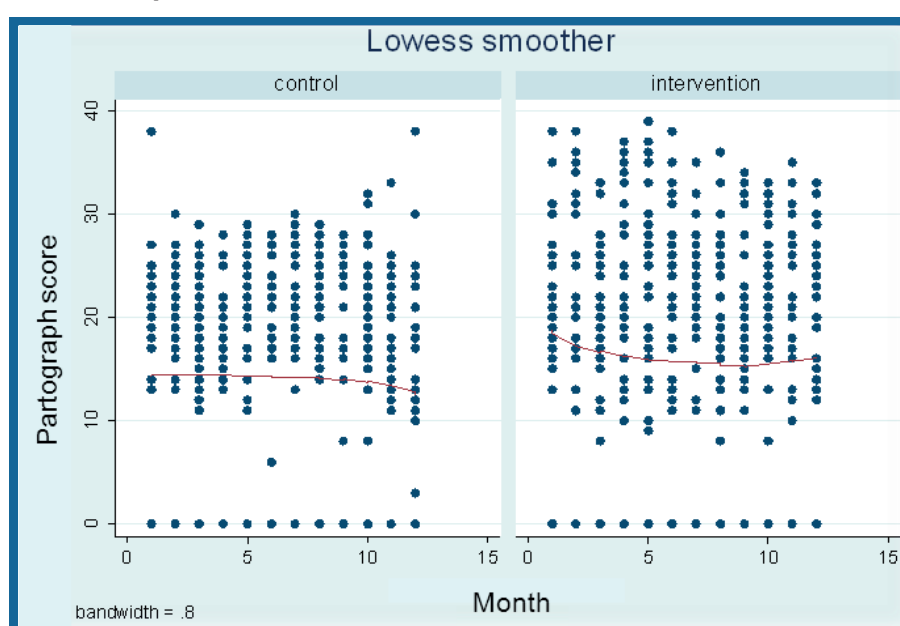


4.3.2.1.2 *The effect of time*

The fact that the partographs were collected over a 12 month period afforded the opportunity to see if there were any particular patterns in the scores in relation to time since the intervention. Lowess is a data analysis technique for

producing a smooth set of values from a time series. Figure 4.3.7 shows that a level of difference between the two arms was achieved and maintained over this period. The profiles were generally parallel over time. The pattern illustrates that the control group had a lower level of scores which, after a plateau, fell off in months 10-12, whereas in the intervention group the level started higher, stabilised in month five and then rose at months 11 and 12. Given that the intervention arms received training in months one and two and testing a month later, this could account for the higher levels in the first few months.

Figure 4.3.7 Comparison of scores between arms over time⁴⁸



Using mixed effects regression analysis, the effect over time was 0.2 marks with $p=0.085$ (95% CI : -0.3 to 0.0) (Table 4.3.7), meaning that the partograph scoring was consistent for the 12 months of the study and there was no interaction effect over time. This illustrates further underlines that the randomisation was balanced.

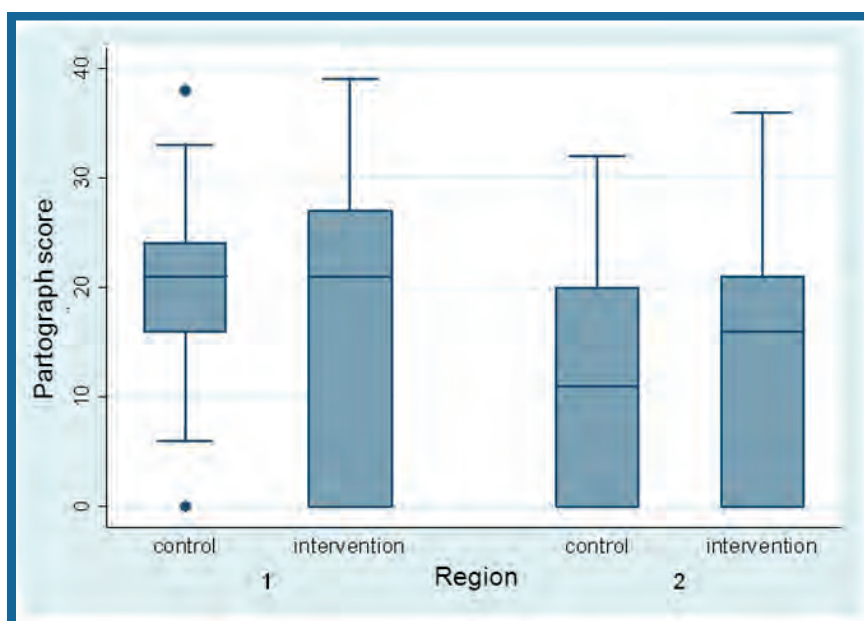
4.3.2.1.3 *The effect of region*

Figure 4.3.8 shows that in region one the median score was maintained across both arms and there was a very wide range for the intervention group with

⁴⁸ Bandwidth refers to the size of the smoothing window. 0.8 means that 80% of the horizontal axis variable is covered.

higher scores in the upper quartile range. In region two the median position improved in the intervention arm and the scores were found in a higher range. Overall, region one had higher mean scores than region two.

Figure 4.3.8 Box plot of partograph scores per arm and per region

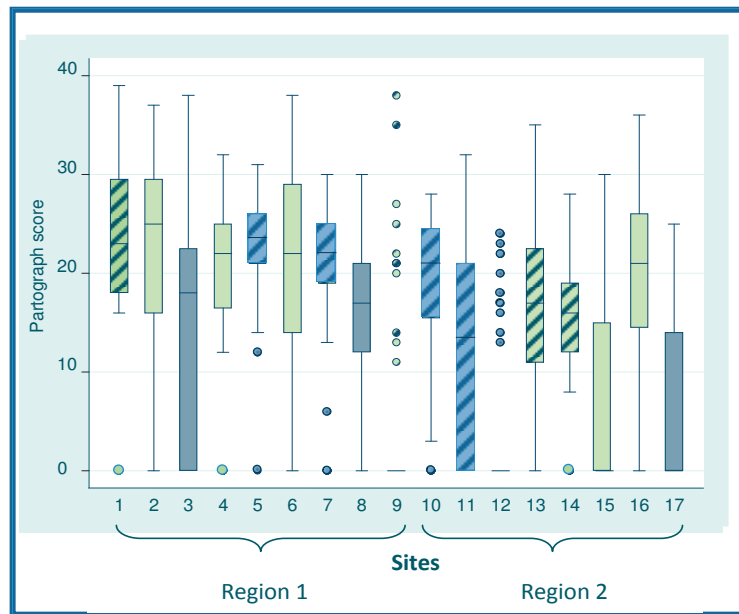


Using the linear mixed effects regression analysis the effect for the regions was -5.5 marks with $p=0.009$ (95% CI : -9.7 to -1.4) (Table 4.3.7), thus indicating that there was a statistically significant difference between the partograph scores of the two regions, with region two obtaining lower scores, but these were independent of the intervention.

4.3.2.1.4 *Intra-cluster correlation coefficient (ICC)*

The boxplot of scores obtained across the 17 sites (Figure 4.3.9) indicates that in four sites at least 25% of partographs obtained a score of zero. There were two sites in particular - both in moderate-to-large sites and one in each arm of the study (sites nine and twelve) – where there were high levels of non-use of the partograph. This impacted the ICC which represents the variation between clusters. In this instance the ICC was 0.16, which is high. This differed from the ICC of 0.28 determined for the partograph score in Table 4.3.3 but this change was due to the fact that a number of adjustments had been made to account for possible confounders.

Figure 4.3.9 Boxplot of partograph scores for all 17 sites



Legend

Intervention sites	Control sites	Small sites	Moderate-to-large sites
			
			Unhatched

4.3.2.2 Conclusion for partograph utilisation score

Overall, in respect of the partograph utilisation score the following conclusions can be made :

- ◆ The utilisation of partographs was higher in the intervention group, but did not reach statistical significance.
- ◆ The mean scores for the total partograph were not different between the two arms.
- ◆ In an intention-to-treat analysis which categorised the partographs into four categories (not used, and scores of 1-17, 18–26 and 27-max) there was a significant difference in the categorical distribution of the two arms.
- ◆ The proportion of partographs with a score of 27 or more was significantly greater in the intervention arm.
- ◆ The intervention effect was consistent over time after the intervention.

- ◆ Regional differences were statistically significant in favour of region one, and independent of the intervention.
- ◆ Size of the service - the intervention effect was not homogeneous across the sizes of service. There was no intervention effect in the small services, but a statistically significant effect was observed in the moderate-to-large services.

4.3.3 Sector scores

In addition to the total partograph score, sector scores were measured in order to identify trends of weakness and strength in order to identify future training and management interventions. These results were derived from the records where partographs were used (n=714). The descriptive statistics are presented in Table 4.3.9.

Variable (max. possible score)	Arm	N	mean	sd	min	p25	p50	p75	max
Medico-legal (8)	Intervention	388	6.1	1.2	2	5	6	7	8
	Control	326	6.1	1.3	2	5	6	7	8
Maternal observations (8)	Intervention	388	4.3	2.7	0	2	4	7	8
	Control	326	4.8	2.7	0	2	6	7	8
Fetal observations (5)	Intervention	388	4.0	1.1	0	3	4	5	5
	Control	326	3.8	1.0	0	3	4	4	5
Progress of labour (6)	Intervention	388	4.1	1.0	1	3	4	5	6
	Control	326	3.8	1.0	0	3	4	4	6
Combined observations (19)	Intervention	388	12.4	3.7	1	10	12	16	19
	Control	326	12.5	3.4	0	10	13	15	19
Recognition of Risk status (3)	Intervention	388	1.1	0.9	0	0	1	2	3
	Control	326	0.8	0.8	0	0	1	1	3
Alignment with guidelines (15)	Intervention	388	2.4	2.3	0	1	2	4	11
	Control	326	1.7	1.9	0	0	1	3	12
Frequency of observations (10)	Intervention	388	1.2	1.7	0	0	0	2	8
	Control	326	0.9	1.3	0	0	0	2	7

Table 4.3.9 Descriptive statistics for scores for the completion of the partograph in respect of sector scores

The scales for the first four variables and ‘recognition of risk status’ were too limited to use as continuous data hence a linear regression model could not be used. Therefore ordered logistic regression was performed and Odds Ratios were determined with adjustment for cluster (Table 4.3.10). None of these Odds Ratios were statistically significant.

Variable	Odds Ratio	Robust Std Error	95% Confidence Interval		p	ICC
			Low	High		
Medico-legal	1.0	0.3	0.6	1.8	0.910	0.26
Maternal	0.8	0.3	0.4	1.4	0.388	0.24
Fetal	1.4	0.7	0.6	3.7	0.442	0.25
Progress of labour	1.6	0.9	0.6	4.6	0.368	0.27
Recognition of risk	1.7	0.5	1.0	2.9	0.064	0.13

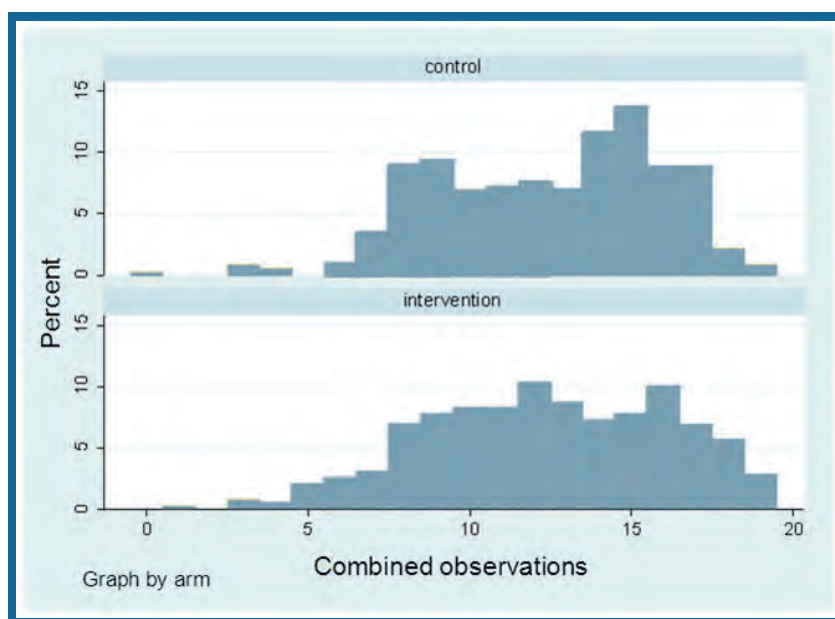
Table 4.3.10 Odds Ratios for various subsector scores (with standard error adjusted for clusters)

The results for the remaining three variables – ‘combined observations’, ‘alignment with clinical management guidelines’ and its subscore, ‘frequency of observations’ - are presented in 4.3.3.1 – 4.3.3.3.

4.3.3.1 Combined observations

Due to the scales for maternal observations, fetal observations and progress of labour variables being very small, it was decided to group these together and see if there was a discernible difference where the scale was larger. The descriptive statistics are presented in table 4.3.9 above and are graphically presented in figure 4.3.10 below.

Figure 4.3.10 Distribution of combined observations scores



There was some evidence of a bimodal distribution of these data (an artefact of the maternal urine observations which had an ‘all or nothing’ pattern), which was more pronounced in the control arm.

Variable (max possible score)	Intervention	Control	Coefficient (intervention effect)	95% Confidence Interval		ρ
				Low	High	
Combined observations (19)	12.40	12.46	-0.06	-1.7	1.5	0.940

Table 4.3.11 Difference in combined observations (maternal, fetal and progress of labour) between arms, adjusted for 17 clusters

The mean difference in the score was 0.06 with $p = 0.94$ (95% CI : -1.7 to 1.5) (Table 4.3.11). The difference relative to the total score using the overall scale was 0.3%. The conclusion remained unchanged with there being no statistically significant difference in the mean score for combined observations between arms.

4.3.3.2 Adherence to clinical management guidelines

The score for 'alignment to clinical guidelines' included the appropriate use of fluids and drugs, appropriate clinical conclusions and decisions including referral, as well as frequency of observations. The descriptive statistics for 'alignment with clinical guidelines' and its sub-score 'frequency of observations' can be found in Table 4.3.9 above. The boxplot for the former is presented in Figure 4.3.11.

Figure 4.3.11 Boxplot of 'alignment to clinical management guidelines' score

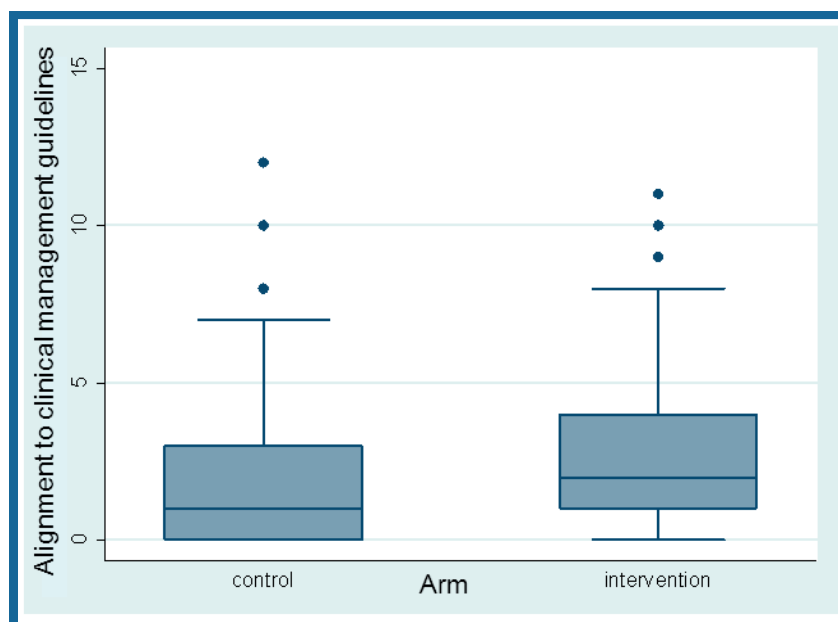


Table 4.3.12 compares the scores for these two variables.

Variable (max possible score)	Intervention	Control	Coefficient (intervention effect)	95% Confidence Interval		p	ICC
				Low	High		
Alignment to clinical guidelines (15)	2.4	1.72	0.7	0.2	1.3	0.014	0.13
Frequency of observations (10)	1.2	0.9	0.3	0.0	0.6	0.037	0.07

Table 4.3.12 Comparison of scores obtained in respect of 'alignment to clinical guidelines' and 'frequency of observations'

The difference between study arms in respect of the alignment of practice to the clinical management guidelines was significant ($p=0.014$) with the intervention effect being 0.7 (95% CI : 0.2 to 1.3). The medians of the two groups were different by one score unit favouring the intervention group (Figure 4.3.11). However, the scores in both arms were low. The difference expressed as a percentage of the overall scale was 4.9%.

‘Frequency of observations’ was a reduced score from the ‘adherence to clinical guidelines’, so it is not unexpected that the significant difference ($p=0.037$) between the two arms was maintained. For this variable the intervention effect was less (0.32) (95% CI : 0.0 to 0.6). Both arms show a poor score (out of a maximum of 10 points).

4.3.3.3 Conclusion regarding sector scores

This section of results looked at trends rather than significance. In general the scores were low with the median percentage score lower than 50% in both arms for the following variables - alignment with clinical guidelines, frequency of observations, recognition of risk status and comprehensive assessment. There were positive differences in most parameters but these were found to be not significant in many cases. Incrementally they contributed to a positive trend in the intervention arm. The differences for ‘alignment with clinical guidelines’ and ‘frequency of observations’ were statistically significant between the two arms of the study in favour of the intervention arm.

4.3.4 Scores related to specific innovations

There were two specific innovations introduced at this time. These were the observation and recording of the fetal heart rate and pattern before and after a contraction (rather than one or other), and a requirement to do and record a critical comprehensive assessment and management plan every four hours (or more frequently if indicated). The instructions for these were included in the written guidelines so were available to all sites through the usual dissemination strategy. These two scores (which were pre-specified) were treated as ordinal data and the Odds Ratios determined (Table 4.3.13).

Variable	Odds Ratio	Robust Std Error	95% Confidence Interval		ρ	ICC
			Low	High		
Fetal heart obs pre- AND post-contraction	2.0	0.3	0.6	6.3	0.237	0.32
Critical comprehensive assessment	12.6	7.3	4.1	39.2	<0.001	0.32

Table 4.3.13 Odds Ratios for innovations

4.3.4.1 Fetal heart observations

The Odds Ratio was 2.0, meaning that the odds of the intervention arm using this pattern of observation was twice that of the control arm. However, this was not statistically significant ($\rho = 0.237$) and the confidence interval was wide. Table 4.3.14 illustrates that the uptake of this innovation was greater in the intervention arm with 8.8% of records completed as required compared to 1.2%. Of those records where there was any evidence of the pre- AND post-contraction fetal heart observations being done, the intervention arm showed a higher percentage adoption of this policy (64.2% compared to 50%).

Fetal observations done pre- AND post-contraction	Arm		Percentage of overall total
	Intervention	Control	
Never	139 (35.8%)	163 (50.0%)	302 (42.3%)
Sometimes	215 (55.4%)	159 (48.8%)	374 (52.4%)
≥75% of the time	34 (8.8%)	4 (1.2%)	38 (5.3%)
Percentage of overall total	388 (100%)	326 (100%)	714 (100%)

Table 4.3.14 Comparison between arms of pre- AND post-contraction fetal heart observations in the records with partographs used

4.3.4.2 Critical comprehensive assessment and management plan

The critical comprehensive assessment and management plan required the midwife/clinician to review the course of the labour over the preceding four hours (or earlier if this was clinically indicated) and make a decision about the future management. This could require following up with the next comprehensive assessment sooner, if indicated. It is possible that this

assessment might not be clinically indicated, e.g. due to the birth or referral occurring before the first four hourly assessment was due to be done.

Of the 1020 partograph records sampled

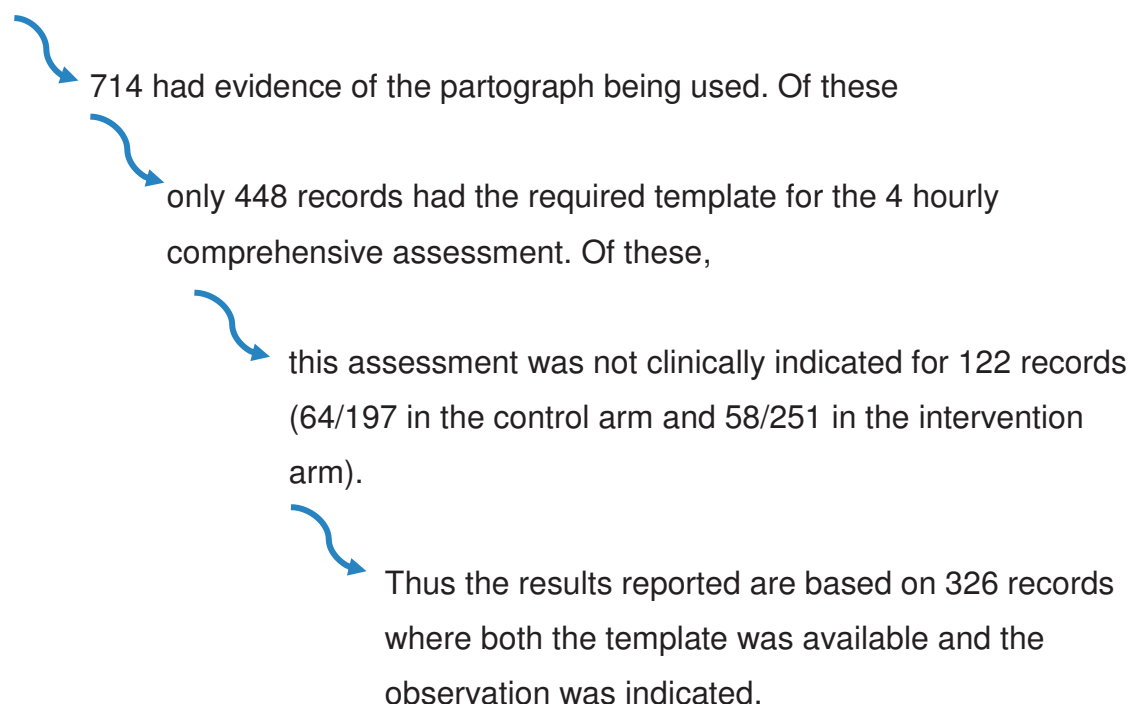


Table 4.3.15 shows that the median for the control group was zero and for the intervention group was two (which represents a percentage score of 50%). The score of four at the 75th centile indicates that at least 25% of the (eligible) intervention arm records had completed comprehensive assessments compared to 0% in the control arm.

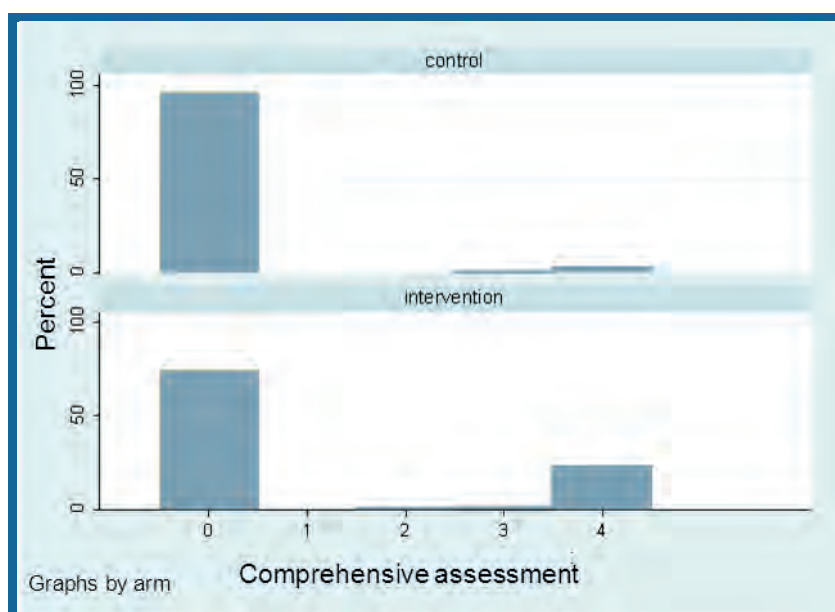
Variable (max. possible score)	Arm	N	mean	sd	min	p25	p50	p75	max
Comprehensive assessment (4)	Intervention	193	2.0	2.0	0	0	2	4	4
	Control	133	0.3	1.0	0	0	0	0	4

Table 4.3.15 Descriptive statistics for scores for the critical comprehensive assessment

The distribution of the scores illustrates this pattern (Figure 4.3.12). This ‘all or nothing’ picture is to be expected where there is a large number of records

without partographs and another group of records missing the required template.

Figure 4.3.12 Distribution of scores for the critical comprehensive assessment



The Odds Ratio of 12.6 (Table 4.3.13 above) indicates that the intervention arm was much more likely to have the comprehensive assessment completed (where indicated and where there was a template) than the control arm, and that this was statistically significant ($p < 0.001$).

Table 4.3.16 illustrates that 47.2% of the intervention arm had this assessment completed appropriately as opposed to 6.8% in the control arm.

Comprehensive assessment	Arm		Percentage of overall total
	Intervention	Control	
<i>Not indicated</i>	58	64	122
Not used	95 (49.2%)	123 (92.5%)	216 (66.9%)
Partially completed	7 (3.6%)	1 (0.7%)	8 (2.4%)
≥75% used	91 (47.2%)	9 (6.8%)	100 (30.7%)
Percentage of overall total	193 (100%)	133 (100%)	326 (100%)

Table 4.3.16 Comparison of critical comprehensive assessment for those records where this was indicated

Overall there was a difference between the arms in an aspect that was an innovation (*viz.* four hourly critical comprehensive assessment and management plan) at the time of the guidelines being introduced to these regions. This indicates that the uptake of this innovation was much higher in the intervention arm than in the control arm.

4.3.5 Contextual reasons for not completing the partograph

During the intervention training, it was emphasised that if observations could not be done for various operational reasons, these reasons should be indicated on the partograph, rather than leaving it blank. This data was collected to give context and explanation, rather than to draw inferences from it. This data did not form part of the total score for the partograph and so, although there were 1020 records evaluated, there was no information regarding this section on a few of the data sheets, thus giving a total of 1009 instead of 1020. These data are summarised in Table 4.3.17.

Reasons recorded	Arm		Percentage of overall total
	Intervention	Control	
Yes	41 (7.6%)	19 (4.0%)	60 (6.0%)
No	498 (92.4%)	451 (96.0%)	949 (94.0%)
Percentage of overall total	539 (100%)	470 (100%)	1009 (100%)

Table 4.3.17 Partographs where reasons were recorded for non-completion of the partograph

Table 4.3.18 shows that there was a statistically significant difference between the arms, with the intervention arm showing a 3.6% higher response than the control.

Control	Risk difference	Semi-robust Std Error	95% Confidence Interval		p
			Low	High	
0.04	0.36	0.17	0.00	0.07	0.041

Table 4.3.18 Comparison between arms of the study in respect of recording reasons for non-completion of the partograph

4.3.5.1 Reasons recorded on the partograph

Multiple reasons could be recorded on a partograph so table 4.3.19 reflects more responses than the number of partographs recorded above.

	Reason	Intervention (40 respondents)	Control (19 respondents)
Operational	Very busy	0	3
	Ward busy	7	7
	With other patients	4	1
	Doing a delivery	12	2
	Shift handover	0	1
	Doctors' rounds	3	0
	Lunch / tea break	0	1
	In another department	5	3
	Observations done by doctor and not recorded/incomplete	2	1
	Subtotal	33	16
Patient related	Not in labour	2	0
	Mobilising / In bath	7	4
	Eating / sleeping	8	0
	Visitors	1	0
	Unco-operative / refused	3	2
	Elsewhere	2	0
	Subtotal	23	6

Table 4.3.19 Reasons recorded for observations not being done

4.3.5.2 Time of birth in relation to last observation

If the birth or transfer took place within 1.5 hours⁴⁹ of the last observation, this was recorded by the partograph evaluator. If the birth did take place within 1.5 hours of the last observation, it was deemed acceptable in the context of this study⁵⁰.

⁴⁹ This time period was crudely determined based on the frequency of observations in active phase of labour being half-hourly, at which time the midwife might find the woman in second stage of labour, which could last approximately one hour.

⁵⁰ This does not imply that this period of absence of recorded observations is acceptable in practice.

If incomplete, did birth / transfer occur within 1.5 hours of the last observation?	Arm		Total
	Intervention	Control	
Yes	146 (29.4%)	143 (31.8%)	289 (30.6%)
No	108 (21.8%)	99 (22.0%)	207 (21.9%)
Not applicable (observations not missing during this period)	93 (18.8%)	66 (14.7%)	159 (16.7%)
No record (no partograph used)	149 (30.0%)	142 (31.5%)	291 (30.8%)
Total	496 (100%)	450 (100%)	946 (100%)

Table 4.3.20 Births occurring within 1.5 hours of the last observation

Table 4.3.20 shows that where the records were incomplete, the birth occurred within 1.5 hours of the last observation in 289 labours (30.6%). The proportions were similar for both arms. This also highlights that, over and above the records where no partograph was used, a further 207 records (21.9%) where a partograph was used had more than 1.5 hours without observations. The proportions were similar in both arms.

4.3.5.3 Time of birth in relation to the commencement of the partograph

A factor that could have contributed to a partograph appearing to be complete could have been a birth occurring within a short period after the admission. This could have had the effect of inflating the score for completed partographs.

If only one set of observations could be expected to be done (in terms of the frequency specified) from the time the partograph was started until the birth had taken place (either within one hour if observations commenced in latent phase, or within half an hour if observations commenced in the active phase), this was recorded by the evaluator (Table 4.3.21).

If complete, could only 1 set of observations be done?	Arm		Total
	Intervention	Control	
Yes	6 (1.1%)	4 (0.8%)	10 (1.0%)
No	3 (0.6%)	3 (0.6%)	6 (0.6%)
Not applicable (partograph not used)	112 (20.8%)	117 (24.4%)	229 (22.5%)
Not relevant (record not complete)	418 (77.5%)	356 (74.2%)	774 (75.9.0%)
Total	480	539	1019

Table 4.3.21 Births occurring before a second set of observations could be done

Only ten out of the 1019 records fell into this category, thus not contributing to an inflated completion rate. There was no significant difference between the two groups using Pearson's χ^2 test ($Pr = 0.559$).

4.3.5.4 Conclusion regarding contextual factors

The contextual information does not illustrate any specific differences between the two arms in terms of timing. It does show a significantly higher response in the intervention arm indicating reasons for the non-completion of the partograph. While limited, this is quite revealing in terms of operational reasons (for all sites).

4.3.6 Partograph utilisation audit conclusion

There was a small and consistent but not statistically significant effect on a number of the elements in this phase of the study which cumulatively resulted in the shift to a higher range. Operational and staffing constraints made it difficult to comply with the requirements for the frequency of observations as set out in the provincial and national guidelines for intrapartum care.

4.4 Midwives' knowledge and skills (Objective 2)

In this section the dynamics of the midwife population during the study period will be described in terms of those joining and leaving the study. The results of test three (conducted at month 12 of the study) will be presented comparing the performance of the two arms. The unit of analysis was at the individual level.

Due to the staffing fluctuations, ‘intention-to-treat’ (ITT) analyses as well as ‘per-protocol analyses’ (PPA) were conducted as appropriate, and are presented. Thereafter, changes within the intervention arm over time will be presented.

4.4.1 Study population dynamics

One hundred and fifty-four registered midwives consented to participate in the study. However, at no single testing timepoint were all 154 involved. This was due to staffing fluctuations throughout the study period, and is consistent with a pragmatic study. These dynamics are illustrated in Figure 4.4.1. Contributing circumstances to these dynamics are discussed in 5.3.1.2. Participants who completed tests one and three and the related descriptive statistics are reflected in Tables 4.4.1 and 4.4.2, indicating similarity across arms.

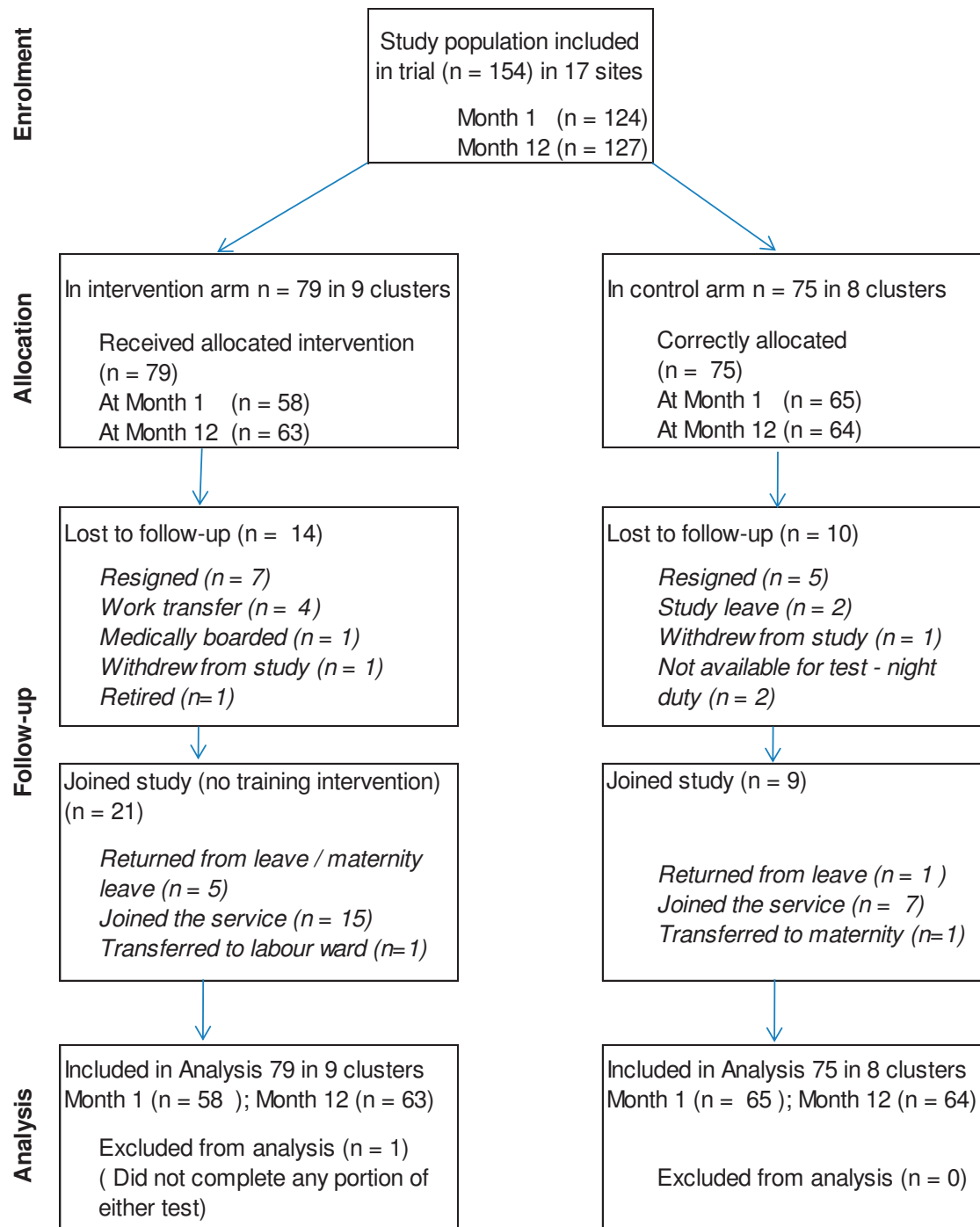
Arm	Sites (code)	Total midwives enrolled	Midwives undertaking	
			Test 1	Test 3
Intervention	1	14	12	13
	2	9	8	7
	4	8	3	6
	6	11	9	8
	9	12	8	11
	13	4	4	3
	14	4	4	3
	15	9	3	7
	16	8	7	5
	Total	79	58	63
Control	3	8	8	7
	5	8	8	5
	7	13	12	12
	8	14	13	13
	10	7	3	6
	11	6	5	6
	12	9	7	7
	17	10	9	8
	Total	75	65	64

Table 4.4.1 Registered midwives who enrolled in the trial (per site) and participated at each timepoint

Number of participants per cluster	N	Arm	Mean	Min	Max
Total	79	Intervention	8	4	14
	75	Control	9.4	6	14

Table 4.4.2 Descriptive statistics of numbers of midwives enrolled per arm

Figure 4.4.1 Participant level (registered midwife) flow diagram ⁵¹



⁵¹ 'Correctly allocated to control' means that the allocation to the control arm was done as specified in the protocol, i.e. there were no protocol violations.

4.4.2 Secondary outcome : Comparison of midwives' final test between arms

One hundred and twenty-seven midwives in 17 clusters undertook the final test (test three) in month 12. Due to the different compositions of the test participants (at different timepoints and for different components of the tests, viz. theoretical and clinical) the analysis is presented in a number of ways :

- ◆ Where there were complete data (theory and clinical components) for test three (month 12) (96 participants in 17 clusters) (ITT)
- ◆ Where there were complete data (theory and clinical components) for test one (baseline) and test three (month 12) (70 participants in 17 clusters) ('per protocol analysis' - PPA)
- ◆ Where only the theory component (104) or clinical component (118) was undertaken for test three (17 clusters)

Figure 4.4.2 provides a results map to assist in navigating section 4.4.2.

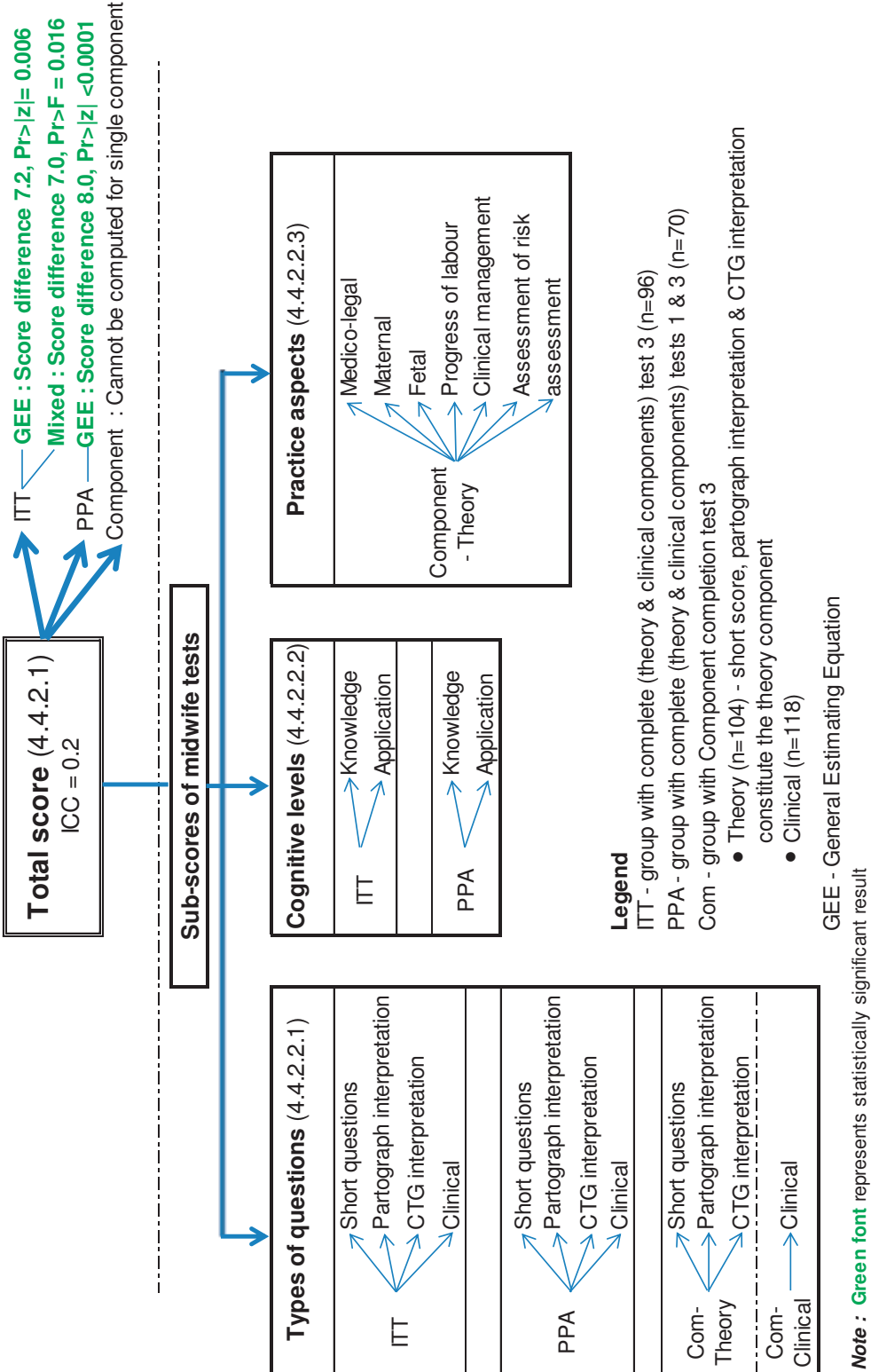
Tables 4.4.3a and b illustrate similar patterns of completions of theory and clinical components for test three for each arm.

		Clinical Complete	Clinical Incomplete	Clinical not undertaken	Total / Percentage of overall total
Theory Complete	n	48 ⁵²	4	0	52
	%	92.3	7.7	0.0	65.8
Theory Incomplete	n	10	1	0	11
	%	90.9	9.1	0.0	13.9
Theory not undertaken	n	0	0	16	16
	%	0.0	0.0	20.3	20.3
Total / Percentage of total	n	58	5	16	79
	%	73.4	6.3	20.3	100

Table 4.4.3a Intervention arm - Test three theory vs clinical completions

⁵² The shaded block comprises the 'intention to treat' (ITT) analysis group.

Figure 4.4.2 : Map of presentation of results and main findings - midwife tests



		Clinical Complete	Clinical Incomplete	Clinical not undertaken	Total / Percentage of overall total
Theory Complete	N	48 ⁵³	4	0	52
	%	92.3	7.7	0.0	69.3
Theory Incomplete	n	12	0	0	12
	%	100.0	0.0	0.0	16.0
Theory not undertaken	n	0	0	11	11
	%	0.0	0.0	100.0	14.7
Total / Percentage of overall total	N	60	4	11	75
	%	80.0	5.4	14.7	100

Table 4.4.3b Control arm - Test three theory vs clinical completions

4.4.2.1 Total scores

The total test score will be presented first. Thereafter, sub-sector scores for types of questions, cognitive levels and practice areas will be presented.

4.4.2.1.1 Test three complete ('intention to treat')

The total score is presented for the sub-group of 96 midwives who did test three where there was complete information for the theory and clinical portions of the test. The descriptive statistics are presented in Table 4.4.4.

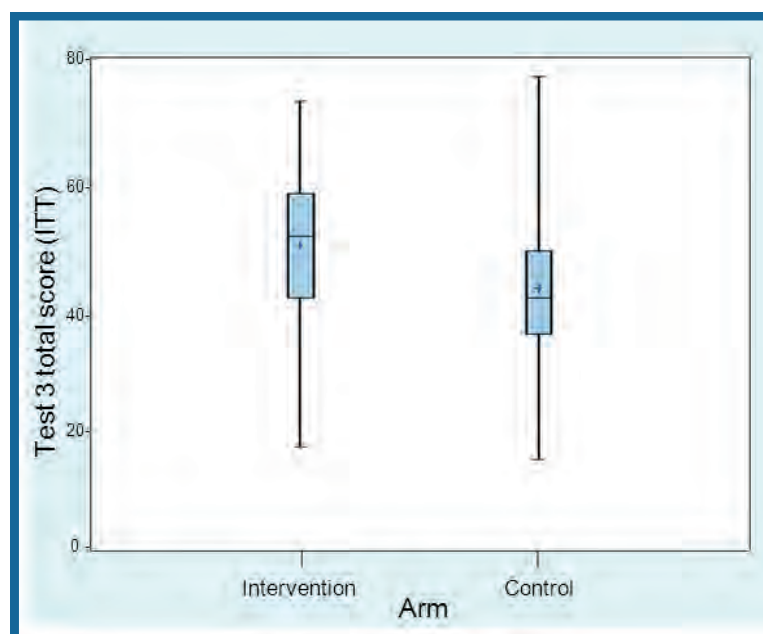
Score (max possible score)	Obs	N	Arm	Mean score	SD	min	max
Total (119)	48	48	Intervention	50.0	12.2	16	74
	48	48	Control	42.7	13.0	14	78

Table 4.4.4 Descriptive statistics for total scores for test three

Total scores for test three are reflected graphically in a box plot in Figure 4.4.3 which illustrates the intervention effect.

⁵³ The shaded block comprises the 'intention to treat' (ITT) analysis group.

Figure 4.4.3 Comparison of test three total scores in intervention and control arms ('intention to treat')



The GEE method was used to compare the two arms for the total score (Table 4.4.5). Given that there was a relatively small number of clusters available, the mixed effects model was also used to check the consistency of the GEE. This takes account of the clustering effect in a different way from the GEE. (This is described in 3.3.5.2).

Score (max possible score)	Obs	N	Test	Difference in scores (intervention – control)	95% Confidence Interval		Pr> z (GEE) Pr>F (Mixed) ⁵⁴	ICC
					Low	High		
Total score (119)	96	96	GEE	7.2	2.1	12.3	0.006	0.2
	96	96	Mixed effects	7.0	1.4	12.7	0.016	

Table 4.4.5 Total scores for test three ('intention to treat')

⁵⁴ The probability value is expressed as Pr>|z| and Pr>F for GEE and Mixed Effects tests respectively

Both these tests indicated a significant intervention effect and a similar degree of effect. In the intervention arm the overall score was 7 points higher than the control equivalent to 6% of the total score (7/119) with $p=0.006$ (95% CI 2.1 to 12.3) for the GEE model and $p=0.016$ (95% CI: 1.4 to 12.7) for the mixed effects model. This showed a moderate effect which was statistically significant. The results obtained from the GEE and mixed effects model indicated a consistency between them, and so the GEE could be used with confidence as the preferred method. The ICC of 0.2 indicated that 20% of the variability in the total midwives' test score could be explained by variability between sites / clusters.

4.4.2.1.2 *Test three complete where test one was also complete (per protocol analysis)*

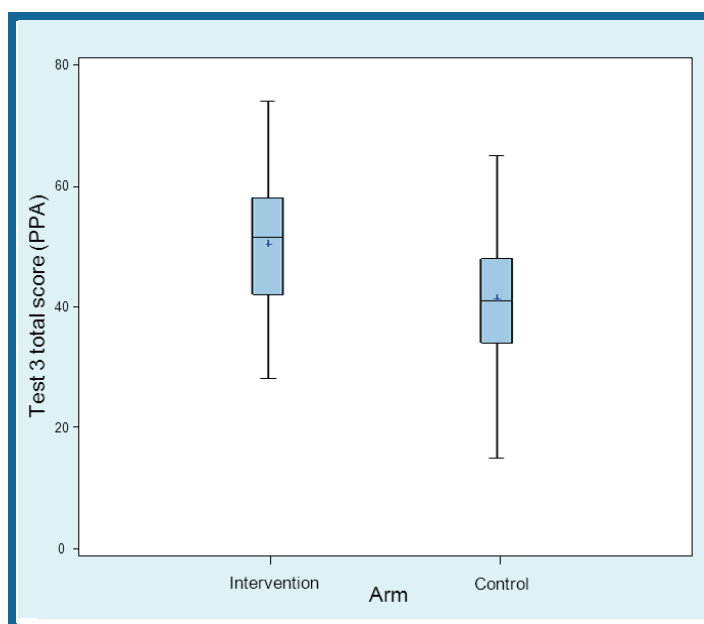
Of the 154 midwives who enrolled, 70 undertook and completed test one (baseline) and test three (month 12) in full. All 17 clusters in the study were represented in this sub-group of midwives. The descriptive statistics are presented in Table 4.4.6.

Test	Max. possible score	Obs	N	Test	Arm	Mean score	SD	min	max
Test 1	129	34	34	1	Intervention	47.8	8.3	30	64
		36	36	1	Control	48.3	11.6	19	69
Test 3	119	34	34	3	Intervention	50.5	10.8	28	74
		36	36	3	Control	42.3	12.2	15	73

Table 4.4.6 Descriptive statistics for test one (baseline) and test three (month 12) total scores ('per protocol')

The divergent pattern speaks both to maintenance and slight improvement in the intervention arm but more particularly, a marked deterioration in performance in the control arm.

Fig 4.4.4 Comparison of test three total scores in intervention and control arms ('per protocol')



The boxplot above (Figure 4.4.4) reflects graphically the descriptive data for test three, and illustrates the intervention effect. The means for test three for each arm were very similar to those in the larger 'intention to treat' group.

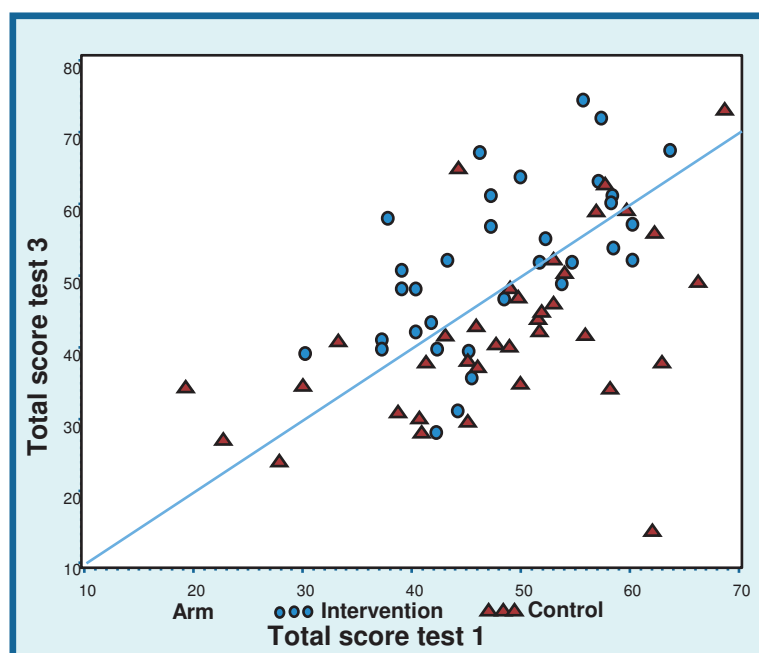
Table 4.4.7 shows a highly significant effect ($p < 0.0001$) with much better precision (95% CI: 4.7 to 11.3) than the intervention effect if adjusted for baseline.

Score	Obs	N	Difference in scores (intervention – control)	95% Confidence Interval		p
				Low	High	
Total score	70	70	8.0	4.7	11.3	<0.0001

Table 4.4.7 GEE result for total score for test three 'per protocol' adjusted for baseline

As the subgroups were not exactly the same, direct comparisons between the two tests could not be done, but are plotted in figure 4.4.5, where the diagonal line represents the line of no change. It can be seen that the control arm scores generally dropped below the line, whereas the intervention arm scores moved slightly above the line.

Figure 4.4.5 Scatterplot of test three versus test one total scores for the two arms



4.4.2.1.3 *Test three component completions*

Of the 127 midwives who were enrolled at the time of test three, there were similar proportions of completions / part completions for both components, but a slightly larger proportion (and identical numbers for each component) of 'not undertaken' in the intervention arm (Tables 4.4.8a and b).

Theory component		Complete	Incomplete	Not undertaken	Total / Overall percentage of total
Intervention arm	n	52	11	16	79
	%	65.8	13.9	20.3	51.3
Control arm	n	52	12	11	75
	%	69.3	16.0	14.7	48.7
Total / Overall percentage of total	n	104	23	27	154
	%	67.5	15.0	17.5	100

Table 4.4.8a Pattern of completions for test three - theory component

Clinical component		Complete	Incomplete	Not undertaken	Total / Overall percentage of total
Intervention arm	N	58	5	16	79
	%	73.4	6.3	20.3	51.3
Control arm	N	60	4	11	75
	%	80.0	5.3	14.7	48.7
Total / Overall percentage of total	N	118	9	27	154
	%	76.6	5.8	17.6	100

Table 4.4.8b Pattern of completions for test three - clinical component

Twenty-seven midwives did not undertake any portion of test three – 16 in the intervention arm and 11 in the control arm (20.2% and 14.6% of these arms respectively). As has been shown in the participant flow diagram (Figure 4.4.1) and in the completions per site (Table 4.4.1), there was no appreciable difference between the two arms regarding loss to follow-up. Given that the pattern was the same for each arm, one can conclude that this would not introduce bias at this point of the study, and no further analysis was done on this aspect.

4.4.2.1.4 Conclusion regarding midwives' total scores (secondary outcome)

Regardless of the type of analysis carried out (ITT or PPA) the results demonstrated consistently higher scores of seven to eight points in the intervention arm compared with the control arm, and these findings attained pre-specified levels of statistical significance.

4.4.2.2 Sub-scores of midwives' tests

The analysis of the tests was broken down in a number of ways to try to identify where there were any particular strengths or weaknesses, so that this could inform any recommendations arising from this study and in particular to direct any changes to the intervention programme. These sub-scores were :

- ◆ the types of questions, e.g. short questions, partograph / case-based questions, cardiotocograph description and interpretation, clinical skills

- ◆ cognitive levels at which the questions were asked, using Bloom's taxonomy. (For the purposes of this analysis the first two categories of the taxonomy were grouped as 'knowledge', and the next three categories were grouped as 'application')
- ◆ assessment and management practice aspects of intrapartum care, e.g. maternal, fetal, recognition of risk.

Having established the GEE as being suitable for this study, the remainder of the analyses will use this single method.

4.4.2.2.1 *Types of questions*

There were four exclusive sections in the tests each of which concentrated on different aspects of intrapartum monitoring. These were :

- ◆ short questions which included definitions, observations and clinical values
- ◆ clinical case-based questions requiring interpretation of a partograph and appropriate clinical management
- ◆ description and interpretation of a cardiotocograph (CTG)
- ◆ clinical observations and skills.

The first three were combined to give a theory score.

The scores for types of question were analysed in the same way as for the total score, viz. GEE model. The analysis will be presented for the different subgroups in this study :

- ◆ the 96 who were analysed on an 'intention to treat' basis (ITT)
- ◆ the 70 who were analysed 'per protocol' (PPA)
- ◆ those who completed theory (104) or clinical (118) components.

4.4.2.2.1.1 *Test three complete – 'Intention to treat'*

In the subgroup of 96 midwives who completed test three in full (intention to treat), the theory score (with its three subsections) and clinical score presented the profiles shown in Table 4.4.9.

Type of question (max. possible score)	Obs	N	Arm	Mean score	SD	min	max
Theory total (83)	48	47	Intervention	39.1	9.5	18	57
	48	43	Control	35.4	10.8	19	60
Short questions (24)	48	47	Intervention	17.5	2.8	11	22
	48	43	Control	15.9	3.1	9	22
Case-based partograph interpretation (49)	48	47	Intervention	18.9	7.5	4	34
	48	43	Control	17.7	8.5	3	34
CTG interpretation (10)	48	46	Intervention	2.7	1.2	1	5
	48	42	Control	1.8	1.0	0	5
Clinical (36)	48	48	Intervention	14.9	4.7	5	23
	48	43	Control	13.1	4.9	4	27

Table 4.4.9 Descriptive statistics for types of questions for test three ('intention to treat')

The mean scores for all types of questions were higher in the intervention arm. By contrast, the maximum scores for the theory component (made up of the short questions, partograph interpretation and CTG analysis) and the clinical component, were both higher in the control group than in the intervention group. The GEE for the theory total score (and one of its subsections, *viz.* CTG interpretation) and clinical score both showed an improvement in the intervention arm which was statistically significant (Table 4.4.10).

Type of question (max possible score)	Obs	N	Difference in scores (intervention – control)	95% Confidence Interval		p
				Low	High	
Theory total (83)	96	93	4.9	1.2	8.5	0.009
Short questions (24)	96	90	1.3	-2.7	0.1	0.06
Case-based partograph interpretation (49)	96	93	2.3	0.0	4.7	0.52
CTG interpretation (10)	96	91	1.1	0.7	1.4	<0.0001
Clinical (36)	96	96	1.95	0.63	3.26	<0.0001

Table 4.4.10 GEE results for types of questions for test three ('intention to treat')

The differences expressed as a percentage of the overall scale were :

♦ theory component total	5.9%
♦ short questions	5.4%
♦ case based – partograph interpretation	4.7%
♦ CTG interpretation	11.0%
♦ clinical component	5.6%

4.4.2.2.1.2 *Test three complete where test one was also complete– ‘Per protocol analysis’*

In the group of 70 midwives who completed tests one and three, the following descriptive statistics apply (Table 4.4.11).

Type of question	Max. possible score	Obs	N	Test	Arm	Mean score	SD	min	max
Theory total	93	34	34	1	Intervention	35.5	6.2	21	47
		36	36	1	Control	35.3	9.4	13	54
	83	34	34	3	Intervention	35.7	7.5	18	53
		36	36	3	Control	30.1	7.7	16	47
Short questions	30	34	34	1	Intervention	18.4	3.0	11	23
		36	36	1	Control	19.3	3.9	10	26
	24	34	34	3	Intervention	18.2	2.9	11	23
		35	34	3	Control	16.0	3.0	9	21
Case-based partograph interpretation	53	34	34	1	Intervention	15.3	5.0	6	26
		36	35	1	Control	15.1	6.1	2	24
	49	34	34	3	Intervention	14.8	4.9	4	29
		35	34	3	Control	12.4	5.1	3	21
CTG interpretation	10	34	31	1	Intervention	1.9	1.4	0	5
		36	36	1	Control	1.8	1.6	0	7
	10	34	33	3	Intervention	2.8	1.2	1	5
		35	33	3	Control	1.7	1.1	0	5
Clinical	36	34	33	1	Intervention	12.7	3.4	8	19
		36	34	1	Control	12.4	3.6	6	19
	36	34	34	3	Intervention	14.8	4.3	5	22
		36	36	3	Control	12.2	4.3	4	20

Table 4.4.11 Descriptive statistics for types of questions for test three ('per protocol')

The section scores were analysed in the same way as for the total score, viz. GEE, adjusted for baseline (Table 4.4.12).

Type of question	Obs	N	Difference in scores (intervention – control)	95% Confidence Interval		p
				Low	High	
Theory total	70	69	5.7	2.17	9.19	<0.0001
Short questions	69	68	2.2	0.7	3.6	0.003
Case-based partograph interpretation	69	68	2.4	0.4	4.4	0.018
CTG interpretation	69	66	1.1	0.6	1.6	<0.0001
Clinical	70	70	2.9	2.34	3.53	<0.0001

Table 4.4.12 GEE results for types of questions for test three ('per protocol')

All sections demonstrated a significant difference between the intervention and control group adjusted for the baseline score, and three of these sections, viz. short questions, CTG interpretation and clinical were highly significant. The CTG interpretation and clinical scores showed not only a deterioration in the control arm but an improvement in the intervention arm. It is acknowledged that this analysis was done on a reduced sample size.

The differences expressed as a percentage of the overall scale were :

- ♦ theory component total 6.9%
- ♦ short questions 8.3%
- ♦ case based – partograph interpretation 4.1%
- ♦ CTG interpretation 11.0%
- ♦ clinical component 8.1%

4.4.2.2.1.3 Theory or clinical component completers

One hundred and four midwives completed test three theory component, but there were some values missing, thus this is based on 101 respondents. One hundred and eighteen midwives completed the clinical component of test three

(see tables 4.4.3a and b above). Table 4.4.13 presents the descriptive statistics. All sections showed a higher score in the intervention arm.

Type of question (max possible score)	Obs	N	Arm	Mean score	SD	min	Max
Theory total (83)	52	51	Intervention	35.7	7.8	18	53
	52	50	Control	31.0	8.4	16	51
Short questions (24)	52	51	Intervention	17.9	3.0	11	23
	52	50	Control	16.4	3.2	9	23
Case-based partograph interpretation (49)	52	51	Intervention	15.1	5.7	4	29
	52	50	Control	12.8	5.5	3	24
CTG interpretation (10)	52	50	Intervention	2.7	1.3	1	6
	52	49	Control	1.8	1.1	0	5
Clinical (36)	58	58	Intervention	15.0	4.6	5	23
	60	60	Control	13.0	4.6	4	27

Table 4.4.13 Descriptive statistics for types of questions for those who completed theory or clinical components

The GEE analysis for this sub-group of midwives is reflected in Table 4.4.14.

Type of question	Obs	N	Difference in scores (intervention – control)	95% Confidence Interval		p
				Low	High	
Theory total	104	101	4.6	1.2	7.9	0.008
Short questions	104	101	1.2	0.2	2.6	0.10
Case-based partograph interpretation	104	101	2.3	0.1	4.4	0.05
CTG interpretation	104	99	1.0	0.7	1.4	<0.0001
Clinical	118	118	2.2	1.1	3.3	<0.0001

Table 4.4.14 GEE results for types of questions for test three for those who completed theory or clinical components

In this large subgroup of midwives the two main components, viz. theory total and clinical, showed highly significant differences between the two arms. The intervention arm achieved 4.6 points more than the control group for the total

theory score with $p = 0.008$ (95% CI: 1.2 to 7.9). The clinical score was 2.2 points more in the intervention arm than the control group with $p < 0.0001$ (95% CI: 1.1 to 3.3). When analysing this larger group ($n = 101$, representing 79.5% of the baseline sample size), where there was no adjustment for the baseline, the short question score was not significantly different between the two arms. However, the other sector scores show a statistically significant difference, most especially for the description and interpretation of the cardiotocograph.

The differences expressed as a percentage of the overall scale were :

◆ theory component total	5.5%
◆ short questions	5.0%
◆ case based – partograph interpretation	4.7%
◆ CTG interpretation	10.0%
◆ clinical component	6.1%

4.4.2.2.1.4 Conclusion regarding types of questions

All three subgroups analysed showed statistically significant differences in favour of the intervention arm for both the theory and clinical components. The differences between arms for each type of question were statistically significant for all types of analysis with the exception of the CTG interpretation in the ITT group and the short questions in the theory component analysis.

4.4.2.2.2 Cognitive levels

The knowledge and application scores were sub-scores of the total score and evaluated the difference in cognitive levels of the questions requiring different skills from the midwives. Cognitive levels were evaluated across the theory and clinical components of the test and so could not be meaningfully analysed or compared for any participants who did not complete both components.

Consistent with what was done previously the results are presented for different subgroups, viz.

- ◆ where theory and clinical components were undertaken for test three (intention to treat analysis - 96 participants)
- ◆ where all components were undertaken at the two testing timepoints (per protocol analysis - 70 participants).

4.4.2.2.1 'Intention to treat' analysis group

In the subgroup of 96 midwives who completed test three in full, the scores for test three were higher in the intervention arm for both cognitive levels (Table 4.4.15).

Cognitive level (max possible score)	Obs	N	Arm	Mean score	SD	min	max
Knowledge (58)	48	48	Intervention	27.5	6.8	11	40
	48	48	Control	23.2	6.6	9	42
Application (61)	48	48	Intervention	22.5	6.4	5	36
	48	48	Control	19.2	7.2	2	37

Table 4.4.15 Descriptive statistics for knowledge and application scores ('intention to treat')

The GEE for the knowledge score showed a highly significant intervention effect of 4.4 with $p=0.0002$ (95% CI : 2.0 to 6.7). The GEE for the application score showed an intervention effect of 3.2 with $p=0.05$ (95% CI: 0.1 to 6.4). This is statistically significant (Table 4.4.16).

Cognitive level	Obs	N	Difference in scores (intervention – control)	95% Confidence Interval		ρ
				Low	High	
Knowledge	96	96	4.4	2.0	6.7	0.0002
Application	96	96	3.2	0.1	6.4	0.05

Table 4.4.16 GEE model results for knowledge and application scores for test three ('intention to treat')

The differences expressed as a percentage of the overall scale were 7.6% for knowledge and 5.3% for application.

4.4.2.2.2.2 'Per protocol' analysis group

The descriptive statistics for the subgroup of 70 midwives who had completed the tests at both timepoints are reflected in Table 4.4.17. This shows that the knowledge score measured at test three increased in the intervention arm but decreased in the control arm compared to baseline, whereas for the application score the intervention arm score increased but the control arm retained the same level in both tests.

Cognitive level	Max possible score	Obs	N	Test	Arm	Mean score	SD	min	max
Knowledge	71	34	34	1	Intervention	26.3	5.3	16	37
		36	36	1	Control	26.6	6.1	13	37
	58	34	34	3	Intervention	28.0	6.1	17	38
		36	36	3	Control	22.7	6.2	9	40
Application	58	34	34	1	Intervention	18.8	4.5	10	28
		36	36	1	Control	19.1	6.6	5	33
	61	34	34	3	Intervention	22.6	5.8	11	36
		36	36	3	Control	19.2	6.7	5	33

Table 4.4.17 Descriptive statistics for test one (baseline) and test three (month 12) knowledge and application scores ('per protocol')

The GEE for the knowledge score showed an intervention effect of 4.6 with $p < 0.0001$ (95% CI : 3.6 to 5.6). The GEE for the application score showed an intervention effect of 3.5 with $p = 0.02$ (95% CI: 0.5 to 6.5). Thus both scores showed a significant effect (Table 4.4.18).

Cognitive level	Obs	N	Difference in scores (intervention – control)	95% Confidence Interval		p
				Low	High	
Knowledge	70	70	4.6	3.6	5.6	<0.0001
Application	70	70	3.5	0.5	6.5	0.02

Table 4.4.18 GEE model results for knowledge and application scores for test three (per protocol analysis) (adjusted for baseline score)

The differences expressed as a percentage of the overall scale were 7.9% for knowledge and 5.7% for application.

4.4.2.2.3 Conclusion for cognitive level results

The mean scores for the control and intervention arms for the two cognitive levels were similar. Both subgroups analysed showed higher knowledge scores in the intervention arm which were statistically highly significant. The differences between arms for the application scores for both subgroups were statistically significant.

4.4.2.2.3 Practice aspects

The third way of interrogating the data was to look at practice aspects⁵⁵. The practice aspect scores were not mutually exclusive scores as some features are common to more than one score e.g. fetal distress could be a component of 'fetal' and 'recognition of risk' and 'critical management decisions'. Similar categories to those discussed in the partograph utilisation audit were used.

One hundred and four midwives undertook the theory component of test three. Three of these had missing data so the analysis was done on 101. As has been seen in other sections in which other breakdowns of the scores have been analysed, the descriptive statistics in table 4.4.19 show that the scores tended to be low, often not reaching a 50% 'pass mark'.

⁵⁵ Due to the difficulties with different completions of theory and clinical components and because the clinical score was a smaller component, which only contributed to two of the six practice aspects, only the theory score data is presented.

Type of question (max possible score)	Obs	N	Arm	Mean score	SD	min	max
<i>Theory total</i> ⁵⁶ (83)	52	51	<i>Intervention</i>	35.7	7.8	18	53
	52	50	<i>Control</i>	31.0	8.4	16	51
Maternal (26)	52	51	Intervention	12.7	2.9	6	19
	52	50	Control	11.5	2.8	6	18
Fetal (17)	52	51	Intervention	7.6	2.3	3	13
	52	50	Control	6.5	2.4	2	11
Progress of labour (25)	52	51	Intervention	12.2	3.3	4	20
	52	50	Control	10.4	3.6	2	17
Clinical management (15)	52	51	Intervention	4.3	2.4	1	10
	52	50	Control	3.8	2.5	1	13
Recognition of risk (11)	52	51	Intervention	2.9	1.6	1	7
	52	47	Control	2.4	1.1	1	6
Critical management decisions (7)	52	43	Intervention	2.1	0.9	0	4
	52	42	Control	1.9	1.4	0	6

Table 4.4.19 Descriptive statistics for test three theory component and practice aspect scores ('intention to treat')

The GEE analysis (Table 4.4.20) showed that three of the six sections achieved a significant difference between the two arms – maternal, fetal, and progress of labour.

⁵⁶ The theory total results have been reported in Tables 4.4.13 and 4.4.14 but are included in Tables 4.4.19 and 4.4.20 for ease of reference, and are differentiated by an italic font.

Type of question	Obs	N	Difference in scores (intervention – control)	95% Confidence Interval		p
				Low	High	
Theory total	104	101	4.6	1.2	7.9	0.008
Maternal	104	101	1.2	0.2	2.4	0.047
Fetal	104	101	1.1	0.2	2.1	0.022
Progress of labour	104	101	1.7	0.4	3.1	0.014
Clinical management	104	101	0.5	0.5	1.6	0.337
Recognition of risk	104	98	0.5	0.1	1.1	0.083
Critical management decisions	104	85	0.2	0.3	0.7	0.430

Table 4.4.20 GEE results for test three theory component and sector scores ‘intention to treat’

The differences relative to the total score using the overall scale for these scores were :

♦ Maternal	4.6%
♦ Fetal	6.5%
♦ Progress of labour	6.8%
♦ Clinical management	3.3%
♦ Recognition of risk	4.5%
♦ Critical management decisions	2.8%

4.4.2.2.3.1 Conclusion for practice aspects

The aspects relevant to clinical decision-making and more complex cognitive behaviours are the aspects whose differences between arms are not statistically significant.

4.4.2.3 Determining sensitivity of results

In order to determine if there was introduction of any bias due to the different analysis groups (i.e. ITT or PPA) or by analysing the theory and clinical portions of the tests separately, comparisons were made to determine the sensitivity of the results across these various groupings. Where compared scores fell within

one score unit (approximately 1%), this was regarded as comparable, i.e. no bias. Three sets of data are presented (all baseline scores, PPA group, and all test three scores) and a conclusion stated. Theory scores are presented, followed by clinical scores.

4.4.2.3.1 Theory scores

The maximum possible theory score for test one was 93 and for test three, 83.

Group	Obs		N	Mean Score	Std dev	Min	Max
Intervention	58		58	36.0	7.0	19	56
Control	61		61	35.9	9.5	13	54

Table 4.4.21 : All subjects with baseline scores - Test one (Baseline randomisation)

Group	Obs	Variable	N	Mean score	Std dev	Min	Max
Intervention	34	Theory T1	34	35.8	6.2	21.0	47.0
	34	Theory T3	34	39.0	7.5	18.0	53.0
Control	36	Theory T1	36	36.2	9.4	13.0	54.0
	36	Theory T3	36	35.6	7.9	16.0	47.0

Table 4.4.22 : Subjects with baseline and follow-up scores - Test one (Theory T1) and Test three (Theory T3) ('per protocol analysis')

Group	Obs		N	Mean score	Std dev	Min	Max
Intervention	52		52	38.8	7.8	18	53
Control	52		49	34.8	8.4	16	51

Table 4.4.23 : Subjects with follow-up scores - Test three

The subgroup with both baseline and follow-up scores (PPA) (Table 4.4.22) had baseline scores that were similar (i.e. within one score unit) to the group with valid baseline scores only (Table 4.4.21), and follow-up scores that were similar (i.e. within one score unit) to the group with valid follow-up scores only (Table 4.4.23). One can conclude that there was no selection bias in respect of the analysis of the theory scores, regardless of the type of analysis grouping.

4.4.2.3.2 Clinical scores

The maximum possible clinical score for tests one and three was the same, *viz.* 36.

Group	Obs		N	Mean score	Std dev	Min	Max
Intervention	58		57	13.9	4.8	8	35
Control	61		61	13.3	3.7	6	22

Table 4.4.24 : All subjects with baseline clinical scores - Test one (Baseline randomisation)

Group	Obs	Variable	N	Mean score	Std dev	Min	Max
Intervention	34	Clinical T1	33	12.7	3.4	8	19
	34	Clinical T3	34	14.8	4.3	5	22
Control	36	Clinical T1	36	12.4	3.6	6	19
	36	Clinical T3	36	12.2	4.3	4	20

Table 4.4.25 : Subjects with baseline and follow-up clinical scores - Test one (Clinical T1) and Test three (Clinical T3) ('per protocol' analysis)

Group	Obs		N	Mean score	Std dev	Min	Max
Intervention	58		58	15.0	4.6	5	23
Control	60		60	13.0	4.6	4	27

Table 4.4.26 : Subjects with follow-up clinical scores - Test three

The subgroup with both baseline and follow-up clinical scores (PPA) (Table 4.4.25) had a baseline clinical score that was similar (within one score unit) in the control arm but slightly lower in the intervention arm in comparison to the group with baseline clinical scores only (Table 4.4.24). The subgroup with both baseline and follow-up clinical scores (PPA) (Table 4.4.25) had follow-up clinical scores that were similar to the group with follow-up clinical scores only (Table 4.4.26). One can conclude that there was minimal selection bias.

4.4.2.4 Midwives' tests conclusion

The secondary outcome was the total score obtained in the midwives' tests. The GEE and mixed effects models were applied and found to be consistent so the GEE was used throughout. The intervention has shown a moderate effect which was statistically significant, and when compared to the control mean for test three the intervention arm had a mean increase of 17.1%. Irrespective of the subgroup analysis (intention to treat or per protocol) the size of the effect and significance was consistent. Although there were losses (and gains) to follow-up, the proportion of completions for each arm was virtually the same. The sensitivity analysis indicates there was minimal selection bias.

4.4.3 Changes within intervention arm over time (Objective 3)

The intention of this objective was to identify if there were any patterns in the midwives' knowledge and skills that might inform the timing of any education interventions, by testing at three timepoints. Further, this might have indicated some links to patterns in the partograph audit.

Although 79 midwives were enrolled in the intervention arm of the study, only 58 of these commenced the study at the time of the baseline test (test one). At the end of the 12 month period only 22 participants (8 clusters) completed the tests at the three timepoints. Given the low numbers of participants whose data were eligible for inclusion, a formal analysis was abandoned. This section therefore presents a broad perspective. Scores were translated into percentages of the maximum possible score in order to be able to make comparisons.

The baseline test was done in the month preceding the commencement of the partograph audit. Test two was done in month three and test three in month 12 - the last month of the partograph audit. Some portions of the test at timepoint two could not be included in the analysis because the theory test was administered at the incorrect time in one cluster. Twenty-one participants were included in the analysis of the total score and 19 for the clinical score. The total score at each timepoint followed by the changes for the sector scores (whose

compositions have been described in 4.4.2.2) are reported. Apart from the total score and its major components (theory and clinical), descriptive statistics will not be presented.

4.4.3.1 Total score

The descriptive statistics for the total scores at the three timepoints (Table 4.4.27) show that there was a reduction in the mean total score at timepoint two; at timepoint three the mean total score was higher than that of timepoint one; and the maximum scores increased at each timepoint.

Timepoint	N	Mean score	Percentage score	Std Dev	Minimum	Maximum
1, baseline	22	48.5	37.6	10.0	30	60
2, month 3	21	42.4	33.4	14.2	10	69
3, month 12	22	53.1	44.6	13.0	31	72

Table 4.4.27 Descriptive statistics for total score (reflected as a percentage) at three timepoints

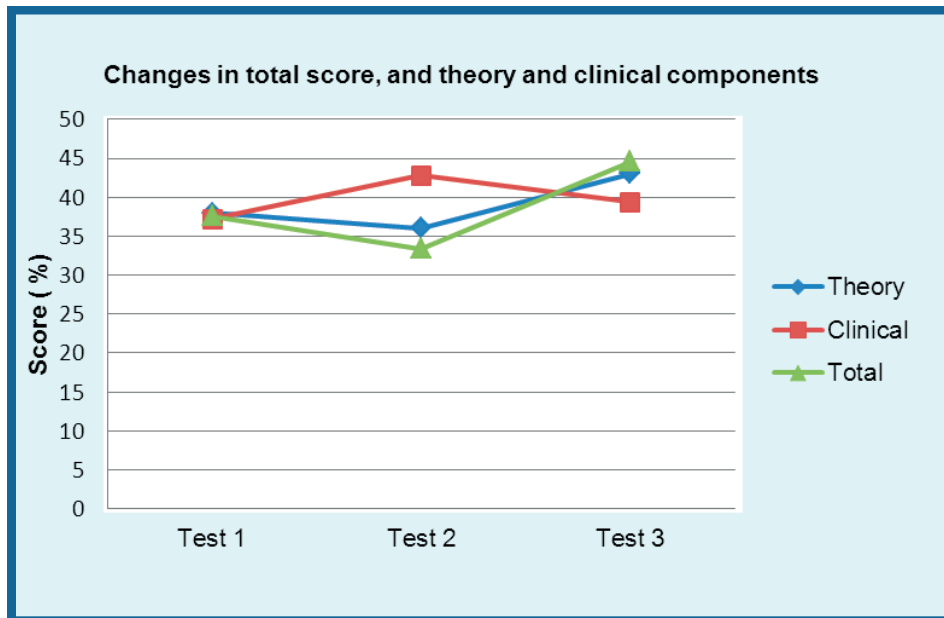
Table 4.4.28 illustrates that the theoretical component had a similar pattern to the overall score, except the maximum score which showed a slight increase in test two. The clinical component (which represents clinical assessment) shows a peak in mean score at month three. This degree of increase is not maintained but the score at month 12 is higher than that at baseline. The maximum score in the clinical component remained fairly constant.

Variable	Time-point	N	Mean score	Percentage score	Std Dev	Minimum	Maximum
Theory	1	22	35.1	37.8	7.2	21	47
	2	22	33.0	36.2	9.5	11	50
	3	22	35.6	42.9	7.4	18	48
Clinical	1	22	13.4	37.2	3.9	9	21
	2	19	15.4	42.8	4.6	7	23
	3	22	14.2	39.4	4.4	5	22

Table 4.4.28 Descriptive statistics for theory and clinical components at three timepoints

The changes in total, theory and clinical scores are represented graphically in Figure 4.4.6.

Figure 4.4.6 Changes in total percentage score, and theory and clinical components, at three timepoints



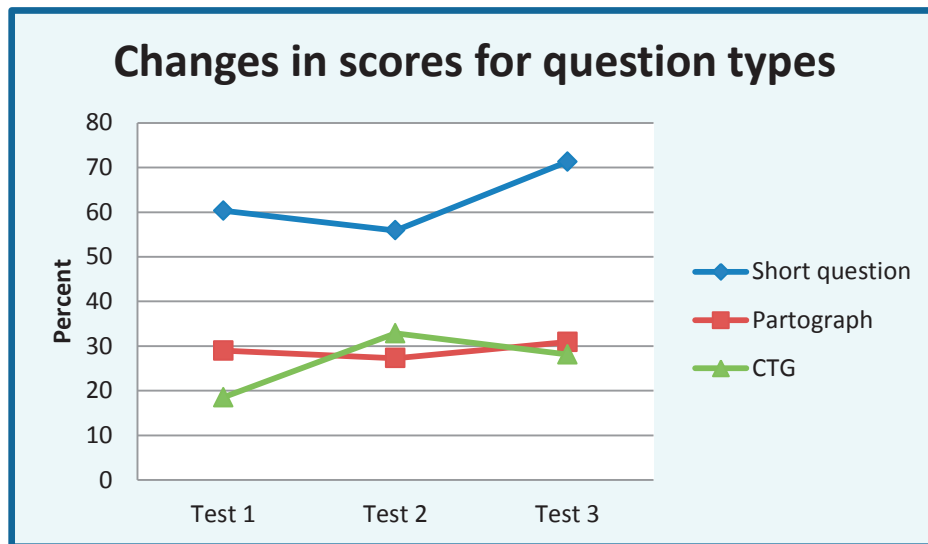
4.4.3.2 Sub-scores

As has been presented in the rest of this study, the sub-scores represented types of questions, cognitive levels and practice aspects.

Overall changes apparent at test two (Figures 4.4.7 – 4.4.9) were that three scores virtually maintained the same level as at baseline, *viz.* interpretation of the partograph, theory score, and critical management decisions. Three scores increased from the baseline level. These were interpretation of the cardiotocograph (very markedly), knowledge level questions, and fetal. The other six scores decreased – three markedly *viz.*, maternal, progress of labour and clinical management, and the other three more modestly, *viz.*, recognition of risk, application (higher order cognitive questions), short question score.

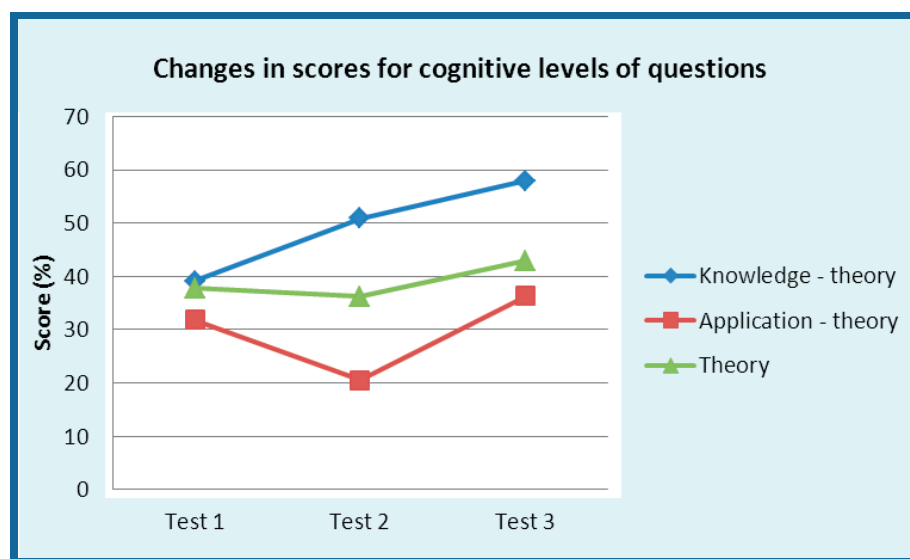
At test three, of the scores dealing with types of questions (Figure 4.4.7), the short question score improved after test two and achieved a higher level than the original baseline; the case based interpretation of a partograph maintained its level throughout; and the CTG score dropped off slightly but was still well above the baseline level.

Figure 4.4.7 Changes in percentage scores for question types



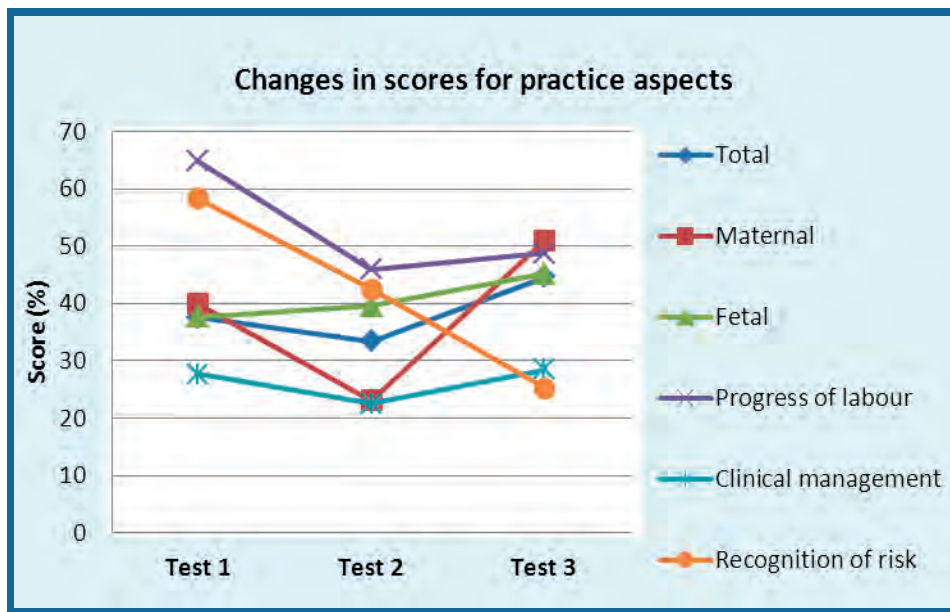
At test three, of the scores dealing with cognitive levels of questions (Figure 4.4.8) knowledge scores increased further while application scores recovered from the loss seen in test two and increased above the baseline level.

Figure 4.4.8 Changes in percentage scores for cognitive levels



At test three, of the scores dealing with practice aspects (Figure 4.4.9), two, *viz.* maternal and fetal scores, improved after test two and achieved a higher level than the original baseline; two, *viz.* progress of labour and recognition of risk, did not regain the original baseline level (progress of labour did improve but recognition of risk fell at both test occasions); two, *viz.* clinical management and critical management decisions, regained the original baseline level.

Figure 4.4.9 Changes in percentage scores for practice aspects



In summary, over the study period, improvements were shown in the total score, theoretical component, clinical component (mostly in the earlier period), short questions, interpretation of a cardiotocograph, cognitive levels (theoretical), knowledge and, to a lesser extent, application, maternal assessment, and fetal assessment. The following aspects were constant - partograph interpretation, clinical management and critical management decisions, while a decrease in scores was found in recognition of risk and progress of labour. However, given the very small sub-group of the intervention arm that this represents, these trends should be viewed with caution.

4.5 Comparison between structured versus unplanned *ad hoc* approach to staff training (Objective 4)

This study compared a focused and limited intrapartum training programme combined with clinical facilitation involving all relevant registered midwives in each intervention cluster, against the usual method of offering in-service training to midwives if it was possible for them to be released from the service for this purpose.

The primary outcome was the standard of practice using the **partograph utilisation audit** as a proxy measure to evaluate this. The findings were as follows :

- ◆ The overall partograph score did not show a difference that was statistically significant, but it was shown that the higher scores were significantly more likely to occur in the intervention arm (4.3.2). Most sector scores showed improvement but did not achieve statistical significance (4.3.3).
- ◆ Specific innovations linked to the provincial guidelines were 'alignment to clinical guidelines' which showed a significant difference in favour of the intervention arm (4.3.3.2), and fetal heart observations pre- AND post-contraction (4.3.4) which showed a higher adoption rate in the intervention sites. The Odds Ratio was 2.0 in favour of the intervention arm, but $p=0.237$, thus this difference was not statistically significant. The third innovation was the comprehensive 4 hourly assessment (4.3.4) which had an Odds Ratio of 12.6 indicating that the intervention arm was much more likely to have the comprehensive assessment completed (where indicated and where there was a template) than the control arm, and that this was statistically significant ($p<0.001$).

The **midwife tests** (secondary outcome) showed a statistically significant difference between the two arms, marked by deterioration in the control arm.

In the intervention arm :

- ◆ CTG interpretation showed a significant improvement which was generally maintained (4.4.2.2.1).
- ◆ Case-based interpretation of a partograph and clinical skills had significantly higher scores (4.4.2.2.1).
- ◆ Cognitive scores showed significant improvement (4.4.2.2.2).
- ◆ Aspects of practice that were significantly higher were maternal, fetal and progress of labour, where the midwife was able to build on familiar concepts, knowledge and skills (4.4.2.2.3).
- ◆ There was no significant difference for recognition of risk, clinical management and critical management decisions (4.4.2.2.3).

The intervention contributed to higher scores or improvements in knowledge and clinical skills. The control arm showed low scores which deteriorated over time, in response to the usual practice of opportunistic cover of in-service training.

4.6 Conclusion

This chapter has presented the findings of the primary and secondary outcomes specified for this study. Detailed analysis has demonstrated for the primary outcome that higher scores were more likely to be obtained in the intervention arm, as well as for some of the specific innovations included in the intervention package, viz. alignment with clinical guidelines and four hourly critical comprehensive assessment and management. In the secondary outcome, despite the staff dynamics (expected in a pragmatic study), there was consistency of results irrespective of the statistical tests applied (GEE and the mixed effects model), and between groups analysed (ITT and PPA).

Chapter 5 Discussion

5.1 Introduction

This chapter presents the following for the partograph utilisation audit and midwife test results :

- ◆ a synopsis of the key findings, reflecting these against the study hypotheses, sources of potential bias or imprecision, adverse event reporting and dangers associated with multiplicity of analysis and outcomes (5.2.1 and 5.3.1)
- ◆ consideration of possible mechanisms and explanations for these results (5.2.2 and 5.3.2)

Thereafter a comparison with relevant findings from current available literature is presented (5.4), followed by the strengths, weaknesses and limitations (5.5) for the full trial. The generalisability of the findings and a discussion of possible differences in other settings particularly in respect of health service organisation and levels of care are presented (5.5). Given the pragmatic nature of the study and the role of the researcher in the intervention and data collection process, a reflexive section (5.6) is presented, prior to the chapter's conclusion (5.7).

5.2 Partograph utilisation audit

5.2.1 Key findings

The utilisation of partographs (70%) was comparable in both arms (5.2.2.1 and 5.2.2.2). The mean scores for the total partograph did not show a statistically significant difference between the two arms (5.2.2.3). Bootstrap regression analysis showed parity in the lower levels of completion, but a statistically significant difference in those with better quality completions favouring the intervention arm. Of the specific innovations introduced fetal heart observations pre- AND post-contraction showed a higher adoption rate in the intervention arm than in the control arm. The Odds Ratio was 2.0 with $p=0.237$, thus not statistically significant. The Odds Ratio for the critical comprehensive assessment was 12.6 ($p<0.001$) favouring the intervention arm, for those

records where a partograph was used. Operational and staffing constraints hampered compliance with the specified frequency of observations (5.2.2.2).

5.2.1.1 Hypothesis testing

Hypothesis 1 : A well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators improves the standard of practice of registered midwives.

Null hypothesis 1 : A well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators does not improve the standard of practice of registered midwives.

The data presented indicates that in terms of the partograph score that there was no significant difference between the arms, thus supporting null hypothesis one. However, there was a significant difference between the two arms in the higher level scores, and the pattern exhibited in the time series gave evidence of a sustainability pattern, which was not apparent in the control arm.

5.2.1.2 Bias

The sample was obtained as planned. The profile of the records indicates that the arms were comparable and that further analysis could be undertaken with confidence.

5.2.1.3 Adverse events

The data used in this study were obtained from completed clinical records. Adverse events that might be related to this part of study would include coding errors or revealing of the identity of a person's record. This did not occur. The finding that 30% (306/1020) of records lacked any partograph reflected the practice current at the time of the data being collected and was not directly attributable to the study.

5.2.1.4 Dangers associated with multiplicity of analysis and outcomes

The main results related to the total score which was the single primary outcome. All other analyses reported in this section were secondary and explanatory with the purpose of being able to explain certain practice dynamics.

5.2.2 Possible mechanisms and explanations for these results

5.2.2.1 Non-utilisation of partographs

Nearly 30% of women who should have had partographs for recording their labours did not, meaning that the labouring women (and their fetuses) in these two regions did not experience the level of surveillance regarded as necessary for a safe and well managed labour.

A number of midwives seemed to lack confidence in commencing the use of the partograph for maintaining this record, preferring to write (sometimes extensive) notes in long hand. The 'problem orientated patient record' was used as a 'catch-all' by the nursing staff. Frequently it was not used as intended, and clinically significant aspects of the progress of labour were not captured. Some records revealed the use of a vital signs observation sheet or 'toxaemia' chart for labour, suggesting familiarity (and preference) with the use of a chart over that of the partograph, but this meant that the advantage of the alert and action lines for determining prolonged labours was not utilised. In some instances observations were recorded both on the partograph and in free text format⁵⁷.

This raised the question as to why some midwives do not utilise the partograph. This was explored briefly in the workshop format at the beginning of the two days of training and informed some of the educational activities. Responses relating to capacity (not enough time, shortage of staff), administration (difficulty in sourcing the document), lack of knowledge and clinical skill (don't understand it, never taught how to use it, uncertain about clinical findings) were given. This resulted in lack of confidence in using it or being unconfident to start a

⁵⁷ The only document that should be used in monitoring and documenting labour is the partograph. There is a free text record sheet which is used for any other observations antenatally, e.g. when a woman is admitted but not in labour, or which some midwives elect to use when they choose not to use the partograph (contrary to the policy and standard). It did not form part of the partograph utilisation evaluation except to determine patient records which were eligible for inclusion in the study but where the partograph was not used, despite the woman being in labour. There is also a 'toxaemia' chart which is a vital signs chart modified to record signs associated with gestational proteinuric hypertension. It is not the policy for this to be used to record observations in labour, but evidence of this practice was found. This did not form part of the partograph utilisation evaluation.

partograph but willing to continue it once it is started. Most of these highlight teaching and learning opportunities, and the need for clinical facilitation.

The lack of confidence / understanding of how to utilise the partograph was illustrated by the time being recorded incorrectly on the partograph thus rendering it ineffective. This could have medico-legal consequences where parturient women are referred too late, or an intervention is commenced inappropriately early, e.g. caesarean section for poor progress of labour, which was in fact normal progress. The incorrect plotting of observations in relation to time could also influence the frequency with which observations are performed, e.g. adjudged to be in latent phase of labour instead of active phase of labour and thus observations performed less frequently according to the guidelines.

5.2.2.2 Non-completion of partographs

While data was not available to explain all incomplete partographs, table 4.3.19 gives an indication of factors contributing to non-completion. A number of partographs had multiple reasons recorded. One can assume that the same reasons are relevant for both study arms.

5.2.2.2.1 Staffing levels for intrapartum care

The majority of the reasons were operational rather than patient-related in nature, suggesting that pressure on staff time contributed to the non-completion of the observations and recordings. It appeared that inadequate staffing to manage the workload in the labour wards contributed to non-completion of observations and recording on the partograph as set out in the guidelines for intrapartum care (Medical Research Council Unit for Maternal and Infant Health Care Strategies 2006).

There are no staffing norms set for South Africa for midwives per population or midwives per number of births per labour ward. The latter would not be applicable to the majority of the sites in the study as all but two of them are labour rooms accommodated within a ward which attends to varying scopes of illness requiring hospitalisation (described in 1.2.2). The smaller hospitals (240-500 births per year) would generally have one registered nurse-midwife on duty

per shift who would be attending to all nursing and midwifery responsibilities. The largest site in this study, which was a midwifery-specific site, has approximately 1200 births per annum, but also covers all antenatal care in addition to postnatal and neonatal follow-up. There would be one to two registered midwives on duty during the day and one at night at this site.

5.2.2.2.2 *Inappropriate deployment of staff*

What was noteworthy was how often midwives were working outside the labour ward at a time when they were officially allocated to the labour ward. The records showed that midwives were redeployed during their shift when women in labour required their care and sent to the outpatients department, casualty, theatre, 'another ward, and 'another department'.

5.2.2.2.3 *Failure to value the importance of observations and documenting thereof*

While midwives often plead that there is not enough time to record their findings on a partograph, time is wasted documenting observations on an inappropriate document in detailed long-hand. In a number of cases unnecessary and frequent observations were performed (including unwarranted vaginal examinations), yet labouring women who qualified for observations did not always receive them according to the guidelines. There was a sense that midwives were taken by surprise by births because of lack of monitoring. In many instances one or two observations were done and, after a gap of several hours, the birth occurred. It appeared that the CTG machines were used as 'babysitters'. Often the fetal heart rate was not even recorded but a note was written 'CTG' thus depriving a clinician or auditor of any useful information. This suggests an element of being seen to be doing what is required, without understanding the importance of one's observations and findings.

Often doctors did not make their notes on the partograph or admission notes, but recorded these separately. This duplication of information could have resulted in important information being overlooked, leading to an inappropriate management decision being made.

5.2.2.2.4 *Lack of second stage observations recorded*

Documenting of labour often seemed to cease once full cervical dilatation was reached. The partograph does not accommodate observations for the second stage of labour, thus any observations that are performed do not have a place to be recorded formally.

5.2.2.2.5 *Birth or transfer*

Of the 714 records where the partograph was utilised, 496 were incomplete. Of these, 289 labours resulted in birth or transfer within 1.5 hours of the last observation, which might explain the missing observations.

5.2.2.3 *Low scores*

There was a significant difference in favour of the intervention arm in respect of the better quality completions (at a score of ≥ 27), but this needs to be interpreted against the implications of these scores. The maximum possible score for the partograph was 47, so a score of 27 represents 57%. This means that virtually none of the control group reached what would be regarded as a 'pass mark', if this were set at the usual 50% level. Median percentage scores lower than 50% were found in both arms for 'alignment with clinical guidelines', 'frequency of observations', 'recognition of risk status' and 'comprehensive assessment'. This cannot be an acceptable standard of recording of observations on which clinical judgement and management decisions depend and which are necessary to provide safe and effective intrapartum care as set out in the guidelines. This illustrates the need to address most of the parameters in any future in-service training programme and practice improvement initiative.

5.2.2.3.1 *Clinical skills deficit*

There were numerous instances where clinical skills and interpretation of observations were lacking. Some clinical findings were not believable, e.g. where there was rapid cervical dilatation, where moulding and caput were recorded without a vaginal examination being recorded. In some instances there was a significant change in the quality of the observations and recording after the change of shift.

5.2.2.3.2 *Lack of understanding of labour / clinical relevance*

There seemed to be a lack of understanding of the labour process, e.g. recording the presenting part ascending as labour progressed. There was also a lack of understanding of clinical relevance of the observations, e.g. the frequent recording of 'toilet' in place of urinary observations, rather than appreciating if the bladder was empty, and what the urinary findings were. Frequently there were inexplicable clinical management decisions made in relation to the observations recorded. This may have been because the graph was plotted incorrectly or simply be contrary to the clinical guidelines for labour, e.g. continuing monitoring for a further four to eight hours in the presence of poor progress of labour, and cervical dilatation crossing the action line.

5.2.2.3.3 *Compliance with clinical guidelines*

It is not sufficient to make and record certain observations. The clinical guidelines indicate the frequency with which these should be done and how management should proceed based on the findings.

'Alignment with clinical guidelines' related to frequency of observations, and appropriate responses in terms of clinical management (including referral) and further observation intervals. This was promoted during the training programme and was in line with the clinical guidelines which had been distributed to the sites through the usual route, i.e. provincial circular. The intervention highlighted the requirements in the guidelines, indicated the relevance and importance of applying these guidelines, and gave midwives the skills to make the necessary observations. The results confirmed that active dissemination of information was superior to circular style communication which may not be applied at the clinical interface.

Common departures from the clinical guidelines were that women were not in labour (and should not have had a partograph commenced), there was a failure to monitor, non-adherence to guidelines (e.g. fetal heart observations, response to crossing of alert or action lines), multiple unjustified vaginal examinations, lack of recognition of risk, and continuation of observations onto a second partograph.

Differences in terms of observations and actions become apparent when the midwife was required to make a decision based on the clinical observations and then apply the appropriate clinical guidelines. During the teaching portion of the intervention it was clear that midwives were unable to articulate their clinical findings, and often were quite inadequate in undertaking these. This links with the findings in midwife tests where the improvement in the clinical score was significant (4.4.2.2.1).

Given that the scores in the control arm were so low, and even the statistically significant improvement in the intervention arm was low, one has to consider the clinical implications of this and the implications of human resource constraints which prevent the effective monitoring of women in labour and timeous responses to a changing clinical picture.

5.2.2.4 Specific innovations

There were two innovations introduced in the province during the study in respect of intrapartum monitoring, *viz.* comprehensive assessment and management plan, and fetal heart monitoring pre- AND post-contraction. These were communicated to these regions through the usual dissemination route, *viz.* provincial circular (document) and *ad hoc* in-service training, and incorporated into the in-service training for the intervention sites.

Formal assessment and documenting of the comprehensive assessment and management plan had not been a stipulated requirement prior to this study and thus could be regarded as an indication of the impact of the intervention. Among the 326 records eligible for this recording, the Odds Ratio of 12.6 ($p < 0.001$) in favour of the intervention arm indicated that there was a much higher uptake. The template for recording the critical comprehensive assessment and management plan was on the reverse of the partograph so was not immediately visible unless one specifically looked for it. Further, labour had to be assessed as a trend rather than a single timepoint. The nature of this score required clinical analysis to be done, which speaks to a more complex cognitive

behaviour, to be discussed in 5.3.2.2.2. The intervention arm showed a positive response and suggested that actively mediating an innovation is more likely to yield a positive response than the usual passive dissemination strategy.

By contrast the other innovation - monitoring of the fetal heart pre- AND post-contraction - had an Odds Ratio of 2.0, but with $p=0.237$ this was not statistically significant. One reason might be that there was some familiarity with fetal heart observations so the usual dissemination strategy might have resulted in a better uptake in the control arm when the policy was introduced. The uptake was 50% in the control arm as opposed to 64% in the intervention arm. This might be attributed to the trigger on the partograph where the legend indicated that there should be two fetal heart readings at each timepoint. Further, fetal heart recording is familiar and should pose no difficulty as no new skills are required.

5.2.2.5 Institutional profiles

The partograph scores for the moderate-to-large clinics scored 6.4 units more in the intervention arm than the control arm, this difference being statistically significant, $p=0.029$ (95% CI : 0.7 to 12.1). With the maximum possible partograph score being 47, this equated to a 13.6% difference. The implication for the larger sites is that there would be greater possibility for the midwives to implement what had been learnt. The clinical facilitation role was likely to be more effectively implemented in the larger services where one of the clinical facilitators was likely to be on duty with another registered midwife.

Smaller sites showed no difference between arms. Fewer births could expose a lack of familiarity with clinical management. Some of the smaller services have between 22 and 30 births per month, i.e. less than one per day on average. Spread over eight to twelve staff, an individual midwife might attend two to three births per month, interspersed with other general nursing duties and challenges, thus not having the possibility to put into practice what was learnt.

5.2.2.6 Clinical facilitation

Region one had the advantage of appointing the planned number of clinical facilitators at the beginning of the programme, and having an actively involved MCWH co-ordinator supporting them. Thus a system level strategy was in place. This contrasted with region two which did not achieve the full complement (see 5.4.2.2.1). The difference between the regions for the moderate-to-large sites in the intervention arm suggested that this difference in clinical facilitation strength might have contributed to the higher scores in region one. However, the study was not powered to detect this degree of difference.

The pattern of scores in relation to time show that the intervention arm scores started at a higher level (in the month following the commencement of the training component) and were maintained at a higher level throughout the data collection period, whereas in the control arm scores ultimately deteriorated. One reason for maintaining the higher level could have been the contribution of the clinical facilitators which was more comprehensive in region one.

5.2.3 Comparison with relevant findings from current available literature

The 'Saving Mothers' and 'Saving Babies' reports which relate to mortality data all refer to infrequent observations, lack of partograph utilisation, abnormal observations being ignored or referred after a delay (Pattinson 2003, Department of Health 2006a, Pattinson 2007, Department of Health 2009, Department of Health 2012). What this study shows is that the standard of partograph utilisation is poor even where there are non-morbid outcomes.

5.2.3.1 Utilisation of a partograph

Great variety in completion rates for partograph utilisation have been reported from various settings : 85% during a Safe Motherhood Demonstration project ⁵⁸ versus 11% prior to the project (Wamwana, Ndavi *et al.* 2007); 25% in Cambodia⁵⁹ (Ith, Dawson *et al.* 2012); only 12% of women having an

⁵⁸ A provincial general hospital and the surrounding four districts in one province, in Kenya

⁵⁹ Purposively selected public sector primary health centres, secondary referral and provincial hospitals in one province, covering a range of services offering basic and comprehensive emergency obstetric care

emergency caesarean section in Ethiopia were monitored utilising a partograph (Fesseha, Getachew *et al.* 2011); 58% were satisfactorily completed in southern Tanzania⁶⁰ (Bosse, Massawe *et al.* 2002); and 70% use in a sub-district of Uganda was recorded (Ogwang, Karyabakabo *et al.* 2009). Rotich, Maina *et al.* (2011), who directly observed the clinical observations being done⁶¹, found that not one of the 234 partographs utilised was recorded completely correctly. This current study, with a 70% utilisation rate, compares reasonably favourably to the literature.

There might be variation between the components of the partograph. Osungbade, Oginni *et al.* (2010) reported optimal care in respect of vaginal examination, fetal heart monitoring and blood pressure measurement for 243/338 (71.9%), 73/338 (21.6%) and 52/338 (15.4%) parturients respectively which diminished as labour progressed. Rotich, Maina *et al.* (2011) reported similar issues to this study regarding poor recording of certain parameters, e.g. descent of the head, urine observations. Other observations by Rotich, Maina *et al.* (2011) obtained lower percentage scores than this study, e.g. degree of moulding, half-hourly observations of the fetal heart rate plotted, blood pressure, temperature.

Reasons for non-recording of observations reported in the literature included certain observations leading to immediate transfer and the observations not being recorded; during second stage observations were often done but if there was no change in status or cause for concern these observations were not recorded in the notes (Pitchforth, Lilford *et al.* 2010). It is possible that these might be relevant to this study as well. In this study where records were incomplete, there was minimal recording of observations in the second stage and 289/946 (30.6%) births occurred within 1.5 hours of the last observation.

⁶⁰ Two government hospitals and one mission hospital

⁶¹ Two national referral hospitals in Kenya

While some report difficulty in the use of the partograph⁶² (Pettersson, Svensson *et al.* 2000) or lack of guidelines and teaching in its use (Ogwang, Karyabakabo *et al.* 2009), others report on its benefits and ease of use (Kwast, Poovan *et al.* 2008, Lavender, Hart *et al.* 2008, Rotich personal communication 2012).

When comparing the user-friendliness of a simplified partograph to the 'composite'⁶³ partograph, the simplified partograph gained higher scores (Mathews, Rajaratnam *et al.* 2007). However, that study had several limitations: it was not a RCT, it was conducted with physicians in a teaching hospital and not midwives in a primary or secondary level of service, and it lost the strength of a graphical 'snapshot' by recording columns of figures.

5.2.3.2 Policy into practice needs to be actively mediated

The need for active rather than passive dissemination of information for clinical practice change is recognised (Oxman, Thomson *et al.* 1995, Grimshaw, Thomas *et al.* 2004, Althabe, Buekens *et al.* 2008). Introduction of new or revised guidelines needs to be accompanied by an active implementation strategy involving all relevant participants (Grimshaw, Thomas *et al.* 2004, Menon, Korner-Bitensky *et al.* 2009) as well as ongoing orientation of new staff (Rotich personal communication 2012) and monitoring of the application of clinical guidelines (Sprague, Oppenheimer *et al.* 2008). This can be contrasted with the introduction of Reproductive Health Library in Mexico and Thailand without active mediation, which showed no change in the birth attendants' practices leading to Gülmezoglu, Langer *et al.* (2007) recognising that access to knowledge alone is insufficient to change health providers' behaviour.

⁶² This study refers to the 'Angolan model of the WHO partograph', but is not described. It assumed that this resembled the standard WHO partograph as it included the latent phase.

⁶³ Other authors refer to this as the standard WHO partograph, which includes the latent phase. The modified and simplified versions of the partograph both exclude the latent phase.

5.2.3.3 Use of multifaceted intervention

Educational material on its own shows modest improvement, but when used as part of a multifaceted strategy which includes an interactive process, e.g. educational outreach, supported by ongoing audit and feedback, its effectiveness is potentiated. While it appears that the use of more than one intervention usually is more effective than a single intervention, it was not clear what the ideal number of strategies should be (O'Brien, Rogers *et al.* 2007, Menon, Korner-Bitensky *et al.* 2009, Ivers, Jamtvedt *et al.* 2012) or what combinations are optimal (Grimshaw, Thomas *et al.* 2004). A multi-faceted intervention using the conceptual framework of diffusion theory was found to be effective in changing behaviour (Rollnick, Mason *et al.* 1999, Althabe, Buekens *et al.* 2008). This study was guided by the health behaviour change model of Rollnick, Mason *et al.* (1999) which was based on the 'Stages of Change' trans-theoretical model developed by Prochaska and DiClemente (1982).

5.2.3.4 Under-resourced settings

The introduction of a perinatal mortality audit process in low and middle income countries demonstrated a reduction in perinatal deaths (Pattinson, Kerber *et al.* 2009). Audit of records and the quality of decisions made should be an action-orientated activity to improve the quality of care, and participation by all professional staff involved in intrapartum care should be a requirement (Department of Health 2006a). Souza suggests that in under-resourced settings a local opinion leader may be the single available resource for his / her peers and could act as an agent of change, and that continued support through academic detailing (educational outreach) for opinion leaders could be used to keep them committed to changing the behaviour of their peers (Souza 2007a, Souza 2007b).

In this study, conducted in an under-resourced setting, informal audit and ongoing feedback and support built on input by an educationally influential individual (outreach activity) and educational materials, which were technologically appropriate for the setting.

5.2.3.5 ICC values

Generally in public health research the ICC is set at 0.05 - 0.1 (Adams, Gulliford *et al.* 2004). In this study the calculations of the sample size were performed on the basis of an assumed ICC of 0.05 for the partograph utilisation audit. The ICC measures obtained using the partograph score data (post intervention; pooled over control and intervention arms) actually yielded an ICC of 0.28. The ICC was reduced to 0.16 when adjusted for interactions. However, this is still higher than the values normally associated with this type of study, although it does fall within the range for ICCs for process indicators described previously, i.e. 0 to 0.415 (Campbell, Fayers *et al.* 2005). It seems that the context and type of intervention needs more careful consideration when determining ICCs.

5.2.3.6 Professional / workplace culture

5.2.3.6.1 Professional differences

As discussed in the literature review, systematic reviews looking at change in professional practice source studies generally target physicians (Grimshaw, Thomas *et al.* 2004, Cheater, Baker *et al.* 2005, Doumit, Gattellari *et al.* 2007, Jamtvedt, Young *et al.* 2007, O'Brien, Rogers *et al.* 2007, Forsetlund, Bjorndal *et al.* 2009). Many of these reviews report prescribing behaviour.

The settings in which Nursing care is offered, are often in institutions with highly developed hierarchical structures where individual professional nurses may have a weak sense of autonomy or agency, and where there may or may not be a recognised role of clinical leadership / clinical governance for nurses. Where Midwifery services are offered within such health systems, there may be a similar sense of disempowerment. One cannot extrapolate findings from predominately medically oriented systematic reviews without understanding the different focus, history, culture, educational preparation, status, and social dimensions of the different professions, (Kirshbaum 2008) and the context in which the professional practice takes place.

Given that there are differences in professional practice and culture, and unequal power relations, a tailored approach to interventions has merit. However, any such studies would need to clearly specify what behaviour

change was targetted, how the barriers or supports for change were identified, if this was a comprehensive set of valid issues, and which of these were being targeted in the study (Cheater, Baker *et al.* 2005, Michie, van Stralen *et al.* 2011). These aspects were articulated in this study.

5.2.3.6.2 *Staff utilisation and support*

The availability of certain resources including staffing was raised by Pitchforth, Lilford *et al.* (2010) in the Ethiopian study previously described, who identified the lack of senior members of staff, and ineffective and inefficient employment of staff as problems, rather than an absolute shortage of staff *per se*. Similar conclusions were expressed by nurse managers in a South African study, although they did concede that staff had often been prevented from getting further training due to staffing limitations. The State of the World's Midwifery report 2014 cautions against considering only a headcount of appropriate personnel, and that the Midwifery workforce should be assessed on the basis of full-time-equivalent availability (United Nations Population Fund 2014). Penn-Kekana, McPake *et al.* (2007) reported that 84% of nurse respondents felt there were not enough staff to do the work. These findings resonate with the reporting by participants in this study.

In the South African context the appropriate use of staff (instead of frequent rotation of skilled midwives out of maternity) has been raised in the 'Saving Babies' reports (Pattinson 2007) and 73.4% of hospitals were reported to regularly rotate nurses through maternity services (Penn-Kekana, McPake *et al.* 2007). This management practice has important implications for many initiatives that attempt to develop staff working in maternity departments, as it undermines their ability to consolidate their learning and clinical skills.

Some human resource policies have militated against midwives remaining in Midwifery and midwives have transferred in order to obtain better salary packages which have not been extended to Midwifery. Penn-Kekana, McPake *et al.* (2007) reported that more than 30% of facilities had advanced midwives who were not working in the maternity section. Many of those midwives were

working in management positions as these offered promotion opportunities (Penn-Kekana, McPake *et al.* 2007). In addition, 40% of nurses working in the maternity section felt stressed, demotivated and burnt out, and were thinking about transferring to another ward. Their sense of feeling unsupported by management meant that the environment was not conducive to implementing policy interventions [clinical practice improvement] (Penn-Kekana, McPake *et al.* 2007). Such dynamics confirm the need for institutional and system level support for practice change initiatives, as was done in this study.

5.2.3.6.3 Practice settings

A restricted CRT was conducted in a portion of one of the regions where the current study was undertaken (Groenewald-Neethling 2010). This compared the use of face-to-face with training-of-trainer methods for introducing the WHO's Basic Antenatal Care package (BANC). This involved all the clinic staff participating in a two- to three-hour session weekly for five to six weeks. The outcome measure was an audit of antenatal cards according to a 24 point checklist which specified certain actions for each clinic visit. This showed remarkable changes for both training methods (mean difference overall of 2.9/24 (12%), $p < 0.001$) and, more particularly, the mean difference between the methods post-intervention was 3.32/24 ($p < 0.001$) in favour of the F2F method.

One might have expected the intervention effect reported to be similar to the current study. However, there are some key differences, i.e. the BANC package was new to all the participants; many of the participants had not been working with antenatal clients and so their knowledge and skills needed refreshing; and the BANC programme had a reasonably high profile in the health department, with one of the BANC targets being a Department of Health performance target / process indicator. (Services were diverted to other clinics during the training sessions in order to facilitate 100% attendance.) Thus there was a high degree of expectation and support for the initiative.

By contrast, in the intrapartum setting, there are midwives who are uncertain of, or resistant to, utilising the partograph for various reasons. The practitioner-

patient interaction is vastly different between the clinic-based one-to-one consultation in a limited time period, and the one-to-many responsibility in the ward setting (not limited to maternity) where 'triage care' is often practised in order to ensure that those most acutely in need of attention receive it. In the clinic setting the consultation occurs, all the observations and appropriate care are given and documented, and the interaction is completed, after which the practitioner attends to the next person. It is clear from the information captured in 4.3.5.1 (Table 4.3.19) that there are many obstacles to fulfilling all the obligations for observation and care of all patients for whom the midwife is responsible, including those in labour. It is in this setting that the partograph is meant to be completed by doing half-hourly observations lasting a minimum of ten minutes each.

It is of interest that in Theron's study of the use of the PEP maternal care manual, the improvement within the intervention arm was greater for use of the antenatal card (33%, $p < 0.001$) than for the partograph (17.5%, $p = 0.001$) (Theron 1999b). This might point to the different context or the greater complexity of clinical judgment that is required during the intrapartum period.

5.3 Midwife tests

5.3.1 Key findings

Regardless of the type of analysis carried out (ITT or PPA), the scores in the intervention arm were seven to eight points higher, and this was statistically significant. This change was due mainly to deterioration in the control arm scores.

When types of questions were analysed (ITT and PPA), significant differences were found between the two arms in favour of the intervention arm for short questions, interpretation of the CTG and clinical skills. The difference in the score for short questions was not significant in the subgroup of midwives who completed only the theory portion. However, in the PPA group this aspect also achieved a difference between the two arms which was significant and in favour of the intervention arm. ITT and PPA analysis of cognitive levels of questions

exhibited a statistically significant difference in favour of the intervention arm for knowledge and application scores. Three of the six practice aspects showed a significant difference between the two arms – maternal, fetal, and progress of labour. What was of concern was that those aspects which required clinical judgment and management decisions to be taken had very low scores with no significant difference between the two arms.

5.3.1.1 Hypothesis testing

Hypothesis 2 : A well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators improves the knowledge and skills of registered midwives.

Null hypothesis 2 : A well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators does not improve the knowledge and skills of registered midwives.

The data presented shows that there were higher scores for the intervention arm, and thus supports hypothesis 2. It is acknowledged that the higher scores in the intervention arm reflects the deterioration in the control arm more than an improvement in the intervention arm.

5.3.1.2 Bias

The randomisation process yielded a profile of midwives in respect of their personal attributes and their baseline test scores which was comparable between the two arms. Despite the changes in the participants, the overall profile showed no bias. The reasons for joining or departing reflected the expected dynamics of staff turnover and were not related to the study *per se*.

There were six withdrawals from the trial, where participants did not want to undertake further tests. All agreed that completed tests could be retained in the study and analysed.

- ◆ There were two withdrawals from the trial after test one (one in each arm of the study).

- ◆ Three individuals (one from the intervention arm and two from the control arm) completed the clinical test at month 12 but refused to do the theory test.
- ◆ One individual (in the control arm) wrote the theory test at month 12 but refused to do the clinical test.

One participant (intervention arm) who joined the study after the training days, completed neither theory nor clinical components at month 12. Her duties prevented her from being able to undertake these tests at the times available. However, she did not withdraw from the study, but was not included in the analysis.

There were a number of incomplete tests where either the theoretical or clinical portion was not done. The reasons were :

- ◆ midwives being on leave at the time the theory test was administered or only joining the service after the theory test had been administered, but being present when the clinical testing was done.
- ◆ due to the limited capacity of the single researcher / evaluator to conduct clinical evaluations on multiple occasions at multiple sites across a large geographic area, it was not always possible to match the researcher's visits with the availability of participants. Each site was visited on a number of occasions to cover day and night shifts as well as different shifts during the week.

The pattern of incomplete components was similar for the two arms, as were the components not undertaken, i.e. where the individuals were no longer part of the service and thus no longer in the study. The sensitivity analysis (4.4.2.3) compared the theory and clinical scores obtained for the ITT and PPA groups. The conclusion was that there was minimal bias.

Contamination cannot be addressed through randomisation. The pragmatic nature of this trial recognised the real-world setting of this study. What was different in the intervention arm was the added both activity, i.e. focused

in-service training and clinical facilitation. It was anticipated that some midwives might move out of their particular site but remain in the district in another part of the health service that had no link to the intrapartum service under study, and this occurred. None moved from an intervention site to a control site, or *vice versa*. The evaluation personnel were limited to the single researcher. The procedure outlined in 3.7.7 regarding data management was followed.

5.3.1.3 Adverse events

Although this was an intervention study it was not a treatment study. One would therefore not anticipate there being any adverse events related to the intervention *per se*. The only potential risks were those identified in 3.8.7 regarding exposure of an individual midwife's identity (this was not necessary) or harmful practices (3.8.8.2). In a couple of instances, where there were significant departures from the guidelines, these were brought to the attention of the regional clinician who was able to deal with these through the regular site meetings which were linked to mortality and morbidity meetings. This occurred after the evaluation of the tests was performed, which took place after the completion of all the tests, so this did not introduce bias in certain sites. There were no untoward events amongst the staff that related to this study. There were no major changes in management during the study period, although there may have been some anticipation of restructuring to form the new districts which were implemented a year after the data collection period.

5.3.1.4 Dangers associated with multiplicity of analysis and outcomes

The main results related to the total score which was the single secondary outcome. All other analyses for this section were subsidiary with the purpose of being able to explain certain learning dynamics. The analyses were not 'stacked' so that more select groups could be analysed.

5.3.2 Possible mechanisms and explanations

This section highlights noteworthy findings/trends and to interpret the findings by providing possible mechanisms or explanations for the results obtained.

5.3.2.1 Pattern of change in scores

The two tests used to compare the total score (i.e. GEE and mixed effects) indicated a similar intervention effect and both were statistically significant ($\Pr>|z|=0.006$ and $\Pr>F=0.016$ respectively). The difference between arms was seven points or 6% of the total score (7/119). While this was statistically significant in terms of relative change, the effect was moderate. Insofar as the knowledge and skills of the midwives are concerned, one would like to see an improvement in the scores of the intervention group. However, the difference between the scores was due to the small improvement and maintenance of scores in the intervention group, coupled with a reduction in scores in the control group. This illustrates the rapid deterioration of knowledge and skills that can take place if there is no sustained, focused continuing professional development programme established. It is suggested that the health service change in the intervention arm (i.e. the introduction of clinical facilitation) might have contributed to ensuring that the standard of care did not decrease.

5.3.2.2 Low scores

The mean scores were low in both arms, with only the intervention arm at test three achieving a mean percentage score of 50% ('pass mark') (4.4.2.1.1). This is a sobering finding where midwives appear to be functioning with such low levels of basic information.

During the marking of the tests the following observations were made regarding the midwives' skills in undertaking tests :

- ♦ Test preparedness : The method of evaluation of the midwives' knowledge and skills was a formal test. For many of the midwives, it would have been some years since they functioned under such conditions. While the intention was to test the working knowledge and skills of the midwives as they go about their daily work, rather than to reflect a contrived ('swotted') level, the fact that some of them were using an unfamiliar method of communication could have caused them to score at a lower level than their actual knowledge and ability.
- ♦ Expressing oneself in writing : The midwives could choose to do the test in either English or Afrikaans. However, some of the midwives

indicated that they were not used to having to express themselves in writing.

- ◆ Midwives appeared to be more comfortable to report on findings rather than making a conclusion as to the implications of their findings.

Therefore it is possible that the scores were a conservative representation of the midwives' knowledge and skills. This would have been similar for both arms.

5.3.2.2.1 *Cardiotocograph skills*

Cardiotocograph machines were introduced to all primary sites offering intrapartum care in the year preceding this study. A number of midwives had some knowledge of how to attach it to get a reading, but were unfamiliar with the interpretation of the graphical record. The positive response (1.1/10 (11%), $p < 0.0001$) should be considered in relation to a very poor (almost absent) level of knowledge of this aspect.

5.3.2.2.2 *Cognitive levels*

It is not sufficient to address the 'what' (knowledge). One needs to develop skills in the 'why' and 'what if' (analysis and evaluation) which are more complex cognitive skills. The difference relative to the total score was greater for knowledge than application, but both were statistically significant. A drop in the knowledge score in the control arm contributed to this pattern. The knowledge score encompassed the less complex levels as defined in the Bloom's taxonomy and relates to information acquisition. However, the application score remained constant in the control arm and increased in the intervention arm. The more complex cognitive skills which are required for application, analysis, synthesis and evaluation and which are necessary for clinical judgment are not so easy to acquire without a planned educational programme or necessary guidance, and may have been facilitated by the clinical facilitation. This has implications for change in practice and is perhaps a key to unlocking the low levels of change in aspects of practice seen in the partograph utilisation audit results, specifically the sector scores (4.3.3).

5.3.2.2.3 *Alignment with clinical guidelines*

Both arms had woefully low scores for this variable, but the intervention arm showed a statistically significant increase in aligning with the clinical guidelines

suggesting greater familiarity with, or commitment to, the guidelines. As this is an essential aspect of setting and attaining an acceptable standard of care, this finding is both alarming for the baseline and control arm scores, and mildly encouraging by the improvement in the intervention arm.

5.3.2.3 Tracking test performance of a subgroup of midwives at different timepoints

In the intervention arm an additional test was done at month three to ascertain if there was any pattern in relation to the timing of the training update. Two short term improvements were determining abdominal and vaginal findings (clinical score at test two), and knowledge and short question scores. Due to the low numbers of participants who completed the full test at all three timepoints formal analysis was not done, and the descriptive data should be treated with caution. Overall there is significant scope for improvement.

5.3.2.4 Human resource management

What was noteworthy was the number of midwives who had been qualified for many years but with little experience in Midwifery, who continue to be licensed and employed to practise as a midwife⁶⁴. There are no designated registered midwife posts. Registered midwives are appointed into professional nurse posts, many of which were vacant and unfilled. The general policy was to rotate staff through all departments of the service including the labour ward (sometimes even as the clinical supervisor due to seniority, when rotated into this department). A number of registered nurses never work as a midwife.

Barriers that prevent completion of partographs were highlighted in the feedback from midwives as well as in the comments captured during the training workshops. At the individual level, lack of confidence and skills were identified. At the health service level there may well be human resource constraints or unhelpful human resource utilisation such as rotation or redeployment, or system / process issues such as lack of clinical leadership and support, and lack of quality assurance monitoring.

⁶⁴ At the time of the study there were no mandated requirements to demonstrate current competency as a midwife after initial registration. Currently there are plans for a continuing professional development system to be implemented in 2015-16 but this is awaiting regulations to support the legislation.

The approach to in-service training appears to be *ad hoc* rather than linked to goals for health outcomes. There are no formalised links between primary level services and academic units in these regions. The lack of contact and exposure to innovations in Midwifery and related disciplines are likely to continue to produce 'more of the same' which is inadequate for the demands of the health service. Further, the lack of designated clinical leadership is an obstacle to midwives gaining confidence in their practise and clinical decision making.

5.3.3 Comparison with relevant findings from current available evidence

5.3.3.1 Location of intervention

Access to training is often raised as a concern especially if people need to leave their normal place of residence / work for any period of time.

Participants in the Perinatal Education Programme (PEP) do not need to leave their place of employment while studying the programme (Theron 1999a). In a 70 item multiple choice test, the cognitive score improved within the intervention arm by 22 marks (32%) ($p < 0.0001$), whereas within the control arm the mean improvement was 1.8 marks ($p = 0.185$) (Theron 1999a). By contrast, a comparison was made between hospitals which had on-site facilitation and those which had off-site facilitation for the implementation of Kangaroo Mother care. The researchers concluded that there was no significant difference in effectiveness between the two sites of facilitation, although they conceded that more hospitals would need to be covered before conclusive evidence could be established (Davis, Bergh *et al.* 2006). Longer-term evaluations need to be performed to establish the sustainability of the learning. Further, the PEP intervention cannot be directly compared to other programmes where the intervention includes not only the location but the inclusion of experienced trainers / clinicians and other contextual factors.

In this study the intervention training was located within the region in which people worked and was within normal work-related travel distances. Further, access was facilitated with transport being supplied by the employer.

5.3.3.2 Role of expert trainer

In the PEP programme the responsibility for learning is placed on the student and a formal tutor is not required. The principle is of co-operative learning, with groups being managed by a co-ordinator, often a peer (Theron 1999a). None of the co-ordinators had any training preparation. It was not clear how many co-ordinators there were and what influence they had on the process. In the BANC rollout study which compared the use of differently prepared trainers (an expert trainer delivering face-to-face training versus a non-expert trainer prepared using the training-of-trainer approach), while both arms showed significant improvements, there was a mean difference of 3.32 ($p < 0.0001$) in favour of the face-to-face arm (Groenewald-Neethling 2010). Both studies were conducted in similar settings in South Africa.

In this study the educational intervention was delivered by an experienced educator / clinician using the face-to-face approach. Similar degrees of change in midwives' knowledge and skills were obtained.

5.3.3.3 Accessing and applying knowledge

The PEP study also tested application of knowledge from clinical data (Theron 1999b). Midwives working in the antenatal clinic were tested on the antenatal card and those working in the labour ward were tested on the partograph. The same documents were used for both tests which could have resulted in some recognition from the participants. While improvements were found after the completion of the maternal care manual, this was greater for the antenatal card (Theron 1999b). The results for the partogram showed an improvement within the intervention arm of 3.5 score units (17.5%) which was statistically significant ($p = 0.001$) and no change within the control arm.

There was a difference in the baselines between the two arms in respect of the partograph score. This was explained by the fact that the midwife in charge of the labour ward in the study site had realised the value of the correct use of the partograph before the study commenced (Theron 1999b). This suggests a contextual factor related to the intervention site in addition to the location and facilitation of the training.

The ability to apply knowledge for clinical decision making requires current and correct knowledge as well as the confidence to use this. Midwives in rural settings need to be able to access evidence that will inform their practice. This should be available through evidence-based clinical guidelines (of which national and provincial guidelines have been developed for a variety of situations), via the PEP (which is aligned to these guidelines and other clinical policies) (Woods and Theron 1994, Theron 1999) or through the use of the Reproductive Health Library (which is available on compact disc (CD) to practitioners in developing country settings). However, as has been made abundantly clear by numerous authors (Rollnick, Mason *et al.* 1999, Grimshaw, Thomas *et al.* 2004, Gülmezoglu, Villar *et al.* 2004), the availability of information will not lead to change in behaviour without this being actively mediated. It will therefore be necessary to employ an appropriate implementation package consisting of elements presented in this study.

New knowledge needs to be accessed and synthesised which requires specific skills including interpretation of research reports. In order to effectively utilise research findings one should have the authority and personal agency to change practice. Kirshbaum (2008) notes that these barriers, along with insufficient time to read research, appear to be widespread and ingrained in nursing culture and practice and require multiple, systematic, and innovative strategies to eliminate obstacles and enhance facilitating aspects. This could be applied to midwives in the South African context where similar barriers are embedded in Nursing structures and education, and stifle innovation. In addition, there is a need for consistent clinical leadership. Where there is uncertainty regarding clinical management, practitioners have been found to be ineffective (Scott, VandenBeld *et al.* 2011).

5.3.3.4 The extent of change

The midwife test finding (GEE model 6% higher total score in the intervention arm, $p=0.006$) is consistent with those in systematic reviews for changes in professional performance due to continuing education meetings alone (median adjusted RD 6%) (Forsetlund, Bjorndal *et al.* 2009) and educational outreach

visits (median adjusted RD 6%) (O'Brien, Rogers *et al.* 2007). Educational outreach visits (EOV) found an overall median adjusted risk difference of 5.6% when measuring change in professional practice (O'Brien, Rogers *et al.* 2007), and, when directly compared against audit and feedback, EOVs appeared to be slightly superior. Grimshaw, Thomas *et al.* (2004) concluded that there was an imperfect evidence base to support decisions about which strategies are likely to be effective under various conditions. However, mean absolute improvements for dissemination of educational materials (8.1%), audit and feedback (7%) and multi-faceted including educational outreach (6%) showed modest improvement (Grimshaw, Thomas *et al.* 2004). There is a consistency about the extent of change among existing literature, and this study achieved a similar degree of difference.

5.3.3.5 Number of interactions

It is not clear how many visits or what duration of outreach is most effective for practice change. Educational outreach or academic detailing suggests that there should be at least two to three follow-up visits. Other authors have indicated that multiple interactions were more effective than single visits (Soumerai and Avorn 1990, O'Brien, Rogers *et al.* 2007, Ivers, Jamtvedt *et al.* 2012). Inferences could not be drawn from the within-arm analysis of this study but it was clear that midwives in the intervention arms increased their knowledge and skills and maintained these over a 12-month period whereas those in the control arm deteriorated. Although it is thought that more visits may be costly it is possible that the benefits may outweigh the costs. O'Brien, Rogers *et al.* (2007) suggest that economic evaluation of different strategies is required.

5.3.3.6 Barriers and supports for professional change

There are various models in the literature relating to barriers to change which could apply to partograph utilisation or midwife capacity⁶⁵. The findings of this study are applied to the categorisation developed by various authors.

⁶⁵ Given that this study involved an intervention package, and that outcomes related both to standard of care, and knowledge and skills of the midwives, some of the literature applies to both primary and secondary outcomes. Rather than repeating information, this has been accommodated where there was the best fit, with the recognition that there was some cross-over for certain aspects.

This study reflected six of the seven categories of barriers to optimal practice identified in a systematic review by Cochrane, Olson *et al.* (2007). The findings from this study are applied to the relevant category. Cognitive-behavioural barriers were identified in the lack of knowledge, awareness, lack of insight into partograph utilisation, lack of capacity in more complex cognitive processing and lack of clinical skill. Attitudinal or rational-emotional barriers were suggested by the lack of efficacy, lack of confidence, lack of sense of authority.

Professional barriers were evidenced in the mismatch between the number of years qualified and years of experience indicated, the lack of opportunity or progress, and the lack of appropriate peer influences or models. Guidelines were available but not necessarily distributed, known and communicated. Human and material resources were lacking, rotation and redeployment of staff, and lack of time to meet clinical care requirements were evident. The lack of system support, e.g. clinical facilitation and leadership, lack of incentives or sanctions to promote better practice and clinical uncertainty were apparent in the control arm. Patient barriers were not evident in this study.

Bingham and Main (2010) advanced another categorisation of potential barriers focusing on project implementation. In this study leader barriers were addressed by involving the leaders in the preparation phase to harness their input and support, and in the clinical facilitation programme to ensure they acquired the requisite knowledge and skills. Clinician barriers were addressed by using the health behaviour change model and adopting a participatory process in the training programmes. Clinical facilitation training included developing a positive and supportive environment, where audit and feedback were anticipated and teamwork was emphasised, thus addressing the implementation climate barrier. Characteristics of the project as barriers were not specifically addressed but involvement of local leaders and regional programme managers facilitated the process.

Depending on the barrier, an implementation plan can be planned systematically by using three broad types of interrelated strategies, i.e. discourse (communication), education (formal and informal) or data (audit)

(Bingham and Main 2010) or by using the Behaviour Change Wheel to determine intervention functions or areas for policy change (Michie, van Stralen *et al.* 2011). In their systematic review, Baker, Camosso-Stepinovic *et al.* (2010) concluded that interventions tailored to prospectively identified barriers are more likely to improve professional practice than no intervention or dissemination of guideline. This study's intervention package was tailored to local context, content was informed by mortality audit data, and the educational process was guided by a model of behaviour change.

5.3.3.7 ICC values

In this study the calculations of the sample size assumed an ICC of 0.1 for the midwives tests whereas the midwives' total test score yielded an ICC of 0.2, indicating a higher degree of variation between clusters than anticipated. When the focus of the intervention is on professional behaviour, with repeated measurements of individuals within a group, rather than on a diagnostic test or illness profile with harder measures, one might need to consider different values for the ICC. In the study by Fairall, Zwarenstein *et al.* (2005) the ICCs were low for almost all parameters except for those which reflected patient satisfaction with clinic health services, e.g. general satisfaction, professional care, depth of relationship and perceived time, where they ranged from 0.16–0.36. The ICC measured in this study, which reflects professional behaviour, also falls within this range.

5.4 Strengths and weaknesses (including limitations)

This section presents the strengths, weaknesses and limitations of this study, and the methods used to minimise or compensate for any limitations.

5.4.1 Strengths

5.4.1.1 Application of this research design to this issue

This study is unique in applying a cluster design to an intervention addressing professional behaviour and clinical management change in an intrapartum setting. The appropriate use of the CRT design in this setting is likely to give a more accurate outcome than the use of designs which do not take account of

clustering. Milne, Scotland *et al.* (2004) have highlighted challenges with the evaluation of complex interventions and recommend that researchers should make explicit the rationale behind the choice of approach, what influenced their decision and appraise this in terms of strengths, weaknesses and transferable lessons. This was done.

5.4.1.2 Application of this research design and question to this professional group

This study is also unique in that it addresses this particular professional group, whereas many of the CRT studies have concentrated on physicians. The dynamics within a nursing administration structure (within which Midwifery falls in the South African context) are significantly different from those of independently practising physicians because nursing administration structures have a greater degree of pressure to conform to institutional norms and to follow protocols devised by others.

5.4.1.3 Application of this research design in this setting

While RCTs are common, there is increasing interest in the area of health promotion and knowledge translation where it is more appropriate to randomise according to a practice setting or other coherent grouping. Cluster randomised trials are appropriate for evaluating behavioural change strategies (Campbell, Steen *et al.* 1999) and for public health type interventions and policy implementation research (Miller, Sloan *et al.* 2003). Most of the published studies of this nature are set in North America and Europe (Gülmezoglu, Villar *et al.* 2004). There has been a major study in one province of South Africa looking at the effect of educational outreach on nurses on tuberculosis case detection and primary care of respiratory illness (Fairall, Zwarenstein *et al.* 2005), and some studies have been conducted in rural settings (Hayes, Changaluchab *et al.* 2005, Koethe, Westfall *et al.* 2010). This study among midwives in a rural setting in South Africa therefore adds to this under-researched area.

5.4.1.4 Pragmatic approach

The RCT approach is limited in that it is subject to strict conditions (Miller, Sloan *et al.* 2003) and does not realistically represent a health service context which is

more likely to require a pragmatic approach. Given the expected dynamics of a functioning health service, this study adopted a pragmatic approach, which was accommodated through design and analysis. Despite the statistical complexity posed by a pragmatic approach, the analysis of midwives' tests was handled systematically and demonstrated no significant difference between the two analysis groups (ITT and PPA). As this approach accommodated the real-world setting of this study, the findings can be expected to be relevant in similar settings.

5.4.1.5 Application of this research design to this educational intervention

This combination of interventions in this study (clinical facilitator development and support, and focussed in-service training in intrapartum care) addressing both individual needs and infrastructural support at a health service ('mini-system') level has not been presented in this type of setting and population. Educational interventions tend to be evaluated at the individual or cohort level but have not taken into account the effect of correlation within sites or differences between sites. This study addresses this particular gap.

5.4.1.6 Design of intervention

'Training' often consists of a one-way information transfer session which does not take account of the learning that needs to support the acquisition of information. The education component for both of these was intentional but had a high degree of flexibility to accommodate the identified learning needs of the participants.

Learning from the work of Gülmezoglu, Villar *et al.* (2004) it was planned to get as full as possible **coverage** as possible of any registered midwife who might be called to work in the intrapartum setting. In this way, all staff would be exposed to the same messages and be aware of the reasons for this programme and the expectations. This would assist those in clinical facilitator positions to work from a basis of common understanding. In all, of the 79

midwives in the intervention arm, 58 were exposed to the training directed to the midwives. The remaining 21 joined the study later, so did not have this training, but would have been exposed to the clinical facilitation activities.

The **presenter** was an experienced teacher and clinician and was prepared to be flexible as far as the timetable, content and methods were concerned (Fullerton, Gherissi *et al.* 2011). She was experienced in diagnosing conceptual learning difficulties and was able to address these as appropriate.

Attention was paid to creating an **environment** that was welcoming, respectful, that felt safe to engage in the programme, that indicated that there was purpose and planning, and was fun (Rogers 1969). The environment was designed to stimulate various senses to ensure that attention was engaged.

The **programme** addressed personal learning, but also highlighted the need for institutional responses to support change. This was developed more in the mentor training programme where they were functioning both at an individual level as well as being part of the institutional response. Interventions which vest in individuals and their capacity to sustain personal behaviour change are unlikely to be sustainable unless there is a consistent institutional / health system dimension supporting positive practices and sanctioning negative practices (Grol and Grimshaw 2003, Bingham and Main 2010, Scott, VandenBeld *et al.* 2011, Siassakos, Fox *et al.* 2011, Ivers, Jamtvedt *et al.* 2012). The intervention in this study, by including a clinical facilitation process (which among other things would employ the partograph utilisation audit as a tool for encouraging change in practice), attempted to develop an action that extended beyond the individual and harnessed a corporate dynamic.

The feedback from the clinical facilitators at the end of their programme was combined (translated from Afrikaans where necessary) and précised (**Appendix 16**). This did not form part of the formal analysis, but is included to illustrate the acceptability of the education programme and the range of responses, which indicates a strength.

5.4.1.7 Partograph utilisation instrument

The partograph utilisation instrument has been subjected to a credible validation process. A high degree of concordance and agreement was found for all measures for intra-reader reliability. For inter-reader reliability the instrument has good concordance and will deliver reliable measures for this population and the level of evaluator used in this study.

5.4.2 Weaknesses / limitations

5.4.2.1 Design issues

5.4.2.1.1 Sample size calculations

The unique context of this study had implications for the design where the ICC had to be estimated. In the absence of known ICC factors for these outcomes and for these populations, the assumptions made were that the ICC for the partograph would be 0.05 and for the midwives would be 0.1. These were based after due consideration of the context as well as ICCs established in similar studies. The ICCs that were calculated during this study were 0.28 for the partograph and 0.2 for the midwives. It appears that the cluster effect was under-estimated and that the variability between clusters was greater than expected. This under-estimation means that the design effect would have been higher and the study could have been under-powered.

The sample size calculations (presented in 3.7.3.1 and 3.7.3.2) indicated that there should be 406 partograph records in each arm of the study which would result in 14 clusters (seven in each arm); as well as 100 registered midwives which would amount to eight clusters (four in each arm). It was decided to use all the available clusters (17) (resulting in 1020 partograph records and 154 registered midwives) to provide insurance for inference of the study. This precaution appears to have been justified.

5.4.2.1.2 Cluster versus individual level analysis

Despite this study having fewer than 15 clusters in each treatment arm where cluster level analysis might be justified (discussed in 3.3.5.1), the disadvantages for the cluster level summaries made this option less attractive. This was particularly in relation to the cluster sizes for the secondary endpoint on midwife

knowledge and skills where all midwives were followed up during the study period, and personnel movements were dynamic.

5.4.2.1.3 *Statistical power*

There was a difference in the strength of the clinical facilitation between regions for the moderate-to-large sites. However, as this was a *post hoc* secondary analysis, the study was not powered to detect this degree of difference.

5.4.2.1.4 *Blinding*

The participatory nature of the design of the intervention package was such that the allocation to study arm could not be concealed from the managers, clinical facilitators or registered midwives. This was acknowledged in the planning of the study.

5.4.2.2 *Intervention issues*

5.4.2.2.1 *Number of clinical facilitators*

The planned study design included two clinical facilitators per site. This was fully implemented in region one. In region two only one clinical facilitator per site was identified. In this region, resignation, transfer out of the maternity department within a couple of months, and re-allocation of personnel all played a part in undermining the role of those who commenced as mentors. In order to try and maintain integrity of the intervention, a follow-up mentorship training programme was run (in month seven). A further three mentors were trained at this time, but the necessary threshold was still not achieved in all sites. Thus, partial remediation was done to maintain the integrity of the mentor portion of the intervention.

5.4.2.2.2 *Attributes of clinical facilitators*

The study design involved the identification of suitable clinical facilitators by the nurse managers of the participating sites according to the collectively pre-determined criteria. However, these were not always applied, but this was outside the control of the researcher. At one site the manager of the institution excluded herself from the mentorship training (while still offering her support), while at another site the manager put forward an inappropriate person as a mentor in the hope that this would address a perceived personal development need. This highlights the competing needs of management objectives and

research requirements. In a pragmatic study it is anticipated that different agendas might be at play. This was minimised as far as possible through the preparatory meetings, follow-up letters and *aides memoire*.

5.4.2.2.3 *Role of the regional MCWH programmes co-ordinator*

The participation of the regional co-ordinator for Maternal, Child and Women's Health programmes in the clinical facilitation programme was planned to contribute to the sustainability of the intervention, rather than being driven externally by the researcher for the duration of the project and then abandoned. The resignation of the programme manager in region two (the same region where only one person per site was identified for mentorship training) was a disappointing loss. Unfortunately, there was no one appointed during the study period who could continue in the supportive role.

5.4.2.3 *Instrument issues*

Some issues had not been identified during the pilot testing of the partograph utilisation instrument, and only became apparent during the evaluation process : there was no category to indicate that the partograph should not have been used because the woman was not in labour; it was not sensitive enough where there were minimal or single recordings; and the comprehensive assessment template did not include date, time and signature of the person making the observations. Two descriptions of item scoring which were of specific importance to this study needed clarification : 1) observations of the fetal heart were required both pre- AND post-contraction. This was accommodated as an additional item in the trial study, and all records checked; and 2) urine volume could not be nil and still have results of protein, blood *etc.* The urine scoring was revised after 120 records were evaluated, and all these records were re-evaluated. This limitation was dealt with satisfactorily.

5.4.2.4 *Testing and analysis issues*

5.4.2.4.1 *Partograph utilisation audit*

5.4.2.4.1.1 *Proxy indicator*

The considerations for appropriate indicators are presented in 3.5. The use of mortality indicators to indicate the standard of care would have required extending the study beyond a single province and this was not feasible for a

single researcher with limited resources. Therefore, the proxy indicator of a partograph utilisation audit was used. It was recognised that this had limitations in that there may be an underlying assumption that a completed observation instrument, such as the partograph, represented a high / acceptable standard of care. The researcher was aware of this and did not hold this view. Rather, the assumption underpinning the use of this proxy indicator was that the absence of a completed partograph limits the clinician's ability to make sound clinical decisions. Further, it was acknowledged that the emphasis was on the recording of observations rather than on the actual management. To this end the instrument was designed to capture the alignment of actions to the guidelines for the utilisation of the partograph and related intrapartum clinical management.

5.4.2.4.1.2 Lack of baseline scores

The partograph utilisation audit was planned to collect data across the 12 months following the intervention. All profiles described indicated that the samples in each arm were comparable and suggested that these were sufficiently robust to be able to determine change and differences between arms. However, it is acknowledged that the lack of baseline prevented the teasing out of some of the dynamics of the partograph scores. This was particularly apparent in relation to the differences between regions. In region one the control and intervention arms exhibited no difference, and had higher scores than region two, whereas in region two there was a clear difference between the two arms (see 4.3.2.1.3). The availability of baseline partograph scores could have helped resolve this.

5.4.2.4.1.3 Non-utilisation of the partograph

The 30% non-utilisation of the partograph could have impacted the analysis. The fact that there was parity between the two arms reduced the likelihood of bias.

5.4.2.4.2 Midwife tests

5.4.2.4.2.1 Different participants at different timepoints

The study population of midwives was not constant at all time points during the data collection period. This was anticipated. Because staffing and health service

dynamics are ever-changing, it would be unrealistic to approach the study as if it were taking place under controlled laboratory conditions. Even if it were possible to do so, the findings would be largely irrelevant as the conditions would not be able to be replicated for a service-wide rollout.

This dynamic presented a challenge for statistical modelling in that the denominators for test one and test three (which included both intervention and control sites) were different. The pattern and reasons for non-completion were similar for both arms of the study and are illustrated in the flow diagram for participants (Figure 4.4.1). Those components answered by the majority were analysed. Where datasets were incomplete, it was decided to split the theoretical and clinical portions of the test, and conduct analyses separately. Analyses were performed both on an 'intention to treat' basis as well as 'per protocol', and reported accordingly. For test three, an analysis performed on the pattern of completions / non-completions showed that there was no appreciable difference in the pattern of non-completions between the two arms of the study (4.4.2.1.3 and tables 4.4.8a and b).

Regardless of the type of analysis carried out (ITT or PPA) the results demonstrated the same pattern and similar levels of statistical significance. The sensitivity analysis (4.4.2.3) indicated that there was minimal selection bias. Thus this limitation was managed effectively.

5.4.2.4.2.2 Within-arm comparisons

This study reported on analyses conducted at timepoint three between arms.

The baseline values did not differ and the groups were randomised so the intervention effect could be validly estimated. Doing within-arm comparisons (which could have caused an intervention effect in the control arm) was not planned for this pragmatic study, and would not have been feasible given that a number of midwives were not available for each of the test timepoints.

5.4.2.4.2.3 Intervention arm analysis – missing timepoint

At one of the intervention sites, test two, which should have been conducted in month three of the study, was only completed at the time of test three in month

twelve. Although the clinical portion of the tests had been done at the correct time, this entire site had to be discarded in terms of test two. Although the discarding of this data was regrettable, it did not contribute to the main outcome of the study and did not materially affect the findings of the study. As indicated in 4.4.3, because the number of participants in the intervention arm who completed all three tests was reduced, this prevented a robust before-and-after analysis being done (Katzenellenbogen, Joubert *et al.* 1997).

5.4.2.5 Health service issues

The approach to this project was to see if the intervention could be sustained within regular health service dynamics. To this end, the researcher adopted a 'light footprint' approach and did not develop an engaged and encouraging relationship with mentors beyond the training period (as one would normally build into the implementation phase for this type of activity). This would have introduced an additional element that would not normally have been present in the service and therefore not sustainable beyond the study. This approach was therefore a stringent test of effectiveness of the intervention.

There were numerous demands being placed on all sites, and these impacted many of the registered nurse-midwives, e.g. pharmacology / dispensing course, ultrasound training, immunisation mobilisation, and initiatives for other programmes and research projects. Within this context, even minimal needs for assistance multiplied across a number of competing activities can place a heavy demand on the sites. While careful attention was paid to negotiating and gaining permission from each level of the health service, with changes in staff and in leadership, carefully developed plans were not always communicated onwards.

5.4.2.6 Method issues

Although pragmatic evaluation has the advantage that its focus is on the information needs of policy or decision makers and therefore is likely to result in the findings being used (Milne, Scotland *et al.* 2004), some authors suggest that this type of evaluation risks compromising methodologic rigor to please policy makers (Pawson and Tilley 1997). Further, there is an expectation that pragmatic studies will yield lower intervention effects. In this study there was

careful attention to design, sample size and sampling framework, attention to potential bias, and appropriate use of analytical models. Although there was support for this study within the provincial health department, it was not a commissioned study so the researcher was independent and able to implement the study according to the required standards for this design. An advantage of this study is that the findings are reality-based and therefore likely to be able to be replicated either for subsequent studies or for programme roll-out initiatives.

Milne, Scotland *et al.* (2004) stated that a study of this nature may not be able to identify which specific component or components are effective and may be limited to evaluating the programme itself. This is indeed the case in this study. However, drawing on the systematic reviews carried out by Grimshaw, Thomas *et al.* (2004) and Oxman, Thomson *et al.* (1995) it is clear that there was no single intervention or particular combination of interventions that could be regarded as a 'magic bullet' in addressing behaviour change. Rather, a particular combination of two or three interventions and strategies tended to yield positive changes.

5.5 Generalisability and external validity

When considering the generalisability of a study one has to determine how representative the sample studied is of the broader population in order to determine if the trends identified in the sample can be applied to the general population, i.e. its external validity. Aspects to consider are the breadth or narrowness of the definition of the accessible population, if the sample was randomly selected, the degree of control in experimental studies, and how bias might have influenced the findings (Hayes and Moulton 2009, Grove, Burns *et al.* 2013).

Given that this study was located in various sites it was recognised that a randomised controlled trial would be inappropriate as such a design makes the assumption that all observations on individuals are statistically independent of one another. When investigating group settings one needs to recognise that there will be some degree of correlation between individuals within a cluster.

Therefore, a cluster randomised trial design was chosen. The rationale for this choice is discussed in 3.2.

There were various levels of sampling. At the cluster level there was a stratification framework based on geo-political region and size of the service. The clusters were randomly assigned (3.7.3.3) and the profile of the services indicated that the stratification had yielded comparable arms (4.2.1). Within clusters the labour records were sampled systematically and the inclusion criteria were broadly representative of a healthy labouring population (3.7.3.1). All midwives working in the study sites were invited to participate in the study thus the sample was the population (3.7.3.2). The profiles of the labour records (4.2.2) and of the midwives (4.2.3) indicated that the randomisation was effective. Drop-outs from the midwives were minimal and balanced, thus not affecting the overall balance.

This study was conducted under real-life and not 'laboratory' conditions. The intervention was a single package consisting of two parts, *viz.* midwives' focused in-service training conducted at the beginning of the study, and clinical facilitation which was planned for the 12 months of the duration of the study. For those midwives who joined the study after its commencement, those who were in the intervention sites would not have had the midwives' training portion but their performance would be influenced by the clinical facilitation. Therefore, the dynamic nature of the staffing fluctuation had an impact on the midwives' availability to undertake testing at various times during the year. The pragmatic nature of the study anticipated the changing numbers of participants during the data collection period and analysis was designed to accommodate this by undertaking intention-to-treat (ITT) and per protocol (PPA) analysis. The sensitivity analysis (4.4.2.3) concluded there was no selection bias in respect of the theory scores and minimal bias in terms of the clinical scores. Bias and contamination are presented and discussed further in 3.3.5.4, 5.2.1.2 and 5.3.1.2.

Grove, Burns *et al.* (2013) caution against making generalisations unless these can be supported by many studies. Given the rigour applied throughout this study and the compliance with the reporting requirements set out in the various CONSORT statements (Zwarenstein, Treweek *et al.* 2008, Moher, Hopewell *et al.* 2010, Campbell, Piaggio *et al.* 2012) enabling comparisons with similar studies, the evidence is credible and likely to be able to be applied in similar settings.

5.5.1 Key aspects of the setting which determined the trial results

The significant interaction was with the intervention arm and size of service. This suggests that there was sufficient opportunity for midwives to apply what they had learnt. Further, it was likely that a clinical facilitator would be on duty in addition to the ward staff and would be able to support and encourage the practices that were covered in the training programme. This contrasts with the smaller services (where there was no difference between arms) where a single registered nurse-midwife would be on duty while attending to patients with a wide variety of illnesses.

During the clinical facilitator development programme there was time for the clinical facilitators to plan their implementation and sustainability strategies. The clinical facilitators in region one (which had a full complement) decided to meet quarterly to support each other and share ideas. The researcher was not included in these meetings. As part of their strategy, the clinical facilitators requested to use the test memoranda for workshopping. Once all sites had completed a test, this information was sent to all clinical facilitators (both regions). The test for each timepoint was different so this could not be used for 'swotting'. Clinical facilitator pairs (per site) adopted or developed strategies for their units which were shared with region two during their training programme, and adapted to local contexts. Some of the strategies communicated to the researcher were regular partograph utilisation audit, rotating responsibility for partograph audits, key messaging, individual feedback (**Appendix 8a**).

While the clinical facilitators might have regarded themselves as supporting the other midwives, they themselves gained skills that they required to perform their own roles. Although they were required to monitor the quality of care given in their services they did not have tools in terms of conducting an audit or managing a clinical review. These were built into the clinical facilitator development programme thus fitting them for purpose. In this way all participants gained something that was professionally valuable.

Given the pragmatic nature of this study, apart from the initial training input (four days for the mentors and two days for the midwives), nothing external to the existing service was added. As the health service was similar to other public sector services, this model should be able to be replicated.

Competing demands on the time of the midwife, especially being rotated out of the maternity labour ward or being redeployed to other departments, prevented optimum benefit being gained.

The intervention was planned with minimal external resources. If this model were to be rolled out by the service, this should be more actively managed, in terms of clinical facilitation and regular focused in-service education to meet the identified needs of the staff and consolidation of clinical allocation.

5.6 Reflexivity and pragmatic approach

This section will present the role and purpose of reflexivity and its application to a pragmatic trial⁶⁶. This will be followed by examining the impact of the researcher on the research process and *vice versa*, before concluding with a reflection.

⁶⁶ This section refers to the human interactions between the clinical facilitators and the midwives, and the researcher. Given that the data collection for the partograph audit was done retrospectively using a systematic sampling technique starting with randomly allocated starting point, there was no influence of the researcher on the selection of partographs. The partograph utilisation was evaluated using a pre-designed and tested partograph utilisation instrument. There was no relational dimension between the partograph data and the researcher, nor was any meaning ascribed to the data through a relational process. Therefore this falls outside the brief for reflexivity.

5.6.1 The role and purpose of reflexivity, and its application to a pragmatic trial

Research designs in the positivist tradition, such as this study, aim to eliminate or minimise bias in order to be able to generalise findings. In such a study design, subjectivity and personal influence in the processes and analysis would be regarded as weaknesses. By contrast, in the naturalistic paradigm the involvement of the researcher can be a strength and should be made visible. Qualitative approaches to research recognise that the researcher becomes part of the process and may impact the way in which the data is collected, analysed and shaped. In order to ensure that the research is credible, the qualitative researcher is required to engage in reflexivity where s/he continuously and consciously reflects on her/his actions and responses by the participants to her/his actions and *vice versa*, and the process of collecting and interpreting the data (Webb 1992, Finlay 2002, Dowling 2006, Lipp 2007, McCabe and Holmes 2009, Marshall, Fraser *et al.* 2010, Grove, Burns *et al.* 2013). In its broadest sense reflexivity can be viewed as a way of contributing to the production of knowledge from experience by considering the impact of one's position and actions (Lipp 2007).

Unlike studies which are qualitative in nature with an interpretive element based on the interaction between the researcher and the participant, this study was pragmatic, the researcher participated in the intervention training, and collected data for which standardised instruments were used. (This is in contrast to many other pragmatic trials where the intervention may be impersonal, e.g. a treatment regime.) Therefore it was prudent to consider the potential effect of the researcher's presence on the intervention training and data collection (particularly of the midwives' clinical skills, which required contact between the participant and researcher) in order provide greater transparency of this study.

There are numerous typologies of reflexivity described (Finlay 2002, Lipp 2007). Given the nature of this study, reflexivity as social critique which examines the impact of power relations *inter alia* in the research setting (Finlay 2002, Lipp 2007), is an appropriate approach. Social critique focusses on how to

acknowledge and manage the power imbalance between the researcher and the participant, e.g. differences in social positions in relation to class, gender and race should be acknowledged (Finlay 2002). Further, the power to determine the research method decisions (e.g. design chosen, inclusion, questions / testing, analysis) and the examination of the limitations of the decisions should be addressed (Fontana 2004).

When reporting on quantitative studies, the convention is to write in the third person, emphasising the neutrality, objectivity and invisibility of the researcher. However, when embarking on a reflexive exercise, whose nature is by definition subjective and responsive to the processes undertaken during the conduct of the research, it is appropriate to write personal reflections in the first person (Webb 1992, Finlay 2002, Marshall, Fraser *et al.* 2010), which I will do for the remainder of this section (5.6).

5.6.2 The impact of the researcher and her presence on the research process

I pondered on how I could have influenced the outcome of the study, while simultaneously having to ensure that I minimised potential bias in other aspects of the research process (to be consistent with the quantitative nature of the study). The concepts I will address are identity, research design, interaction and integrity. For each of these I will explore the inherent power issues as well as the limitations of the decisions made and how potential bias was assessed and addressed. The questions in the reflexive framework proposed by Lipp (2007), which prompt broader and deeper reflexivity, guided my thought process.

5.6.2.1 Identity and power

In South Africa one cannot escape the impact of race on any relationship. As a professional, middle class woman, previously categorised as 'white', I carry social privilege. Further as an academic and leader in my profession I have attained status and respect of my colleagues. However my life experience has been one where I have consistently functioned personally and professionally outside the social categories that could define me. Relating to the participants in the study, we were all registered nurses and midwives (most of whom were

women) so had had similar experiences in our role in the health service. With some participants I shared a racial category, with others a class identity, and with others a mother-tongue. Some participants were active lifelong learners but many had not undertaken further education after qualifying. The diversity in the training groups ensured that there were commonalities and differences across the various relationships. There were similar relational dynamics in the clinical facilitator training group. However given that most of them occupied work-based managerial roles and this could have been difficult to shed in the larger group, the decision was made that the clinical facilitators' training should be separate from the other midwives.

5.6.2.2 Research design and power

I chose a pragmatic approach as this reflected the real-world situation instead of one that was controlled. The intervention was planned to be sustainable within the existing resources of the health service and not dependent on additional external input, materials or ongoing contact.

As the person delivering the intervention training, there was every expectation that my presence and contribution should influence the outcome through the content and process of the training. The educational materials used were provided to all participants meaning that everyone received the intended package. This included personal skills development and strengthening processes within the institutions that would support better quality of care in labour. My intention for the clinical facilitation rollout was that each site's facilitators would develop a plan jointly that would be responsive to the needs of their sites. A standardised plan would be unlikely to suit the institutional dynamics, staffing and learning needs in the diverse sites. This was consistent with the pragmatic nature of the study, and devolved power to the appropriate roleplayers.

My understanding of midwives' realities in rural settings⁶⁷, coupled with my experience and skill as an educator and midwife clinician, equipped me to

⁶⁷ Having worked in far north Queensland, Australia, where the nearest referral hospital was a minimum of four hours flying time away and any medical support required radio

design a training programme that was appropriate and responsive to the learning needs and learning styles of the midwives (see 3.8.2). I built flexibility into the programme to accommodate the dynamic within each training group. This in itself was pragmatic in that I recognised that the learning programme needed to accommodate the diversity in the groups. Although I was comfortable with this approach as an educator, I constantly had to keep alert to the potential research implications for standardisation.

As far as possible all midwives working in the study sites in the two regions were included in the study. This planned saturation coverage aimed for all midwives in the intervention arm to receive the same in-service training. The intention was that on their return to their workplaces they would all be aiming for the same standards for intrapartum care, rather than a couple of enthusiastic participants arriving back to find little interest from colleagues who had not been included, and being disempowered in their endeavours to improve intrapartum management. I was aware that it was highly probable that the midwife participant population would be dynamic and could not be controlled as people made choices about their employment. The various changes in participants had to be documented carefully. Although the fluidity of the participant population felt uncomfortable, I had to trust the integrity of the research design and the rigour of the statistical analysis, e.g. use of both 'intention to treat' and 'per protocol' analyses.

All nursing operational managers were invited to participate in the clinical facilitator training for maximum inclusivity. The plan was to have two clinical facilitators per site to ensure continuity and support, and was justified when a clinical facilitator was redeployed out of maternity (5.4.2.2.1). Unlike the midwives who undertook the in-service training, the clinical facilitators were not required to participate in the theory and clinical tests. This decision was made because in most cases the nursing managers would not have worked in a

communication, a boat trip or helicopter evacuation to a local hospital, I have personal experience of working in a remote area with limited resources, needing to be able to make accurate clinical judgments and instituting the necessary tests and/or treatments.

maternity setting for some time (if ever) and may have been unwilling to participate in the study if they were required to undergo the testing. It was felt that their involvement, learning and support was far more valuable than their test scores.

In planning the study I did not intend to take on the role of mentor to the clinical facilitators. Outside of the training periods, I had no direct contact with them unless they initiated contact. However, during my site visits at month three I attempted to make a brief courtesy visit with each clinical facilitator. Although ongoing interaction could have made for a more robust intervention, it would have been more resource-intensive and would have defeated the intention of sustainability in the long-term. Thus I was dependent on the two regional MCWH managers to take forward the support of the clinical facilitators. The commitment to not introducing any additional support was put under further strain when one of the regional MCWH managers resigned halfway through the intervention period and she was not replaced.

5.6.2.3 Power and interaction

5.6.2.3.1 *Positive professional engagement*

The first contact I had with participants in the intervention sites was during their participation in the training programmes. My first goal was to make them feel welcome, which I did by paying special attention to the environment – well-prepared, warm and fresh atmosphere, educational materials arranged in personal packs, and hot refreshments available (as some might have travelled up to 2.5 hours in an early winter's morning). I recognised that there was a potentially strong power gradient in the relationship between (and among) us (McCabe and Holmes 2009). Therefore I intentionally developed a positive, co-operative, collegial and relational atmosphere to minimise this. We used first names instead of rank titles, we sat in a horseshoe-shape so that there was no identifiable power seat / position, and participants were invited to identify the issues that they wanted or expected to be covered so that they were empowered to influence the programme. These were put on newsprint and displayed, and participants were encouraged to add to them if new needs

surfaced, thus giving them the opportunity to influence the programme. I referred to these issues during the course of the training to ensure that these were covered satisfactorily. I used various teaching methods that emphasised participation and co-operation, e.g. group work, peer learning, roleplay, icebreakers and energisers. The clinical facilitators spent four consecutive days together in a residential setting whereas the midwives' in-service programme consisted of two days, one month apart. This sustained contact (intense or protracted) built some degree of familiarity and comfort among the participants in the intervention sites, and between them and me.

As part of my conduct of a learning programme I asked the participants to give feedback. With the larger groups of midwives they gave (small) group feedback after a participative process where they were required to give three positive and three negative responses. I did this so that they had to engage and think carefully about their feedback and give them 'permission' to express negative responses. The feedback was given verbally by each small group and a picture of the overall group response for the day was developed. This strategy can elicit responses which otherwise might not be expressed, especially when there has been a focus on relationship building. Because the feedback was done as a group there was perceived safety in expressing their feedback. My intention was to be able to reflect on the feedback received and strengthen or limit (as appropriate) the identified factors with subsequent groups.

5.6.2.3.2 *Changing the nature of the interaction*

In being true to the study design, I intentionally terminated the learning relationship with the participants explaining to them why this was necessary. However, when I arrived for the clinical testing I was generally received warmly. Participants were keen to engage and some asked me to conduct a clinical teaching round or consulted on specific clinical cases. I reminded them that I could not as this would alter the intervention and jeopardise the study. This explanation was accepted.

In order to reduce my interactions with the midwives after the training programme was complete, I explored the possibility of someone else doing the data collection for the clinical testing portion of the midwives' tests, but after consultation I decided that there was no single, suitable person for whom this would be feasible. I undertook the testing in order to maintain credibility and consistency in the process. While there was no standard written script I adopted a standard pattern for this activity. I aimed to put people at their ease, to remind them of the anonymous and confidential nature of the test, and to assure them that there was no time limit to complete the test, but I limited any further engagement. The instrument was structured and had the same format for each clinical test which included the instructions. The same clinical models were used at each timepoint although different fetal positions and cervical dilatations were presented.

The midwives in the control group had no knowledge of me from the study as they were not exposed to the intervention (training). The only contact was common personal courtesy in explaining the research study, inviting them to consent to participate, and facilitating the clinical tests. Initially I was greeted warily, but once they realised that the tests results were confidential and would not be shared with their supervisors or managers, and that I was not going to impose my will on them to participate, they were generally willing to enroll. The same pattern for organising the clinical skills tests was adopted as was done for the midwives in the intervention arm. Some midwives in the control arm regarded me as a potential resource but I had to make it clear that I could not conduct a clinical teaching round or be consulted on any clinical cases as this would influence the study. This was accepted. A number of midwives at the control sites expressed their thanks for the tests as they said that these served as a reminder to them of what should be familiar, and some indicated that this has encouraged them to pay more attention to their knowledge base and to developing their clinical skills. This confirmed the wisdom of the decision not to test the control group at month three as this could have resulted in a confounding effect.

5.6.2.4 Integrity – minimising bias

Given that I had met and established relationships with half of the participants at the commencement of the study and was also the data collector / evaluator of clinical skills, I had to consider how I should engage with the data. All midwives received a study code which could not be linked to a site or study arm except by referring to a separately stored code book (3.8.7). For the evaluation of the midwives' tests I used a standard marksheet. During the clinical tests there was minimal interaction (apart from putting people at their ease) and no interpretation as the clinical skills assessments were standardised. Data was not dependent on interaction, nor was any meaning ascribed to the data through a relational process. The data were analysed through the use of recognised statistical processes.

It is possible that the familiarity of the participants in the intervention group with me (at month three (test two) and at month 12 (test three)) could have had some influence. However as the composition of the intervention group at test three was less than 50% of those who had undertaken the training, this is unlikely to have caused any significant bias.

5.6.3 The impact of the research process on the researcher

As an engaged researcher I had to confront the lack of control and lack of contact during the unfolding process. Relinquishing control was a challenge and I had to resist the temptation to step-in by creating clear boundaries for relating to the participants and ensuring that I remained true to the study protocol and trusted the study design. The contradiction between the role of researcher and that of educator was challenging. The researcher role required me to maintain distance and objectivity, whereas in the educator role I was using a participative approach and promoting positive relationships, collegiality and teamwork. Given that fundamentals of mentoring guided some of the intervention training, I would like to have had more engaged, ongoing relationships with the clinical facilitators after their training programme was completed. When I had to relinquish the educator role after the training was complete, I found the distancing in the relationships an uncomfortable space to inhabit.

I also had to accommodate the pressures in the clinical situation as I could not withdraw people from their primary responsibility, *viz.* to render a clinical service, in order to undertake a research study activity, *i.e.* clinical test. Accordingly I arranged to undertake evaluations when it suited the service and the timing would be altered if the service demands changed. This meant that I made multiple visits to each site in my role as evaluator to cover different day and night shifts, and waited around a lot.

During the study I kept field notes on each phase of the process. I spent many hours alone driving long distances between sites and this gave me time to reflect on dynamics and challenges that were occurring. This thinking time was helpful to come to some understanding or resolution about issues. Working in a team would have assisted me by being able to discuss my anxieties and to engage in ongoing reflection with others who understood the context in the field.

I was most grateful for the steadfastness of my supervisors and statistical consultant, and for their wise and encouraging counsel.

5.6.4 Reflection and conclusion

In reflecting on this experience, I can sum it up in one word – paradox. The roles of researcher and educator were quite different in the behaviours required of me. The former required objectivity while the latter required responsive engagement. Engagement itself held a paradox where both interaction and distancing was required. I had to manage the impact of these paradoxes on interpersonal dynamics with care, sensitivity and respect. Another paradox was that of standardisation and flexibility. Although the content of the intervention training was standardised, its delivery displayed some flexibility so that there was integrity of the learning programme which related to the needs of the adult learners. The rollout of the intervention also displayed flexibility / independence as clinical facilitators adopted strategies that were suitable for their settings. I consciously acknowledged these tensions and held them lightly, knowing that the activities were true to the planned design.

The personal paradoxes were of being together yet alone, and passionate yet objective. I was engaged at a health system level with the midwives and required active participation during the training days yet the responsibility for the project was held individually by me. My passion for quality midwifery care and commitment to my colleagues and profession had to be balanced by the engaged yet objective position as a researcher committed to scientific and ethical integrity. This experience has been a life-enriching challenge.

5.7 Conclusion

In 1.8 it was suggested that one needed a fresh perspective – that of appropriate clinical decision making, and thus on improved clinical outcomes, and the following two questions were posed :

1. Is it that the process and progress of labour are inadequately understood, and thus observations are done incompletely resulting in premature or inaccurate conclusions?
2. Is it that clinical skills are poorly developed and that there is little confidence in the ability to make a firm clinical finding?

From the evidence presented, it suggests that there is room for significant strengthening of the registered midwives in terms of their understanding of labour, the associated clinical skills and their interpretation. The control arm (representing the *status quo*) illustrated that without intervention the standard of care, and knowledge and skills of the midwives deteriorated. It is beyond the remit of this study to determine if the weakness could be laid at the door of the basic educational preparation of midwives, or if this degree of attrition is common to nurses and midwives across the board. It is necessary for a purposeful programme of continuing education to be established to support improvements in practice. The intervention arm showed a positive trend in the partograph utilisation and there were higher scores in the knowledge and skills of the midwives indicating that the intervention package was effective.

Various implementation strategies have been shown to be effective in supporting practice change in different settings. These are optimised when they are offered as a multiple package. In this study a multifaceted intervention targeting midwives and intrapartum care was used, thus addressing both the individual and health service level dynamics in a way that was sustainable within existing resources.

Greater attention should be paid to the system in which knowledge translation occurs and in understanding the factors that enhance or hamper the change process. Contextual sensitivity is required in designing the appropriate mix of techniques for a practice improvement intervention. Identifying barriers to change is an area requiring further investigation.

This was a novel study in that it applied a CRT design in a primary level setting in a low-to-middle income country in an under-researched area and professional group. The analysis accommodated a pragmatic approach so that the findings could be replicated in similar settings. Such a combination has not been found in the literature.

Chapter 6 Recommendations and Conclusion

6.1 Introduction

This concluding chapter will present a summary of the study (reflecting on how the chosen methods addressed the aim, objectives and hypotheses) and the particular contribution to the field. Thereafter, the research and clinical implications of the study will be presented followed by the recommendations for research, strengthening midwifery capacity and health system strengthening. Plans for dissemination of the results will precede the final conclusion.

6.2 Summary and reflection of the study

This pragmatic cluster randomised trial was undertaken in a middle-income country, in a non-urban setting amongst registered midwives who work in professional nurse posts (i.e. not designated Midwifery posts). It took account of the dynamics within and between health service sites / clusters, and reporting was done according to the CONSORT requirements.

The aim of the study was to evaluate the effect of an intervention package of focussed in-service intrapartum care training and clinical facilitation on the quality of clinical management in labour by registered midwives in primary level public sector maternity facilities in rural Western Cape province, South Africa.

The in-service training was facilitated by an experienced midwife and educator. A number of learning strategies were employed – workshopping, case-based learning, clinical skills demonstration and development, with a high degree of participation. In addition to the in-service training package for all registered midwives, the clinical facilitators received training in creating a positive climate for change, mentoring, audit and feedback, and motivational interviewing as a strategy to encourage professional behaviour change. Clinical facilitator pairs

(per site) chose and developed site-specific strategies which included key messaging, audit and feedback and a clinical facilitator support group. A major consideration for the intervention package was that it should be a sustainable programme within existing resources rather than one where there was a major external injection which could be withdrawn in a subsequent funding cycle. Except for the training package, no additional resources were introduced into the system thus making this a rigorous test of the effect, feasibility and sustainability of this intervention.

Null hypothesis one relating to the primary outcome was accepted, i.e. a well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators does not improve the standard of practice of registered midwives. The quality of clinical management was measured by an audit of the general labouring population (rather than one where mortality or morbidity was the basis for audit) where partograph utilisation was used as a proxy indicator. Although the partograph utilisation score did not demonstrate a statistically significant change, Bootstrap regression analysis showed a statistically significant difference in favour of the intervention arm in those partographs with better quality completions. The score for 'adherence to clinical management guidelines', which constituted a sub-objective for the primary outcome measure (objective one), did show a statistically significant difference in favour of the intervention arm. The consistent direction of the results suggests that a higher degree of input in the intervention package, e.g. more support of the clinical facilitators in their role as system strengtheners may have enabled a statistically significant result to have been obtained.

The partograph utilisation audit revealed that 30% of eligible women did not have a partograph used to document the process and guide the management of labour, leaving considerable room for improvement. In clusters with >500 births *per annum*, the intervention arm scored 6.4/47 units more than the control arm, this difference being statistically significant. This equates to a 13.6% difference.

Operational and staffing constraints made it difficult to comply with the requirements for the frequency of observations as set out in the provincial and national guidelines for intrapartum care.

Hypothesis two (relating to the secondary outcome (objective two)) was accepted, i.e. a well designed intrapartum care training package accompanied by on-site clinical facilitation by trained clinical facilitators improves the knowledge and skills of registered midwives. There were statistically significant higher test scores for midwives in the intervention arm. Sub-objectives related to objective two included knowledge (theory), clinical judgement, clinical assessment, ability to use and interpret the partograph, adherence to clinical management guidelines, all of which demonstrated statistically significant differences in favour of the intervention arm. Over the same time period the midwives' scores in the control arm deteriorated. This suggests that, without a concerted and sustained programme of professional development, the operational functioning levels are likely to continue to deteriorate.

Objective three also related to the secondary outcome which was to compare the performance of individual midwives at months one, three and 12 within the intervention arm. Given the low numbers that were present at all three timepoints the intended formal analysis was abandoned. This did not materially affect the study. A much larger group of participants would have been required to reach a level suitable for formal analysis.

Objective four was to compare the effectiveness of the current unplanned *ad hoc* approach to training, to full coverage training in combination with the clinical facilitation intervention. The results from the baseline and control arm results which represent the status quo, indicate that the current human resource development programme is insufficient. This study has shown that full coverage training in combination with the clinical facilitation intervention yielded positive results.

The overall challenge remains to convert knowledge and skills into improvement in the standard of care within a supportive health service arrangement, while the local challenge is to address the generally low levels of knowledge and skills among registered midwives, and to improve partograph utilisation.

This study was designed to be pragmatic so that it took account of the reality in the health services rather than testing under laboratory-like conditions which would be unlikely to be able to be replicated in any roll-out programme.

However, this required specific attention to statistical rigour particularly with regard to the changing numbers of midwives in each site over the 12 month period. For this reason both intention-to-treat and per protocol analyses were undertaken to determine if there were any significant biases. None were found.

6.3 The contribution of this study to the field

This was a novel study due to its combination of research approach, design, setting, focus and intervention. Increasingly, the pragmatic approach and CRTs are being employed in public health research where multiple sites are involved. No studies focusing on midwives and intrapartum care using this research design have been identified, nor have such studies in a developing country context been reported, so this study marks a contribution to developing this field.

The literature on the effectiveness of implementation strategies on professional behaviour change is dominated by studies involving physicians and physician interventions. Given the different spaces that are occupied by the different professions in terms of organisational and workplace cultures, findings from one professional group cannot be extrapolated to another. Thus the inclusion of a health service dimension in this study's intervention that aimed to facilitate change was novel. Introducing the role of clinical facilitator to an experienced midwife yoked the supportive, educational dimension to someone who already held clinical responsibility, and was feasible within the capacity of the existing resources of the health service.

Unlike the existing data in South Africa which reports on standard of intrapartum care where a maternal or perinatal death has occurred, the findings in this study highlight that the general standard of intrapartum care and the knowledge and skills of midwives in these two rural regions in South Africa need considerable strengthening to meet the needs of the health service.

The partograph utilisation checklist which was developed demonstrated a high level of content validity and yielded reliable responses for the population studied (3.6). It is suitable for other research studies requiring such an instrument. It may also be used for clinical audit as part of an ongoing quality assurance exercise for intrapartum care.

The ICCs for partograph utilisation evaluation and midwife testing have been established. These are higher than the ICCs usually applied for this type of study and should be considered for similar studies in the future.

6.4 Research and clinical implications

6.4.1 Research implications

6.4.1.1 Establishment of ICCs for this population

The ICC values for this population / setting / professional group can be considered for similar research studies in the future.

6.4.1.2 Availability of a partograph utilisation instrument

The partograph utilisation checklist is suitable for research purposes. It may also be used for clinical audit as part of an ongoing quality assurance exercise for intrapartum care.

6.4.2 Clinical implications

6.4.2.1 Lack of recording of observations

The lack of recording or poor reporting of observations utilising a partograph (with or without the use of the alert and action lines) results in sub-optimal standard of care. In addition to the 30% of records where no partograph was utilised, a further 21.9% of the total records had more than 1.5 hours without

any observations. This sub-optimal monitoring has serious implications for the care of women in labour as approximately 50% of all labours were lacking the observations necessary to guide clinical management decisions and care.

6.4.2.2 Low scores

Low scores, both in the partograph audit and midwife tests, indicate that there are large knowledge and skills deficits in areas where one would expect midwives to be competent. This contributes to a lack of understanding about the importance of the observations. In addition to low scores at baseline for the midwives' tests, in the control group the scores deteriorated indicating the critical need for ongoing in-service training and professional development. In the partograph utilisation audit the scores were low with the median percentage score lower than 50% in both arms for a number of the variables. This indicates that the possibility of identifying clinical problems early is compromised, and thus the clinical management.

6.4.2.3 Lack of clinical judgment

An inadequate knowledge and skills base does not facilitate effective clinical decision-making, and contributes to the 'third delay' described by Thaddeus and Maine (1994), i.e. delay in the provision of adequate health care, which would include implementing clinical management, identifying problems early, and appropriate referral.

6.4.2.4 Second stage observations

A limitation of the partograph design is the lack of accommodation for recording observations in the second stage. The impact of this in this study is that the lack of second stage monitoring made many partographs appear incomplete. There seems to be a lack of importance given to second stage monitoring, e.g. no observations recorded from full cervical dilatation to birth, or cessation of observations after the decision to perform a caesarean section, even though there might be a delay of several hours. At the time that this study was conducted there was a lack of guidelines at national and provincial levels for observations in the second stage of labour and the appropriate frequency.

6.5 Recommendations

Recommendations from this study relate to research, strengthening midwifery capacity and health system strengthening.

6.5.1 Research

6.5.1.1 Instrument design

The partograph evaluation instrument needs to be tested in field settings to determine its utility for routine labour record audit. Minor amendments indicated in 5.4.2.3 should be made.

6.5.1.2 Design of educational outreach

Studies are required that will clarify the optimal pattern of educational outreach activities in terms of frequency, duration and specific strategies employed.

6.5.1.3 Barriers to utilisation of the partograph

Given that the partograph is regarded as the standard of care for labour, it is critical to study systematically the barriers preventing midwives from monitoring labour, utilising the partograph effectively, and acting on these observations. Consideration may need to be given to the development of a simplified partograph which would need to be tested rigorously for effectiveness. A modified version of the partograph, which is colour-coded and omits the descent of the presentation, has been adopted in parts of Asia (Mathews, Rajaratnam *et al.* 2007). Qualitative research could reveal the nature of the barriers so that appropriately designed interventions can be implemented and contribute to improvements in clinical monitoring, recording, and decision-making.

6.5.1.4 Barriers and supports for professional practice change

Barriers and supports to changes in professional practice need to be identified and future studies to be informed by theories of behavioural or organisational change (Fox and Khan 2010, Michie, van Stralen *et al.* 2011), and with due consideration for the social context within which the activity takes place (Penn-Kekana, McPake *et al.* 2007, Konteh, Mannion *et al.* 2008, Mannion, Konteh *et al.* 2009). It is possible that once these are identified and addressed comprehensively, the impact of other interventions like audit and feedback may be more effective and sustained.

Observational and qualitative studies can help to define the context within which practice occurs, and the prevailing social interactions and processes operating within it. This will enable a greater understanding of the critical success factors in behaviour change initiatives and help to identify crucial elements of effective change.

6.5.1.5 Research amongst non-physicians and in developing countries

The bias in the literature towards physician-orientated practice in urban settings and in high-income countries needs to be balanced with research in low-to-middle income countries and among other professional groupings. There is a need for good quality studies of interventions and processes with well-designed implementation strategies (based on an appropriate theoretical framework) to be undertaken, which take account of the contribution of midwives and nurses, and the context of their work and their scope of practice.

With midwives and nurses being more available in rural areas where there is likely to be less support, this is a research area that requires attention in order to deliver better quality health care in these underserved locations. This requires appropriate partnering with academic institutions and / or professional associations / networks / federations, and regional or international partnerships for skills development and multi-professional strengthening where appropriate. This issue could be raised with research councils and funding organisations so that the research agenda can be influenced in this direction.

6.5.1.6 Conceptual clarity in the field of mentorship

A framework to determine and describe the broad field of mentorship, clinical supervision and clinical facilitation is necessary. This could *inter alia* include the purpose, the setting, the target population, the nature of the relationship and the status of the activity (e.g. mandatory or not). Standardised definitions would enable growth in research and understanding of the true contribution of these activities.

6.5.1.7 Effectiveness of mentorship

Effectiveness of mentorship in terms of patient and/or health outcomes needs to be addressed in the research agenda for this field.

6.5.1.8 Costing analysis

Research into practice improvement interventions should be costed as it may be found that some interventions are more cost-effective than expected (Groenewald-Neethling 2010). Further, it is important to know which interventions are affordable and feasible, particularly in low resource settings.

6.5.2 Strengthening Midwifery capacity

6.5.2.1 Continuing Professional Development

Continuing Professional Development should be compulsory for all registered midwives working in maternity (specifically intrapartum) settings. These staff need to have focussed in-service training on a regular basis to ensure that the levels of knowledge and skills are developed and maintained. Confidence should be built in the midwives by developing their capacity to address the knowledge and skills deficits apparent. Any educational activity should be designed using adult education principles where the learner is an active participant in the process as opposed to being a passive recipient. Such a programme should be facilitated and co-ordinated by the Human Resource Development department, and could be supported by the work of District Clinical Specialist Teams (DCST) and linked to academic Midwifery units.

It is recommended that there should be a link to an academic unit for educational outreach activities including regular site-based teaching, role modelling, and for development of critical clinical thinking to improve. This system currently exists for medical professionals only.

Given that the knowledge and skills deficit identified in this study relates to core knowledge and skills required by midwives, it is suggested that there should be a system of regular assessment of competence. This could be linked to annual re-licensing.

6.5.2.2 Clinical facilitation and clinical leadership

A team of clinical facilitators should be established in intrapartum settings to support a better standard of care, to promote evidence based care and to support the creation of a positive practice environment. Clinical facilitators should not be regarded as an extension of an administrative function, but should have clear role identity. Clinical facilitators should be supported by other clinical leaders / practice improvement teams, and should promote and support evidence-based practice.

Clinical leadership by midwives is required to promote a better standard of practice. With the planned establishment of District Clinical Specialist Teams (DCST) there is an opportunity for clinical leadership to be developed among midwives and nurses. Interaction between the clinical facilitators and the DCST could contribute to realising the goal of better care in the intrapartum period and contribute to meeting MDG5. Clinical leaders need to have appropriate knowledge and skills in addition to their clinical expertise in order to fulfil this role. Although the detail is beyond the findings of this study, it is suggested that these could include : teaching skills; the ability to conduct audit and feedback and to develop and implement evidence-based Midwifery care and practice change initiatives; and engaging with health care and health system policy development for quality care of women and babies.

6.5.2.3 Access to appropriate evidence

Midwives in rural settings need to be brought into the mainstream of knowledge translation in order to develop optimum professional practices and interventions. Thus they need to be able to access evidence that will inform their practice and be empowered so that they can apply the evidence. It will be essential to employ an appropriate implementation package consisting of elements presented in this study.

6.5.3 Health system strengthening

6.5.3.1 Clinical practice

Standardised guidelines, with flexibility to meet individual patient needs, should be actively promoted among all professional staff who are involved in intrapartum care. This includes the need for the development of a guideline for optimal monitoring in the second stage of labour. All professional staff have a responsibility to use the guidelines and not to comply with an inappropriate 'instruction'.

Active dissemination strategies are required to encourage the adoption and implementation of clinical guidelines. A multi-method approach is recommended.

6.5.3.2 Clinical records

A single, comprehensive patient-centred record used by all health professionals involved in the care of the parturient woman is required⁶⁸. The provincial nursing documentation committee should endorse the partograph as the key document for recording labour progress. A template for the observations and management of the second stage of labour is required.

6.5.3.3 Clinical audit

Clinical records and clinical management decisions made should be audited regularly using an appropriate standard, e.g. the partograph utilisation instrument for intrapartum care. Baseline performance is inversely associated with the effectiveness of audit and feedback so this is likely to be an effective intervention. There should be a monitoring system to ensure that guidelines and protocols are followed. Feedback should be given and local solutions developed to address problems identified. All relevant people, e.g. doctors, midwives, nurses and administrators, should be involved in the process and take responsibility for the actions decided upon. These should be monitored and reported at each feedback meeting in order to complete the audit loop.

⁶⁸ Subsequent to the study being undertaken there has been a revision of the national maternity record which has been rolled out over a period of 18 months in 2010/2011 across all nine provinces.

6.5.3.4 Institutional culture

An institutional culture needs to be developed where the importance of accurate and timely observations is instilled, supported and evaluated through clinical leadership. Various strategies may be used to do this including clinical facilitation, and audit and feedback. Positive practices should be supported and substandard practices sanctioned. For change to occur, an actively and intentionally managed programme with institutional support / pressure is required.

6.5.3.5 Human resource management

Intrapartum care services need to be staffed by sufficient numbers of registered midwives in permanent posts with the necessary skills to offer the required care. Capacity and competence should be developed among the registered midwives and these should be utilised within the maternity setting. Rotation of staff should be discouraged in favour of developing a strong cohort of competent midwives. Deployment of staff outside the labour ward when there are women in labour requiring their care should be stopped.

There are vastly different needs for staff support amongst services of different capacities for maternal and perinatal care. Health service planners, particularly those with responsibility for midwifery personnel and MCWH services, should develop the necessary capacity in these sites. In-service training should be regarded as a key activity necessary to deliver the health service required. The in-service training programme should be structured and intentional, not opportunistic. Staffing complements should be planned to enable the relevant staff to be released for these in-service training sessions.

6.6 Dissemination of findings

Findings will be disseminated through a research report to the national and provincial government health departments, and feedback will be given to the MCWH co-ordinators, clinical managers and participants in the regions where the study was conducted. The study will be presented at relevant national and international congresses and submitted to peer-reviewed journals.

6.7 Conclusion

In conclusion, this study addressed an issue of professional practice improvement in an under-researched area where the health needs are greater than in urban and high-income country contexts, and where the health care provider and health infrastructure is under-resourced.

This study showed that, although the standard of care in labour (using an audit of partograph utilisation as a proxy indicator) was higher in the intervention arm, this did not reach statistical significance. However, the proportion of partographs with a score of 27 or more was significantly higher in the intervention arm. The secondary outcome of knowledge and skills of midwives showed a statistically significant difference between the two arms, although scores tended to be low. This illustrates that the standard of recording is poor even where there are non-morbid outcomes. This is a critical issue that should be addressed energetically to ensure that lives are not damaged or lost.

This study proposes that educational outreach is essential and feasible, and recommends that barriers to utilisation and completion of the partograph should be determined. Further, it is necessary to consolidate the experience of registered midwives working in maternity settings. There is considerable scope for the Advanced Midwife and other clinicians in the envisaged District Clinical Specialist Team to provide the necessary clinical leadership.

There is no lack of knowledge of how to avert maternal and perinatal deaths. To realise this, the health care system needs appropriate infrastructure to provide excellent care, e.g. appropriate staff training, efficiently functioning systems, monitoring and evaluation of health outcomes. Without this, South Africa will continue to achieve poor outcomes despite meeting international process indicators for preventing maternal deaths. As the deadline for the Millennium Development Goals approaches there is much work still to be done to improve the wellbeing of pregnant women and their newborns. This study can contribute towards this noble and necessary endeavour.

References

Adams, G., M. C. Gulliford, O. C. Okoumunne, S. Eldridge, S. Chinn and M. J. Campbell (2004). "Patterns of intra-cluster correlation from primary care research to inform study design and analysis." Journal of Clinical Epidemiology **57**(8): 784-794.

Allen, D., E. Gillen and L. Rixson (2009). "Systematic review of the effectiveness of integrated care pathways : what works, for whom, in which circumstances?" International Journal of Evidence Based Healthcare **7**: 61-74.

Allen, T. D., L. T. Eby and E. Lentz (2006). "Mentorship Behaviors and Mentorship Quality Associated With Formal Mentoring Programs: Closing the Gap Between Research and Practice." Journal of Applied Psychology **91**(3): 567-578.

Althabe, F. (2007). Local opinion leaders : effects on professional practice and health care outcomes : RHL commentary. Geneva, World Health Organization.

Althabe, F., P. Buekens, E. Bergel, J. Belizán, M. Campbell, N. Moss, T. Hartwell, L. Wright and for the Guidelines Trial Group (2008). "A Behavioral Intervention to Improve Obstetrical Care." New England Journal of Medicine **358**(18): 1929-1940.

Anatole, M., H. Magge, V. Redditt, A. Karamaga, S. Niyonzima, P. Drobac, J. S. Mukherjee, J. Ntaganira, L. Nyirazinyoye and L. R. Hirschhorn (2013). "Nurse mentorship to improve the quality of health care delivery in rural Rwanda." Nursing Outlook **61**: 137-144.

Badawi, N., J. Kurinczuk, J. Keogh, L. Alessandri, F. O'Sullivan, P. Burton, P. Pemberton and F. Stanley (1998a). "Antepartum risk factors for newborn encephalopathy : the Western Australian case-control study." British Medical Journal **317**: 1549-1553.

Badawi, N., J. Kurinczuk, J. Keogh, L. Alessandri, F. O'Sullivan, P. Burton, P. Pemberton and F. Stanley (1998b). "Intrapartum risk factors for newborn encephalopathy : the Western Australian case-control study." British Medical Journal **317**: 1554-1558

Baker, R., J. Camosso-Stefinovic, C. Gillies, E. Shaw, F. Cheater, S. Flottorp and N. Robertson (2010). "Tailored interventions to overcome identified barriers to change : effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews **2010**(3): CD005470.

Ban Ki-Moon (2009). Health systems that deliver for women deliver for all. Secretary-General, United Nations: 2.

Baqui, A. H., S. El-Arifeen, G. L. Darmstadt, S. Ahmed, E. K. Williams, H. R. Seraji, I. Mannan, S. M. Rahman, R. Shah, S. K. Saha, U. Syed, P. J. Winch, A. Lefevre, M. Santosham, R. E. Black and for the Projahnmo Study Group (2008). "Effect of community-based newborn-care intervention package implemented through two service delivery strategies in Sylhet district, Bangladesh : a cluster randomised controlled trial." The Lancet **371**: 1936-1944.

Baum, F., C. MacDougall and D. Smith (2006). "Participatory action research." Journal of Epidemiology and Community Health **60**(10): 854-857.

Bhandari, N., R. Bahl, S. Mazumdar, J. Martines, R. Black, M. Bhan and other members of the Infant Feeding Study Group (2003). "Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illness and growth : a cluster randomised controlled trial." The Lancet **361**: 1418-1423.

Bheekie, A. (2001). Pharmacist educational outreach for improved primary care of asthma in children. . PhD University of the Western Cape.

Bhutta, Z., M. Chopra, H. Axelson, P. Berman, T. Boerma, J. Bryce, F. Busteo, E. Cavagnero, G. Cometto, B. Daelmans, A. de Francisco, H. Fogstad, N. Gupta, L. Laski, J. Lawn, B. Mahqi, E. Mason, C. Pitt, J. Requejo, A. Starrs, C. G. Victora and T. Wardlaw (2010). "Countdown to 2015 decade report (2000–10): taking stock of maternal, newborn, and child survival." Lancet **375** 2032 – 2044.

Bhutta, Z., M. Yakoob, J. Lawn, R. Rizvi, I. Friberg, E. Weismann, E. Buchmann, Goldenberg RL. and for The Lancet's Stillbirths Series steering committee (2011). "Stillbirths: what difference can we make and at what cost?" Lancet **377** (9776): 1523-1538.

Bhutta, Z. A., G. L. Darmstadt, R. A. Haws, M. Y. Yakoob and J. E. Lawn (2009). "Delivering interventions to reduce the global burden of stillbirths : improving service supply and community demand." BMC Pregnancy and Childbirth **9 (Suppl 1)**: S7.

Bingham, D. and E. K. Main (2010). "Effective Implementation Strategies and Tactics for Leading Change on Maternity Units." Journal of Perinatal and Neonatal Nursing **24**(1): 32-42.

Bland, J. M. and D. G. Altman (1986). "Statistical methods for assessing agreement between two methods of clinical measurement." The Lancet **i**: 307-310.

Bloom, B. S., Ed. (1956). Taxonomy of Educational Objectives : The classification of educational goals. . London, Longmans, Green & Co.

Bosch-Capblanch, X. and P. Garner (2008). "Primary health care supervision in developing countries." Tropical Medicine and International Health **13**(3): 369-383.

Bosse, G., S. Massawe and A. Jahn (2002). "The partograph in daily practice: it's quality that matters." International Journal of Gynecology and Obstetrics **77**: 243-244.

Bradley, E. L. and L. G. Blackwood (1989). "Comparing paired data : A simultaneous test for means and variances." The American Statistician **43**(4): 234-235.

Bradshaw, D. and R. E. Dorrington (2012). "Maternal mortality ratio - trends in the vital registration data." South African Journal of Obstetrics and Gynaecology **18**(2): 38-42.

Brunero, S. and J. Stein-Parbury (2008). "The effectiveness of clinical supervision in nursing : an evidenced based literature review." The Australian Journal of Advanced Nursing **25**(3): 86-94.

Buus, N. and H. Gonge (2009). "Empirical studies of clinical supervision in psychiatric nursing: A systematic literature review and methodological critique." International Journal of Mental Health Nursing **18**(4): 250-264.

Campbell, M., R. Fitzpatrick, A. Haines, A. L. Kinmonth, P. Sandercock, D. Spiegelhalter and P. Tyrer (2000). "Framework for design and evaluation of complex interventions to improve health." British Medical Journal: 694-696.

Campbell, M. J., A. Donner and N. Klar (2007). "Developments in cluster randomized trials and *Statistics in Medicine*." Statistics in Medicine **26**(1): 2-19.

Campbell, M. K., P. M. Fayers and J. M. Grimshaw (2005). "Determinants of the intracluster correlation coefficient in cluster randomized trials : the case of implementation research." Clinical Trials **2**(2): 99 -107.

Campbell, M. K., G. Piaggio, D. R. Elbourne, D. G. Altman and for the CONSORT Group (2012). "Consort 2010 statement: extension to cluster randomised trials." British Medical Journal **345**: e5661.

Campbell, M. K., N. Steen, J. M. Grimshaw, M. Eccles, J. Mollinson and C. Lombard (1999). Design and statistical issues in implementation research. Changing Professional Practice : Theory and Practice of Clinical Guidelines Implementation. T. Thorsen and M. Makela. Copenhagen, Danish Institute for Health Services Research and Development: 217.

Canadian Institutes of Health Research. (2009). "Knowledge translation definition." Retrieved 14 September 2012, from <http://www.cihr-irsc.gc.ca/e/39158.html#s11>.

Canadian Nurses Association (2004). Achieving excellence in Professional practice - A Guide to Preceptorship and Mentoring. Ottawa, Canadian Nurses Association.

Cassidy, S. (2013). "Acknowledging hubris in interpretive data analysis." Nurse Researcher **20**(6): 27-31.

Cheater, F., R. Baker, C. Gillies, H. Hearnshaw, S. Flottorp, N. Robertson, E. J. Shaw and A. D. Oxman (2005). "Tailored interventions to overcome identified barriers to change : effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews **2005**(3): CD005470.

Cheater, F., R. Baker, S. Reddish, N. Spiers, A. Wailoo, Gillies C., N. Robertson and C. Cawood (2006). "Cluster Randomized Controlled Trial of the Effectiveness of Audit and Feedback and Educational Outreach on Improving Nursing Practice and Patient Outcomes." Medical Care **44**: 542-551.

Chen, C.-M. and M.-F. Lou (2014). "The effectiveness and application of mentorship programmes for recently registered nurses : a systematic review." Journal of Nursing Management **22**(4): 433-442.

Christie, J., P. O'Halloran and M. Stevenson (2009). "Planning a cluster randomized controlled trial : methodological issues." Nursing Research **58**(2): 128-134.

Clark, P., A. Jamieson, J. Launer, A. Trompetas, J. Whiteman and D. Williamson (2006). "Intending to be a supervisor, mentor or coach? which, what for and why?" Education for Primary Care **17**: 109-116.

Clow, S. E. (2005). Clinical decision making in labour for safer birth. 24th Annual Priorities in Perinatal Care conference, Langebaan, South Africa.

Clow, S. E. (2006). Clinical decision making in labour – how do “part time” midwives fare? 25th annual Priorities in Perinatal Care conference, Champagne Sports Resort, Drakensberg, South Africa.

Clow, S. E. (2012). Modifying and testing an audit instrument for partograph evaluation 31st annual Priorities in Perinatal Care Congress, Kruger Gate, Mpumalanga, South Africa, Priorities in Perinatal Care Association.

Cochrane, L., C. Olson, S. Murray, M. Cupuis, T. Tooman and S. Hayes (2007). "Gaps between knowing and doing : understanding and assessing the barriers to optimal health." Journal of Continuing Education for the Health Professions **27**(2): 94 - 102.

Colton, D. (2000). "Quality improvement in health care. Conceptual and historical foundations." Evaluation & the Health Professions **23**(1): 7-42.

Costa, M. L. B. and J. Khanna (2008). Educational outreach visits : effects on professional practice and health-care outcomes : RHL Commentary. Geneva, World Health Organization.

Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO), Ed. (2002). International Ethical Guidelines for Biomedical Research Involving Human Subjects. Geneva, CIOMS.

Cowan, F., M. Rutherford, F. Groenendaal, P. Eken, E. Mercuri, G. Bydder, L. Meiners, L. Dubowitz and L. deVries (2003). "Origin and timing of brain lesions in term infants with neonatal encephalopathy." The Lancet **361**: 736-742.

Darmstadt, G. L., M. Y. Yakoob, R. A. Haws, V. Menezes Esme, T. Soomro and Z. A. Bhutta (2009). "Reducing stillbirths : interventions during labour." BMC Pregnancy and Childbirth **9 (Suppl I): S6**.

Davies, B., E. Hodnett, M. Hannah, L. O'Brien-Pallas, D. Pringle, G. Wells, members of the Perinatal Partnership Program of Eastern and Southeastern Ontario and Society of Obstetricians and Gynecologists of Canada (2002). "Fetal health surveillance : a community-wide approach versus a tailored intervention for the implementation of clinical practice guidelines." Canadian Medical Association Journal **167**(5): 469-474.

Davis, A., A.-M. Bergh, E. van Rooyen and R. Pattinson (2006). Ukubamba Umtwana kuwe outreach : Implementation of Kangaroo Mother Care in Mpumalanga 25th National Priorities in Perinatal Care, Champagne Sports resort, Drakensberg, South Africa.

Department of Health (1999). Saving Mothers : Report on Confidential enquiries into Maternal Deaths in South Africa 1998. Pretoria, Department of Health

Department of Health (2000). Guidelines for Maternity Care in South Africa : A manual for clinics, community health centres and district hospitals. Pretoria, Department of Health.

Department of Health (2002). Saving Mothers 1999-2001 : Second Report on Confidential enquiries into Maternal Deaths in South Africa. Pretoria, Department of Health.

Department of Health (2004). Ethics in health research : principles, structures and processes. Pretoria, Department of Health.

Department of Health (2006a). Saving Mothers 2002-2004 : Third report on Confidential Enquiries into Maternal Deaths in South Africa. Pretoria, Department of Health.

Department of Health (2006b). Guidelines for Good Practice in the Conduct of Clinical Trials with Human Participants in South Africa 2nd edition. Pretoria, Department of Health.

Department of Health (2007). Guidelines for maternity care in South Africa 2007 : A manual for clinics, community health centres and district hospitals. Pretoria, Department of Health.

Department of Health (2009). Saving Mothers 2005-2007 : Fourth report on Confidential Enquiries into Maternal Deaths in South Africa. Pretoria, Department of Health.

Department of Health (2012). Saving Mothers 2008-2010 : Fifth report on the Confidential Enquiries into maternal deaths in South Africa - short report. Pretoria, Department of Health.

Dick, J., S. Lewin, E. Rose, M. Zwarenstein and H. van der Walt (2004). "Changing professional practice in tuberculosis care : an educational intervention." Journal of Advanced Nursing **48**(5): 434-442.

Dilworth, S., I. Higgins, V. Parker, B. Kelly and J. Turner (2013). "Finding a way forward: A literature review on the current debates around clinical supervision." Contemporary Nurse **45**(1): 22-32.

Doumit, G., M. Gattellari, J. Grimshaw and M. A. O'Brien (2007). "Local opinion leaders: effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews **2007**(4): CD000125.

Dowling, M. (2006). "Approaches to reflexivity in qualitative research." Nurse Researcher **13**(3): 7-21.

Dujardin, B., I. de Schamphelre, H. Sene and F. Ndiaye (1992). "Value of the alert and action lines on the partogram." The Lancet **339**: 1336-1338.

Elashoff, J. D. (2000). nQuery Advisor nQuery Advisor Version 4.0 Users' guide. Los Angeles, Statistical Solutions Ltd.

Elbourne, D. R. and M. K. Campbell (2001). "Extending the CONSORT statement to cluster randomized trials : for discussion." Statistics in Medicine **20**(3): 489-496.

Ellis, M., N. Manandhar, D. S. Manandhar and A. Costello (2000). "Risk factors for neonatal encephalopathy in Kathmandu, Nepal, a developing country : unmatched case-control study." British Medical Journal **320**: 1229-1236.

Eriksson, L., N. Nga, D. Hoa, L. Persson, U. Ewald and L. Wallin (2011). "Newborn care and knowledge translation - perceptions among primary health care staff in northern Vietnam." Implementation Science **6**: 29.

Estabrooks, C., J. Floyd, S. Scott-Findlay, K. O'Leary and M. Gushta (2003). "Individual determinants of research utilization : A systematic review." Journal of Advanced Nursing **43**: 506 - 520.

European Pathway Association. (2007). "Definition of Integrated Care Pathway." Retrieved 22 July 2012, from <http://www.e-p-a.org/index2.html>.

Fahdhy, M. and V. Chongsuvivatwong (2005). "Evaluation of World Health Organization partograph implementation by midwives for maternity home birth in Medan, Indonesia." Midwifery **21**: 301-310.

Fairall, L. F., M. Zwarenstein, E. D. Bateman, M. Bachmann, C. Lombard, B. P. Majara, G. Joubert, R. G. English, A. Bheekie, D. van Rensburg, P. Mayers, Peters Annatjie C. and R. D. Chapman (2005). "Effect of educational outreach to nurses on tuberculosis case detection and primary care of respiratory illness: pragmatic cluster randomised controlled trial." British Medical Journal **331**: 750-754.

Farrell, E.-M. (2005). Labour in the African woman - is there a difference? Intrapartum care in South Africa : Review and guidelines. E.-M. Farrell and R. Pattinson. Pretoria, Medical Research Council.

Fawcus, S., D. Kennedy and M. Kroon (2012). "The effect of maternal HIV status on perinatal outcome at Mowbray Maternity Hospital and referring midwife obstetric units, Cape Town " South African Journal of Obstetrics and Gynaecology **18**(1): 6-10.

Fawcus, S., H. van Zyl, L. Linley, C. Thomson and D. Greenfield (2004). Obstetric risk factors for neonatal encephalopathy. 23rd Annual Priorities in Perinatal Care conference, Limpopo, South Africa.

Fesseha, N., A. Getachew, M. Hiluf, Y. Gebrehiwot and P. Bailey (2011). "A national review of cesarean delivery in Ethiopia." International Journal of Gynecology and Obstetrics **115**: 106-111.

Finlay, L. (2002). "Negotiating the swamp : the opportunity and challenge of reflexivity in research practice." Qualitative Research **2**(2): 209-230.

Fisher, M. and C. Webb (2008). "What do midwifery mentors need? Priorities and impact of experience and qualifications." Learning in Health and Social Care **8**(1): 33-46.

Fitzpatrick, R., C. Davey, M. Buxton and D. Jones (1998). "Evaluating patient based outcome measures for use in clinical trials." Health Technology Assessment **2**(14): 1-74.

Flodgren, G., E. Parmelli, G. Doumit, M. Gattellari, M. O'Brien, J. Grimshaw and M. Eccles (2011). "Local opinion leaders : effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews **2011**(8): CD000125.

Flodgren, G., M. Rojas-Reyes, N. Cole and Foxcroft DR. (2012). "Effectiveness of organisation infrastructures to promote evidence-based nursing practice " Cochrane Database of Systematic Reviews **2012**(2): CD002212.

Fontana, J. S. (2004). "A methodology for critical science." Advances in Nursing Science **27**(2): 93-101.

Forsetlund, L., A. Bjorndal, A. Rashidian, G. Jamtvedt, M. O'Brien, F. Wolf, D. Davis, J. Odgaard-Jensen and A. Oxman (2009). "Continuing education meetings and workshops : effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews **2009**(2): CD003030.

Fox, C. E. and K. S. Khan (2010). Tailored interventions to overcome identified barriers of change : effects on professional practice and health care outcomes : RHL Commentary. Geneva, World Health Organization.

Foy, R., M. Eccles and J. Grimshaw (2001). "Why does primary care need more implementation?" Family Practice **18**(4): 353-355.

Francke, A. L. and F. M. de Graaff (2012). "The effects of group supervision of nurses : A systematic literature review." International Journal of Nursing Studies **49**(9): 1165-1179.

Fraser, D. M. and M. A. Cooper (2003). Myles Textbook for Midwives. Edinburgh, Churchill Livingstone.

Freire, P. (1970). Pedagogy of the oppressed revised edition. London, Penguin

Fullerton, J. T., J. B. Thompson and P. Johnson (2013). "Competency-based education: The essential basis of pre-service education for the professional midwifery workforce." Midwifery **29**(10): 1129-1136.

Gallo, A., C. Weijer, A. White, J. M. Grimshaw, R. Boruch, J. C. Brehaut, A. Donner, M. P. Eccles, A. D. McRae, R. Saginur, M. Zwarenstein and M. Taljaard (2012) "What is the role and authority of gatekeepers in cluster randomized trials in health research?" Trials **13**, 116 DOI: 10.1186/1745-6215-13-116.

Garenne, M. and R. McCaa (2010). "Maternal mortality for 181 countries, 1980-2008." The Lancet **376**: 1389.

Gauge, S. M. and C. Henderson (1992). CTG made easy. Edinburgh, Churchill Livingstone.

Government Communication and Information System (GCIS) (2008). South Africa Year Book 2007/2008. Pretoria, South African Government Communication and Information System.

Graham, I. D., J. Logan, M. B. Harrison, S. E. Straus, J. Tetroe, W. Caswell and N. Robinson (2006). "Lost in Knowledge translation : Time for a map?" The Journal of Continuing Education in the Health Professions **26**(1): 13 -24.

Graham, W. and J. Hussein (2004). "The right to count." The Lancet **363**: 67-68.

Graham, W. J., V. G. A. Fillippi and C. Ronsman (1996). "Demonstrating programme impact on maternal mortality." Health policy and planning **11**(1): 16-20.

Graham, W. J. and M.-L. Newell (2009). Maternal deaths in South Africa : What proportion occur at home? Saving Mothers : Fourth report on Confidential enquiries into maternal deaths in South Africa 2005-2007. R. C. e. Pattinson. Pretoria, Department of Health.

Graham, W. J., P. T. Wagaarachchi, G. C. Penney, A. M. Binns, K. Y. A. Antwi and M. H. Hall (2000). "Criteria for clinical audit of the quality of hospital based obstetric care in developing countries." Bulletin of the World Health Organization **78**(5): 614-620.

Greenfield, D., S. Arends, L. Olivier and H. Odendaal (2005). Perinatal audit using PPIP in the Western Cape : Implications for service delivery. 24th annual Priorities in Perinatal Care conference, Langebaan, South Africa.

Greenfield, D., N. Rhoda and R. C. Pattinson (2011). Chapter 2 - Ten years of the national perinatal care surveys. Saving Babies 2008-2009 : Seventh report on perinatal care in South Africa. R. C. e. Pattinson. Pretoria, Tshepesa Press: 42-59.

Grimshaw, J., L. M. McAuley, L. A. Bero, R. Grilli, A. D. Oxman, C. Ramsay, L. Vale and M. Zwarenstein (2003). "Systematic reviews of the effectiveness of quality improvement strategies and programmes." Quality and Safety in Health Care **12**(4): 298-303.

Grimshaw, J., R. Thomas, G. MacLennan, C. Fraser, C. Ramsay, L. Vale, P. Whitty, M. Eccles, L. Matowe, L. Shirran, M. Wensing, R. Dijkstra and C. Donaldson (2004). "Effectiveness and efficiency of guidelines dissemination and implementation strategies." Health Technology Assessment **8**(6).

Groenewald-Neethling, S. (2010). Implementation of the Basic Antenatal Care Package in primary health care clinics : A comparison of two training methods : face to face and training of trainers. MSc(Nursing), University of Cape Town.

Groeschel, N. and P. Glover (2001). "The partograph. Used daily but rarely questioned." Australian Journal of Midwifery **14**(3): 22-26.

Grol, R., R. Baker and F. Moss (2002). "Quality improvement research : understanding the science of change in health care." Quality and Safety in Health Care **11**: 110-111.

Grol, R. and J. Grimshaw (2003). "From best evidence to best practice : effective implementation of change in patients' care." The Lancet **362**: 1225-1230.

Grosskurth, H., F. Mosha, J. Todd, E. Mwijarubi, A. Klokke, K. Senkoro, P. Mayaud, J. Changalucha, A. Nicoll, G. ka-Gina, J. Newell, K. Mugeye, D. Mabey and R. Hayes (1995). "Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania : randomised controlled trial." The Lancet **346**: 530-536.

Grove, S. K., N. Burns and J. Gray (2013). The Practice of Nursing Research : Appraisal, Synthesis, and Generation of Evidence. St Louis, Elsevier/Saunders.

Gülmezoglu, A., A. Langer, G. Piaggio, P. Lubiganon, J. Villar and J. Grimshaw (2007). "Cluster randomised trial of an active, multifaceted educational intervention based on the WHO Reproductive Health Library to improve obstetric practices." British Journal of Obstetrics and Gynaecology **114**: 16-23.

Gülmezoglu, A. M., J. Villar, J. Grimshaw, G. Piaggio, P. Lumbiganon and A. Langer (2004). "Cluster randomized trial of an active, multifaceted information dissemination intervention based on The WHO Reproductive health library to change obstetric practices: methods and design issues." BMC Medical Research Methodology **4**(2).

Haxton, J. and K. Fahy (2009). "Reducing length of stay for women who present as outpatients to delivery suite : A clinical practice improvement project." Women and Birth **22**(4): 119-127.

Hayes, R. J., J. Changaluchab, D. A. Rossa, A. Gavyole, J. Todd, A. I. N. Obasi, M. L. Plummer, D. Wight, D. C. Mabeya and H. Grosskurth (2005). "The MEMA kwa Vijana Project: Design of a community randomised trial of an innovative adolescent sexual health intervention in rural Tanzania." Contemporary clinical trials **26**: 430-442.

Hayes, R. J. and L. H. Moulton (2009). Cluster randomised trials. Boca Raton, Chapman & Hall / CRC Press.

Headrick, L. (2000). Learning to improve complex systems of care. Washington DC, HRSA/Bureau of Health Professions.

Heritier, S. R., V. R. Gebski and Keech Anthony C. (2003). "Inclusion of patients in clinical trial analysis : the intention-to-treat principle." Medical Journal of Australia **179**: 438-440.

Hesketh, T., W. Zhu and K. Zheng (1994). "Improvement of neonatal care in Zhejiang province, China, through a self-instructional continuing education programme." Medical Educator **28**: 252-259.

Higgins, J. P. T. and S. Green (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0, The Cochrane Collaboration.

Hill, K., C. AbouZhar and T. Wardlaw (2001). "Estimates of maternal mortality for 1995." Bulletin of the World Health Organization **79**(3): 182-193.

Hogan, M. C., K. J. Foreman, M. Naghavi, S. Y. Ahn, M. Wang, S. M. Makela, A. D. Lopez, R. Lozano and C. J. L. Murray (2010). "Maternal mortality for 181 countries, 1980–2008 : a systematic analysis of progress towards Millennium Development Goal 5." The Lancet **375**(9726): 1609-1623.

Horbar, J. D., J. H. Carpenter, J. Buzas, R. F. Soll, G. Suresh, M. B. Bracken, L. C. Leviton, P. E. Plsek and J. C. Sinclair (2004). "Collaborative quality improvement to promote evidence based surfactant for preterm infants: a cluster randomised trial." British Medical Journal **329**(7473).

Horsley, T., C. Hyde, N. Santesso, J. Parkes, R. Milne and R. Stewart (2011). "Teaching critical appraisal skills in health care settings (review)." The Cochrane Database of Systematic Reviews(11): CD001270.

Hotopf, M., R. Churchill and G. Lewis (1999). "Pragmatic randomised control trials in psychiatry." British Journal of Psychiatry **175**: 217-233.

International Confederation of Midwives. (2002). "Essential Competencies for basic Midwifery practice." Retrieved 7 April, 2006.

International Confederation of Midwives (2011a). Essential Competencies for Basic Midwifery Practice. The Hague, The International Confederation of Midwives.

International Confederation of Midwives (2011b). Standards for Midwifery Education. The Hague, International Confederation of Midwives.

International Council of Nurses (2012). ICN Code of Ethics for Nurses. Geneva, International Council of Nurses.

Ith, P., A. Dawson and C. Homer (2012). "Quality of maternity care practices of skilled birth attendants in Cambodia." International Journal of Evidence-Based Healthcare **10**(1): 60-67.

Ivers, N., G. Jamtvedt, S. Flottrop, J. Young, J. Odgaard-Jensen, S. French, M. O'Brien, M. Johansen, J. Grimshaw and A. Oxman (2012). "Audit and feedback: effects on professional practice and healthcare outcomes." Cochrane Database of Systematic Reviews **2012**(6): CD000259.

Jackson, C., F. M. Cheater, W. Harrison, R. Peacock, H. Bekker, R. West and B. Leese (2011) "Randomised cluster trial to support informed parental decision-making for the MMR vaccine." BMC Public Health **11** DOI: 10.1186/1471-2458-11-475.

Jamtvedt, G., J. M. Young, D. T. Kristoffersen, M. A. O'Brien and A. D. Oxman (2006). "Does telling people what they have been doing change what they do? A systematic review of audit and feedback." Quality and Safety in Health Care **15**(6): 433-436.

Jamtvedt, G., J. M. Young, D. T. Kristoffersen, M. A. O'Brien and A. D. Oxman (2007). "Audit and feedback : effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews **4**: CD000259.

Kattwinkel, J., L. Cook, G. Novacek, H. Hallam and J. Short (1979). "Improved perinatal knowledge and care in the community hospital through a programme of self-instruction." Pediatrics **64**(4): 451-458.

Katzenellenbogen, J. M., G. Joubert and S. S. Abdool Karim (1997). Epidemiology : A manual for South Africa. Cape Town, Oxford University Press.

Kemmis, S. (2007). Critical Theory and Participatory Action Research. The SAGE Handbook of Action Research : Participative Inquiry and Practice 2nd edition. H. Bradbury and P. Reason. Los Angeles, SAGE Publishers: 752.

Kerry, S. M. and J. M. Bland (1998). "Sample size in cluster randomisation." British Medical Journal **316**: 549.

Kirshbaum, M. (2008). "Translation to practice : a randomised, controlled study of an evidence based booklet for breast care nurses in the United Kingdom." Worldviews on Evidence Based Nursing **5**(2): 60-74.

Koethe, J. R., A. O. Westfall, D. K. Luhanga, G. M. Clark, J. D. Goldman, P. L. Mulenga, R. A. Cantrell, B. H. Chi, I. Zulu, M. S. Saag and J. S. A. Stringer (2010). "A Cluster Randomized Trial of Routine HIV-1 Viral Load Monitoring in Zambia: Study Design, Implementation, and Baseline Cohort Characteristics." PlosOne **5**(3).

Konteh, F., R. Mannion and H. T. Davies (2008). "Clinical governance views on culture and quality improvement." Clinical Governance : An International Journal **16**(4): 337-352.

Kwast, B. E., C. E. Lennox and T. M. M. Farley. (1994). "World Health Organization partograph in management of labour." The Lancet **343**(8910): 1399-1404.

Kwast, B. E., P. Poovan, E. Vera and E. Kohls (2008). "The modified WHO partograph : do we need a latent phase?" African Journal of Midwifery and Women's Health **2**(3): 143-149

Kwast personal communication (2014). Partograph utilisation checklist. S. Clow: 1.

Langley, G., K. Nolan, T. Nolan, C. Norman and L. Provost (2009). The Improvement Guide: A Practical Approach to Enhancing Organizational Performance (2nd edition). San Francisco, Jossey-Bass Publishers.

Latham, C. L., M. Hogan and K. Ringl (2008). "Nurses Supporting Nurses : Creating a mentoring program for staff nurses to improve the workforce environment." Nursing Administration Quarterly **32**(1): 27-39.

Lavender personal communication (2014). Partograph utilisation instrument. S. Clow: 1.

Lavender, T., Z. Alfirevic and S. Walkinshaw (1998). "Partogram action line study : a randomised trial." British Journal of Obstetrics and Gynaecology **105**: 976-980.

Lavender, T., A. Hart and R. M. Smyth (2008). "Effect of partogram use on outcomes for women in spontaneous labour at term." Cochrane Database of Systematic Reviews **8**(4).

Lavender, T. and I. Malcolmson (1999). " Is the partogram a help or a hindrance?" Practising Midwife **2**: 23-27.

Lawn, J. and K. Kerber, Eds. (2006). Opportunities for Africa's newborns: Practical data, policy and programmatic support for newborn care in Africa. Cape Town, Partnership for Maternal, Newborn and Child Health (PMNCH).

Lawn, J. E., H. Blencowe, R. C. Pattinson, S. Cousens, R. Kumar, I. Ibiebele, J. Gardosi, L. T. Day, C. Stanton and for the Lancet's Stillbirth Series standing committee (2011). "Stillbirths : Where? When? Why? How to make the data count?" The Lancet **377**: 1448-1463.

Lawn, J. E., S. Cousens and J. Zupan (2005). "4 million neonatal deaths : When? Where? Why?" Lancet **365**(9462): 891-900.

Lawn, J. E., M. Y. Yakoob, R. A. Haws, T. Soomro, G. L. Darmstadt and Z. A. Bhutta (2009). "3.2 million stillbirths : epidemiology and overview of the evidence review." Biomed Central Pregnancy and Childbirth **9 (Suppl) 1(S2)**.

le Roux, E., R. Pattinson, W. Tsaku and J. Makin (1998). "Does successful completion of the Perinatal Education Programme result in improved obstetric practice?" South African Medical Journal **88(2)**: 180-187.

Lewycka, S., C. Mwansambo, P. Kazembe, T. Phiri, A. Mganga, M. Rosato, H. Chapota, H. Malamba, S. Vergnano, M.-L. Newell, D. Osrin and A. Costello (2010). "A cluster randomised controlled trial of the community effectiveness of two interventions in rural Malawi to improve health care and to reduce maternal, newborn and infant mortality." Trials **11**: 88.

Lin, L. I.-K. (1989). "A Concordance Correlation Coefficient to Evaluate Reproducibility." Biometrics **45(1)**: 255-268.

Lin, L. I.-K. (2000). "A note on the Concordance Correlation Coefficient (Letter)." Biometrics **56**: 324-325.

Lipp, A. (2007). "Developing the reflexive dimension of reflection : a framework for debate." International Journal of Multiple Research Approaches **1(1)**: 18-26.

Lomas, J., M. Enkin, G. Anderson, W. Hannah, E. Vayda and J. Singer (1991). "Opinion leaders vs audit and feedback to implement practice guidelines." Journal of the American Medical Association **265(17)**: 2202-2207

Lynn, M. R. (1986). "Determination and quantification of content validity." Nursing Research **35(6)**: 382-385.

Macdonald, A. P., G. Bartlett, E. J. Buchmann and A.-M. Goosen (2007). Intrapartum challenges and recommendations. Saving Babies V 2003-2005 : Fifth perinatal care survey. R. C. Pattinson. Pretoria, University of Pretoria, MRC, CDC: 41-58.

- Manandhar, D. S., D. Osrin, B. P. Shrestha, N. Mesko, J. Morrison, K. M. Tumbahangphe, S. Tamang, S. Thapa, D. Shrestha, B. Thapa, J. R. Shrestha, A. Wade, J. Borghi, H. Standing, M. Manandhar, A. Costello and members of the MIRA Makwanpur trial team (2004). "Effect of a participatory intervention with women's groups on birth outcomes in Nepal : cluster-randomised controlled trial." The Lancet **2004**(363): 970-979.
- Mancey-Jones, M. and R. F. Brugha (1997). "Using perinatal audit to promote change : a review." Health Policy and Planning **12**(3): 183-192.
- Mannion, R., H. Davies and Marshall MN. (2005). "Cultural characteristics of "high" and "low" performing hospitals." Journal of Health Organisation and Management **19**(6): 431-439.
- Mannion, R., F. Konteh and H. Davies (2009). "Assessing organisational culture for quality and safety improvement : a national survey of tools and tool use." Quality and Safety in Health Care **18**(4): 153-156.
- Marshall, J. E., D. M. Fraser and P. N. Baker (2010). "Reflexivity : the experience of undertaking an ethnographic study to explore issues of consent to intrapartum procedures." Evidence Based Midwifery **8**(1): 21-25.
- Marshall, M., A. Lockwood, C. Bradley, C. Adams, C. Joy and M. Fenton (2000). "Unpublished rating scales : A major source of bias in randomised control trials of treatments for schizophrenia." British Journal of Psychiatry **176**: 249-252.
- Mathai, M. (2009). "The partograph for prevention of obstructed labour." Clinical Obstetrics and Gynecology **52**(2): 256-269.
- Mathews, J. E., A. Rajaratnam, A. George and M. Mathai (2007). "Comparison of two World Health Organization partographs." International Journal of Gynecology and Obstetrics **96**: 147-150.

McCabe, J. and D. Holmes (2009). "Reflexivity, critical qualitative research and emancipation: a Foucauldian perspective " Journal of Advanced Nursing **65**(7): 1518–1526.

McCloughen, A., L. O' Brien and D. Jackson (2011). "Nurse Leader Mentor as a Mode of Being : Findings From an Australian Hermeneutic Phenomenological Study." Journal of Nursing Scholarship **43**(1): 97-104.

McRae, A. D., C. Weijer, A. Binik, J. M. Grimshaw, R. Boruch, J. C. Brehaut, A. Donner, M. P. Eccles, R. Saginur, A. White and M. Taljaard (2011) "When is informed consent required in cluster randomized trials in health research?" Trials **12** DOI: 10.1186/1745-6215-12-202.

McTaggart, R., Ed. (1997). Participatory Action Research : International Contexts and Consequences. Albany, State University of New York.

Medical Research Council Unit for Maternal and Infant Health Care Strategies (2006). Intrapartum care in South Africa : review and guidelines. E.-M. Farrell and R. C. Pattinson. Pretoria, University of Pretoria.

Menon, A., N. Korner-Bitensky, M. Kastner, K. McKibbin and S. Straus (2009). "Strategies for Rehabilitation Professionals to move Evidence-Based Knowledge into Practice: a Systematic Review." Journal of Rehabilitation Medicine **41**(13): 1024-1032.

Mercer, S. W., K. Sevar and T. D. Sadautshan (2006). "Using clinical audit to improve quality of obstetric care at the Tibetan Delek Hospital in North India " Reproductive Health **3**: 4.

Michie, S., M. M. van Stralen and R. West (2011). "The behaviour change wheel: A new method for characterising and designing behaviour change interventions." Implementation Science **6**: 11.

Mijares, L., S. M. Baxley and M. L. Bond (2013). "Mentoring : A Concept Analysis." The Journal of Theory Construction and Testing **17**(1): 23-28.

Miller, S., N. Sloan, B. Winnikoff, A. Langer and F. Fikree (2003). "Where is the E in MCH? The need for an evidence based approach in safe motherhood." Journal of Midwifery and Women's Health **48**: 10-18.

Mills, J., J. Field and R. Kant (2011). "Factors affecting evidence translation for general practice nurses." International Journal of Nursing Practice **17**: 455 - 463.

Milne, D. L., A. I. Sheikh, S. Pattison and A. Wilkinson (2011). "Evidence-Based Training for Clinical Supervisors : A Systematic Review of 11 Controlled Studies." The Clinical Supervisor **30**: 53-71.

Milne, L., G. Scotland, N. Tagiyeva-Milne and J. Hussein (2004). "Safe Motherhood Program Evaluation : Theory and Practice." Journal of Midwifery and Women's Health **49**(4): 338-344.

Minister of Health (1999). Speech to launch the Saving Mothers report : report on the Confidential Enquiry into Maternal Deaths in South Africa 1998.

Mitchell, S. A., C. A. Fisher, C. E. Hastings, L. B. Silverman and G. R. Wallen (2010). "A thematic analysis of theoretical models for translational science in nursing: Mapping the field." Nursing Outlook **58**(6): 287 - 300.

Moher, D., S. Hopewell, K. F. Schulz, V. Montori, P. C. Gøtzsche, P. J. Devereaux, D. Elbourne, M. Egger and D. G. Altman (2010). "CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials." British Medical Journal **340**: c869.

More, N. S., U. Bapat, S. Das, S. Patil, M. Porel, L. Vaidya, B. Koriya, S. Barnett, A. Costello, A. Fernandez and D. Osrin (2008) "Cluster-randomised controlled trial of community mobilisation in Mumbai slums to improve care during pregnancy, delivery, postpartum and for the newborn." Trials **9** DOI: 10.1186/1745-6215-9-7.

Muir, N. (2004). "Clinical decision-making : theory and practice." Nursing Standard **18**(36): 47-52.

Murray, D. M. (2001). "Statistical models appropriate for designs often used in group randomized trials." Statistics in Medicine **20**: 1373-1385.

New South Wales Health Department (2002). Easy Guide to Clinical Practice Improvement : Guide for Health Care Professionals Sydney, New South Wales Health Department.

NHS Scotland. (2007). "Managing clinical effectiveness." Retrieved 29 July 2012, from <http://www.clinicalgovernance.scot.nhs.uk/section2/pdsa.asp>, .

Nicholl, M. and M. Cattell (2006). "Getting evidence into Obstetric and Midwifery practice: reducing perineal trauma." Australian Health Review **30**(4): 462-467.

Nursing and Midwifery Council (2008). Standards to support learning and assessment in practice. NMC standards for mentors, practice teachers and teachers. London, Nursing and Midwifery Council.

Nyamtema, A. S., D. P. Urassa, A. Massawe, G. Lindmark and J. van Roosmalen (2008). "Partogram use in the Dar es Salaam perinatal care study." International Journal of Gynecology and Obstetrics **100**(1): 37-40.

O'Brien, M. A., S. Rogers, G. Jamtvedt, A. D. Oxman, J. Ordgaard-Jensen, D. T. Kristoffersen, L. Forsetlund, D. Bainbridge, N. Freemantle, D. A. Davis, R. B. Haynes and E. L. Harvey (2007). "Educational outreach visits : effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews **2007**(4): CD000409.

Ogrinc, G., S. E. Mooney, C. Estrada, T. Foster, D. Goldmann, L. W. Hall, M. M. Huizinga, S. K. Liu, P. Mills, J. Neily, W. Nelson, P. J. Pronovost, L. Provost, L. V. Rubenstein, T. Speroff, M. Splaine, R. Thomson, A. M. Tomolo and B. Watts (2008). "The SQUIRE (Standards for Quality Improvement Reporting Excellence) guidelines for quality improvement reporting : explanation and elaboration." Qual Saf Health Care **17** (Suppl 1): i13-i32.

Ogwang, S., Z. Karyabakabo and E. Rutebemberwa (2009). "Assessment of partogram use during labour in Rujumbura Health Sub District, Rukungiri District, Uganda." African Health Sciences **9**(Special issue 1): S27-34.

Osungabade, K. O., S. A. Oginni, E. A. A. Olumide and E. T. Owoaje (2010). "Clinical audit of intra-partum care at secondary health facilities in Nigeria." Nigerian Journal of Clinical Practice **13**(2).

Oxman, A., M. Thomson, D. Davis and R. Haynes (1995). "No magic bullets : a systematic review of 102 trials of interventions to improve professional practice." Canadian Medical Association Journal **153**: 1423-1431.

Padayatchi, N., K. Naidoo, H. Dawood, A. B. M. Kharsany and Q. Abdool Karim (2011). A review of progress on HIV, AIDS and Tuberculosis. The South African Health Review, 2010. S. Fonn and A. Padarath, Health Systems Trust: 87-100.

Page, C. K. (1999). "Performance Improvement Integration: A Whole Systems Approach." Journal of Nursing Care Quality **13**(3): 59-70.

Parmelli, E., G. Flodgren, M. Shaafsma, N. Baillie, F. Beyer and M. Eccles (2011). "The effectiveness of strategies to change organisational culture to improve health care performance." Cochrane Database of Systematic Reviews **2011**(1): CD008315.

Pattinson, R., K. Kerber, E. Buchmann, I. Friberg, M. Belizan, S. Lansky, E. Weissman, M. Mathai, I. Rudan, N. Walker, J. E. Lawn and for The Lancet's Stillbirths Series steering committee (2011). "Stillbirths: how can health systems deliver for mothers and babies?" Lancet.

Pattinson, R., K. Kerber, P. Waiswa, L. T. Day, F. Mussell, S. Asiruddin, H. Blencowe and J. E. Lawn (2009). "Perinatal mortality audit: Counting, accountability, and overcoming challenges in scaling up in low- and middle-income countries." International Journal of Gynecology and Obstetrics **107**: S113–S122.

Pattinson, R. C., Ed. (2003). Saving Babies 2002 : Third Perinatal Care Survey of South Africa. Pretoria, Government Printer.

Pattinson, R. C., Ed. (2007). Saving Babies 2003-2005 : Fifth perinatal care survey of South Africa. Pretoria, University of Pretoria, Medical Research Council, Centre for Disease Control.

Pattinson, R. C., Ed. (2011). Saving Babies 2008-2009 : Seventh report on perinatal care in South Africa. A decade of Perinatal Care Surveys. Pretoria, Medical Research Council Unit for Maternal and Infant Health Care Strategies.

Pawson, R. and N. Tilley (1997). Realistic Evaluation. London, Sage Publications.

Penn-Kekana, L., B. McPake and J. Parkhurst (2007). "Improving Maternal Health : Getting What Works To Happen." Reproductive Health Matters **15**(30): 28-37.

Pettersson, K. O., M.-L. Svensson and K. Christensson (2000). "Evaluation of an adapted model of the World Health Organization partograph used by Angolan midwives in a peripheral delivery unit." Midwifery **16**: 82-88.

Petzold, A., N. Korner-Bitkensky and A. Menon (2010). "Using the knowledge to action process to incite clinical change." Journal of Continuing Education in the Health Professions **30**(3): 167-171.

Pfund, C., S. House, P. Asquith, M. F. Fleming, K. A. Buhr, E. L. Burnham, J. M. Eichenberger Gilmore, W. C. Huskins, R. McGee, K. Schurr, E. D. Shapiro, K. C. Spencer and C. A. Sorkness (2014). "Training mentors of clinical and translational research scholars." Academic Medicine **89**(5): 774-782.

Philpott, R. H. and W. M. Castle (1972a). "Cervicographs and the management of labour in primigravidae : I The alert line for detecting abnormal labour." Journal of Obstetrics and Gynaecology of the British Commonwealth **79**: 592-598.

Philpott, R. H. and W. M. Castle (1972b). "Cervicographs in the management of labour in primigravidae: II The action line and treatment of abnormal labour." The Journal of Obstetrics and Gynaecology of the British Commonwealth **79**: 599-602.

Philpott, R. H. and A. Voce (2005). Quality check of labour records. Intrapartum care in South Africa : Review and guidelines. E.-M. Farrell and R. C. Pattinson. Pretoria, Medical Research Council: 95.

Pick, W., M. Hoffman, J. Myers, J. Klopper and D. Cooper (1990). "Women, health and urbanisation." South African Medical Journal **77**: 553-554.

Pitchforth, E., R. J. Lilford, Y. Kebede, G. Asres, C. Stanford and J. Frost (2010). "Assessing and understanding quality of care in a labour ward : a pilot study combining clinical and social science perspectives in Gondar, Ethiopia." Social Science and Medicine **71**(10): 1739-1748.

Pitman, E. G. (1939). "A note on Normal Correlation." Biometrika **31**: 9-12.

Prochaska, J. O. and C. C. DiClemente (1982). "Transtheoretical therapy: Toward a more integrative model of change. ." Psychotherapy : Theory, Research and Practice **19**(3): 276-288.

Provincial Government of the Western Cape Department of Health Directorate of Information Management personal communication (2012). Total deliveries August 2006 - September 2007. S. Clow. Cape Town.

Quinn, F. (1988). The principles and practice of Nursing Education. London, Chapman & Hall

Rajamani, A., S. Suen, D. Phillips and M. Thomson (2011). "The SCRIPT project : a knowledge translation approach to improve prescription practice in an intensive care unit." Critical Care and Resuscitation **13**(4): 245-251.

Raven, J., R. Tolhurst, S. Tang and N. van den Broek (2011) "What is quality in maternal and neonatal health care?" Midwifery DOI: 10.1016/j.midw.2011.09.003.

Reid, T. P., K. A. Hinderer, J. M. Jarosinski, B. J. Mister and L. A. Seldomridge (2013). "Expert clinician to clinical teacher: Developing a faculty academy and mentoring initiative." Nurse Education in Practice **13**: 288-293.

Rhoda personal communication (2012). Extract of PPIP data for Western Cape 2009-2011. S. Clow. Cape Town: email correspondence.

Rogers, C. (1969). Freedom to learn. Columbus, Ohio, Charles E Merrill Publishing Company

Rollnick, S., C. Butler and N. Stott (1997). "Helping smokers make decisions : the enhancement of brief intervention for general medical practice." Patient Education and Counselling **31**: 191-203.

Rollnick, S., P. Mason and C. Butler (1999). Health Behavior change : a guide for practitioners. Edinburgh, Churchill Livingstone.

Rotich, E., L. Maina, A. Njihia and K. Christensson (2011). "Evaluating partograph use at two main referral hospitals in Kenya." African Journal of Midwifery and Women's Health **5**(1): 21-24.

Rotich personal communication (2012). Clarification on partograph evaluation at two main referral hospitals in Kenya, 2011. S. Clow: email.

Royal College of Nursing (2005). Guidance for mentors of student nurses and midwives : An RCN toolkit. London, Royal College of Nursing.

Sackett, D. L., W. M. C. Rosenberg, J. A. M. Gray, R. B. Haynes and W. S. Richardson (1996). "Evidence based medicine: what it is and what it isn't (editorial)." British Medical Journal **312**: 71-72.

Schulz, K. F., D. G. Altman, D. Moher and for the CONSORT Group (2010) "CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials." BMC Medicine **8**, 18 DOI: 10.1186/1741-7015-8-18.

Scott, J., R. Mannion, H. Davies and Marshall MN. (2003). "Does organisational culture help to explain health care performance?" Journal of Health Services Research and Policy **8**(2): 105-117.

Scott, S., B. VandenBeld and G. Cummings (2011). "Optimizing Clinical Environments for Knowledge Translation : Strategies for Nursing Leaders." Nursing Leadership **24**(3): 73-85.

Sguassero, Y. (2007). Local opinion leaders : effects on professional practice and health care outcomes : RHL commentary. Geneva, World Health Organization.

Shelton, J. D. and R. Gray (2010). "Maternal mortality in 181 countries, 1980-2008." The Lancet **376**: 1389-1390.

Shih, C., C. Aye, M. Cruz, T. Garcia, T. Hui and C. Manosca (2010). "Reducing hospital acquired pressure ulcers by using an evidence based standardised nursing procedure." Journal of Wound, Ostomy and Continence Nursing **37**(35): 567.

Shor, I. and P. Freire (1987). "What is the dialogical method of teaching?" Journal of Education **109**(3): 11-31.

Shung King, M., H. de Pinho and R. E. Mahlanga (2006). Chapter 7 : The context of maternal and child health. The South African Health Review, 2006. Health Systems Trust. Durban, Health Systems Trust.

Siassakos, D., R. Fox, L. Hunt, J. Farey, C. Laxton, C. Winter and T. Draycott (2011). "Attitudes towards safety and teamwork in a maternity unit with embedded team training." American Journal of Medical Quality **26**(2): 132-137.

Simbar, M., F. Ghafari, S. T. Zahrani and H. A. Majd (2009). "Assessment of quality of midwifery care in labour and delivery wards in selected Kordestan Medical Sciences University hospitals." International Journal of Health Care Quality Assurance **22**(3): 266-277.

Simbar personal communication (2012). Assessment of quality of care in labour. Sheila Clow: personal communication email correspondence.

Smith, E. and G. Kirsten (2004). Perinatal asphyxia – are there early predictors for an unfavourable outcome which may disqualify infants from admission to an neonatal intensive care unit with limited resources. 23rd annual Priorities in Perinatal Care,

Limpopo, South Africa.

Smith, H., H. Brown and J. Khanna (2009). Continuing education meetings and workshops : effects on professional practice and health care outcomes : RHL commentary. Geneva, World Health Organization.

Soumerai, S. and J. Avorn (1990). "Principles of educational outreach ('academic detailing') to improve clinical decision making." Journal of the American Medical Association **263**(4): 549-556.

Soumerai, S., S. Salem-Schatz, J. Avorn, C. Casteris, D. Ross-Degnan and M. Popovsky (1993). "A controlled trial of educational outreach to improve blood transfusion practice." Journal of the American Medical Association **270**(8): 961-966.

South African Government Information Service (2008). South Africa Year book 2007/2008, Government Information Service.

South African Nursing Council (1987). R1469 Regulations relating to the scope of practice of persons who are registered or enrolled under the Nursing Act 1978 – Amendment (Chapter 3 – The scope of practice of the registered midwife). Department of Health, South African Nursing Council.

South African Nursing Council (1990). R2488 Regulations relating to the conditions under which registered midwives and enrolled midwives may carry on their profession, South African Nursing Council.

South African Nursing Council (2004). Charter of Nursing Practice South African Nursing Council

South African Qualifications Authority. (2003). "Registered Unit Standard : Conduct on-the-job coaching." Retrieved 6 June 2006, 2006, from <http://regqs.saga.org.za/viewUnitStandards.php?id=7818>.

South African Qualifications Authority. (2004). "Registered Unit Standard : Mentor a colleague to enhance the individual's knowledge, skills, values and attitudes in a selected career path." Retrieved 6 June 2006, 2006, from

<http://regqs.saga.org.za/viewUnitStandards.php?id=114215>.

Souza, J. P. (2007a). Local opinion leaders : effects on professional practice and health care outcomes : RHL commentary. Geneva, World Health Organization.

Souza, J. P. (2007b). Local opinion leaders : effects on professional practice and health care outcomes : RHL practical aspects. Geneva, World Health Organization.

Sprague, A. E., L. Oppenheimer, L. McCabe, I. D. Graham and B. L. Davies (2008). "Knowledge to Action : Implementing a guideline for second stage labour." American Journal of Maternal Child Nursing **33**(3): 179-186.

SPSS Inc. PASW Statistics v18.

Statacorp (2009). Stata v11.0. College Station, Statacorp.

Statistics South Africa. (2007). "Community Survey 2007." Retrieved 6 February 2012, from <http://www.statssa.gov.za/community-new/content.asp>.

Steele, R. (2009). "Gaining competence and confidence as a midwife." British Journal of Midwifery **17**(7): 441-447.

Steiner, J. (2014). "Promoting mentorship in translational research : should we hope for Athena or train mentor?" Academic Medicine **89**(5): 702-704.

Steyn, P., J. Evans and J. Makin (1994). Anthropometric measurements of the black female pelvis. 13th annual Priorities in Perinatal Care congress.

Stiell, I. G., C. M. Clement, J. Grimshaw, R. J. Brison, B. H. Rowe, M. S. Schull, J. S. Lee, J. Brehaut, R. D. McKnight, M. A. Eisenhauer, J. Dreyer, E. Letovsky, T. Rutledge, I. MacPhail, S. Ross, A. Shah, J. J. Perry, B. R. Holroyd, U. Ip, H. Lesiuk and G. A. Wells (2009). "Implementation of the Canadian C-Spine Rule: prospective 12 centre cluster randomised trial." British Medical Journal **339**(b4146).

Straus, S. E., J. Tetroe and I. D. Graham (2011). "Knowledge translation is the use of knowledge in health care decision making." Journal of Clinical Epidemiology **64**(1): 6-10.

Streiner, D. L. and G. R. Norman (2008). Health Measurement scales : A practical guide to their development and use 4th edition. Oxford, Oxford University Press.

Taljaard, M., A. D. McRae, C. Weijer, C. Bennett, S. Dixon, J. Taleban, Z. Skea, M. P. Eccles, J. C. Brehaut, A. Donner, R. Saginur, R. F. Boruch and J. M. Grimshaw (2011) "Inadequate reporting of research ethics review and informed consent in cluster randomised trials: review of random sample of published trials." British Medical Journal **342** DOI: 10.1136/bmj.d2496.

Tanner, C. A. (2006). "Thinking Like a Nurse: A Research-Based Model of Clinical Judgment in Nursing." Journal of Nursing Education **45**(6): 204-211.

Thaddeus, S. and D. Maine (1994). "Too far to walk : Maternal mortality in context." Social Science and Medicine **38**(8): 1091-1110.

Theron, G. (1997). Die effektiwiteit van die verloskunde handleiding van die perinatale self-onderrig program om voorgeboorte en intrapartum sorg te verbeter – 'n gekontroleerde studie. MD, University of Stellenbosch.

Theron, G. B. (1999). "Effect of the maternal care manual of the Perinatal Education Programme on the ability of midwives to interpret antenatal cards and partograms." Journal of Perinatology **19**(6): 432-435.

Theron, G. B. (1999a). "Improved cognitive knowledge of midwives practising in the Eastern Cape province of the Republic of South Africa through the study of a self-education manual." Midwifery **15**(2): 66-71.

Theron, G. B. (1999b). "Effect of the maternal care manual of the perinatal education programme on the ability of midwives to interpret antenatal cards and partograms." Journal of Perinatology **19**(6(Pt1)): 432-435.

- Thompson, D., C. Estabrooks, S. Scott-Findlay, K. Moore and L. Wallin (2007). "Interventions aimed at increasing research use in Nursing : A systematic review." Implementation Science **1**(15): 1-16.
- UNAIDS (2013). Global report : UNAIDS report on the global AIDS epidemic 2013. New York, United Nations.
- United Nations. (2003). "Millennium Development Goals." Retrieved 23 June 2006, 2006, from <http://www.un.org/unsd/mi/pdf/mdglist.pdf>.
- United Nations (2009). The Millennium Development Goals Report 2009. New York, United Nations: 60.
- United Nations Population Fund (2014). The State of the World's Midwifery 2014. A universal pathway : A woman's right to health. New York, United Nations.
- Uys, B. and S. Meyer (2005). "Critical thinking of student nurses during clinical accompaniment " Curationis **28**(3): 11-19.
- Uys, L., L. van Rhyn, N. Gwele, P. McInerney and T. Tanga (2004). "Problem-solving competency of nursing graduates." Journal of Advanced Nursing **48**(5): 500-509.
- van den Broek, N. and A. Falconer (2011). "Maternal mortality and Millennium Development Goal 5." British Medical Bulletin **99**: 25-38.
- Van der Veer, S., N. De Keizer, A. Ravelli, S. Tenkink and K. Jager (2010). "Improving quality of care. A systematic review of how medical registries provide information feedback to health care providers." International Journal of Medical informatics **79**(5): 305-323.
- Villar, J., H. Ba'aqeel, G. Piaggio, P. Lumbiganon, J. M. Belizán, U. Farnot, Y. Al-Mazrou, G. Carroli, A. Pinol, A. Donner, A. Langer, G. Nigenda, M. Mugford, J. Fox-Rushby, G. Hutton, P. Bergsjø, L. Bakketeig, H. Berendes and for the WHO Antenatal Care Trial Research Group (2001). "WHO Antenatal Care randomised trial for the evaluation of a new model of routine antenatal care " The Lancet **357**: 1565-1570.

Voce, A. (2004). Discussion regarding Philpott and Voce labour record review tool. S. Clow: personal communication : email correspondence.

Voce, A. S. (2005). The development of a district-based model of intervention for improving the quality of maternal health care at primary level. PhD, University of KwaZulu-Natal.

Vygostky, L. S. (1987). The Collected works of L.S. Vygotsky. Volume 1 Problems of General Psychology (including the volume "Thinking and Speech"). New York, Plenum.

Wagaarachchi, P. T., W. J. Graham, G. C. Penney, A. McCaw-Binns, Y. K. Antwi and M. H. Hall (2001). "Holding up a mirror : changing obstetric practice through criterion-based clinical audit in developing countries." International Journal of Gynecology and Obstetrics **74**(2): 119-130.

Walker, R., M. Cooke, A. Henderson and D. K. Creedy (2011). "Characteristics of leadership that influence clinical learning: A narrative review." Nurse Education Today **31** 14.

Wallen, G. R., S. A. Mitchell, B. Melnyk, E. Fineout-Overholt , C. Miller-Davis, J. Yates and C. Hastings (2010). "Implementing evidence-based practice : effectiveness of a structured multifaceted mentorship programme." Journal of Advanced Nursing **66**(12): 2761-2770.

Wamwana, E. B., P. M. Ndavi, P. B. Gichangi, J. G. Karanja, E. G. Muia and G. W. Jaldesa (2007). "Quality of record keeping in the intrapartum period at the Provincial General Hospital, Kakamega, Kenya." East African Medical Journal **84**(1): 16-23.

Webb, C. (1992). "The use of the first person in academic writing : objectivity, language and gatekeeping." Journal of Advanced Nursing **17**: 747-752.

Weijer, C., J. M. Grimshaw, M. P. Eccles, A. D. McRae, A. White, J. C. Brehaut, M. Taljaard and the Ottawa Ethics of Cluster Randomized Trials Consensus Group (2012). "The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials." PLoS Med **9**, e1001346 DOI: 10.1371/journal.pmed.1001346.

Weijer, C., J. M. Grimshaw, M. Taljaard, A. Binik, R. Boruch, J. C. Brehaut, A. Donner, M. P. Eccles, A. Gallo, A. D. McRae, R. Saginur and M. Zwarenstein (2011) "Ethical issues posed by cluster randomized trials in health research." Trials **12**, 100 DOI: 10.1186/1745-6215-12-100.

Wensing, M., T. van der Weijden and R. Grol (1998). "Implementing guidelines and innovations in general practice : which interventions are effective?" British Journal of General Practice **48**(427): 991-997.

White, E. and J. Winstanley (2010). "A randomised controlled trial of clinical supervision: selected findings from a novel Australian attempt to establish the evidence base for causal relationships with quality of care and patient outcomes, as an informed contribution to mental health nursing practice development " Journal of Research in Nursing **15**(2): 151-167.

Woods, D., R. Pattinson, D. Greenfield and S. Velaphi (2005). Improving survival rates of newborn infants in South Africa. 24th annual Priorities in Perinatal Care conference, Langebaan, South Africa.

Woods, D. L. and G. B. Theron (1994). "The Perinatal Education Programme." South African Medical Journal **84**(2): 61.

Woods, D. L. and G. B. Theron (1995). "The impact of the perinatal education programme on cognitive knowledge in midwives." South African Medical Journal **85**(3): 150-153.

Workneh, G., L. Scherzer, B. Kirk, H. R. Draper, G. Anabwani, R. S. Wanless, H. Jibril, N. Gaetsewe, B. Thuto and M. A. Tolle (2013). "Evaluation of the effectiveness of an outreach clinical mentoring programme in support of paediatric HIV care scale-up in Botswana." AIDS Care **25**(1): 11-19.

World Health Organization (1989). The partograph – a managerial tool for the prevention of prolonged labour, WHO : Geneva.

World Health Organization (1994). The Partograph : The application of the WHO partograph in the management of labour. Report of a WHO multicentre study 1990-1991. C. Lennox and B. Kwast. Geneva, World Health Organization: 251.

World Health Organization (2000). ACluster version 2.1 : Design and analysis of cluster random trials.

World Health Organization (2001). Safe Motherhood Needs Assessment version 1.1. Geneva, World Health Organization Maternal and Newborn Health / Safe Motherhood.

World Health Organization (2003). Integrated Management of Pregnancy and Childbirth - Pregnancy, childbirth and newborn care : a guide for essential practice, WHO : Geneva.

World Health Organization (2004). Making pregnancy safer : the critical role of the skilled attendant. A joint statement by the World Health Organization, the International Confederation of Midwives and the International Federation of Gynecology and Obstetrics. Geneva, The World Health Organization: 24.

World Health Organization (2009). Monitoring emergency obstetric care : a handbook. Geneva, World Health Organization.

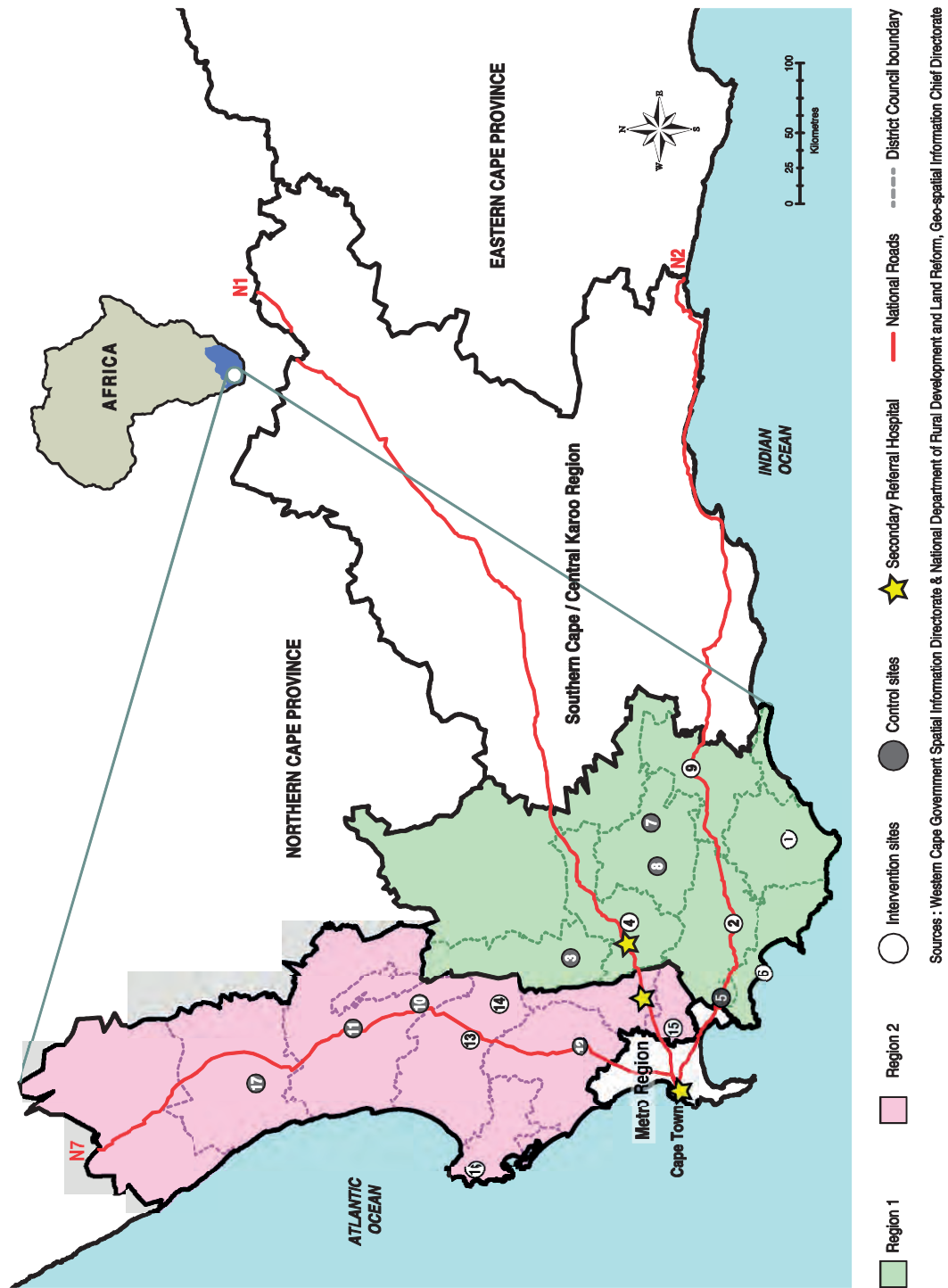
World Health Organization. (2014). "World Health Statistics. Part 1 : Health related Millennium Development Goals." Retrieved 18 September 2014, from http://www.who.int/gho/publications/world_health_statistics/EN_WHS2014_Part1.pdf?ua=1.

World Medical Association. (2013). "WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects." 64th WMA General Assembly, Fortaleza, Brazil. Retrieved 10 August 2014, from <http://www.wma.net/en/30publications/10policies/b3/>.

World Medical Association (2013). "World Medical Association Declaration of Helsinki : Ethical Principles for Medical Research Involving Human Subjects." Journal of the American Medical Association: e1-e4.

- Yaghmaie, F. (2003). "Content validity and its estimation." Journal of Medical Education **3**(1): 25-27.
- Yelland, J., A. Krastev and S. Brown (2007). "Enhancing early postnatal care : findings from a major reform of maternity care in three Australian hospitals." Midwifery **25**(4): 392-402.
- Young, P., R. Hamilton, S. Hodgett, M. Moss, C. Rigby, P. Jones and R. Johanson (2001). "Reducing risk by improving standards of intrapartum fetal care." Journal of the Royal Society of Medicine **94**: 226-231.
- Zhang, W.-H., C. Deneux-Tharaux, P. Brocklehurst, E. Juszczak, M. Joslin, S. Alexander and on behalf of the EUPHRATES Group (2010). "Effect of a collector bag for measurement of postpartum blood loss after vaginal delivery: cluster randomised trial in 13 European countries." British Medical Journal **340**(doi: 10.1136/bmj.c293).
- Zuber-Skerrit, O. (2002). "A model for designing action learning and action research programs." The Learning Organization **9**(4): 143-149.
- Zwarenstein, M., S. Treweek, J. J. Gagnier, D. G. Altman, S. Tunis, B. Haynes, A. D. Oxman, D. Moher and for the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups (2008). "Improving the reporting of pragmatic trials: an extension of the CONSORT statement." British Medical Journal **337**: a2390.

Appendix 1 Map of the Western Cape Province showing the regions, sites and site allocation



Appendix 2 Western Cape Partograph 2006 (page 1)⁶⁹

Provincial Government Western Cape : Department of Health

PARTOGRAM

Name:		Gravity:		Parity:		Gestation:		Spontaneous labour		Induced labour	
Age:		Risk factors:						Time of ROM:			
Folder no.:								Duration of labour (on arrival)			
Maternity Unit:											

LATENT PHASE												ACTIVE PHASE											
Time >												Time >											
Duration in hours >												Duration in hours >											
FETAL CONDITION Fetal heart rate O = before contraction X = after contraction Decelerations * > Liquor * >																							
PROGRESS OF LABOUR Station > Application * > Presenting part * > Caput * > Moulding * > Position e.g. LOA > Cervical dilation Cervical length Head above p/bkts XI																							
CONTRACTIONS CONTRACTIONS > 5 PER 10 MINUTES > 40 sec > 20 - 40 sec < 20 sec OXYTOCIN > Units Rate																							
MATERNAL CONDITION BLOOD PRESSURE & PULSE > URINE > TEMP °C																							
TIME > MANAGEMENT/ MEDICATION/ I.V. FLUID > SIGNATURE & DESIGNATION >												TIME > MANAGEMENT/ MEDICATION/ I.V. FLUID > SIGNATURE & DESIGNATION >											

* Moulding: 0 1+ 2+ 3+ * Caput: 0 1+ 2+ * Liquor: 1 - intact C - clear B - blood stained M - meconium stained * Presenting part: breech - Br face - F vertex - Vx brow - Bw * Application: good - G poor - P * Decelerations: early - E late - L variable - V none - N

PCWCpartogram10.2006(ng)

⁶⁹ The partograph is used as A3 size. It has been reduced to A4 size for inclusion in this dissertation

Appendix 2

Western Cape Partograph 2006 documentation (page 2)

ASSESSMENT No 1	Date _____	Time _____	DOL _____ hrs	DORM _____ hrs
Progress of Labour: good poor none _____				
Maternal condition: _____				
Fetal condition: _____				
Overall assessment & management: _____				
Name (PRINT) _____		Signature & designation _____		
ASSESSMENT No 2	Date _____	Time _____	DOL _____ hrs	DORM _____ hrs
Progress of Labour: good poor none _____				
Maternal condition: _____				
Fetal condition: _____				
Overall assessment & management: _____				
Name (PRINT) _____		Signature & designation _____		
ASSESSMENT No 3	Date _____	Time _____	DOL _____ hrs	DORM _____ hrs
Progress of Labour: good poor none _____				
Maternal condition: _____				
Fetal condition: _____				
Overall assessment & management: _____				
Name (PRINT) _____		Signature & designation _____		
ASSESSMENT No 4	Date _____	Time _____	DOL _____ hrs	DORM _____ hrs
Progress of Labour: good poor none _____				
Maternal condition: _____				
Fetal condition: _____				
Overall assessment & management: _____				
Name (PRINT) _____		Signature & designation _____		

Appendix 3a Summary of Systematic Reviews (*)

Authors, Date / Latest search / Para ref	Intervention / Outcomes	Types of studies and profiles of studies	Results	Author conclusion
Allen, Gillen and Rixson 2009 March 2008 2.4.6.6	Integrated pathways Outcome measures determined by the purposes of the studies selected for review - system, process and clinical outcomes, "What works, for whom, in what contexts?"	RCT 7 studies Health care professionals : Multidisciplinary (5 included nurses) Settings : High income countries (7) Health Service characteristics : Secondary and tertiary levels	7 RCTs reported in 9 papers were included. Meta-analysis and /or qualitative synthesis not possible due to heterogeneity	Integrated care pathways are most effective in contexts where patient care trajectories are predictable, and in bringing about behavioural changes where there are identified deficiencies in the services
Baker, Camosso- Stefinovic, <i>et al.</i> 2010 Oct 2009 2.4.6.3	Tailored interventions to overcome identified barriers to change Outcome : Measure of professional performance	RCT 26 studies Health care professional in charge of patient care: Physicians (16); Nurses (2); Multiprofessional teams (6) Community pharmacist / prescribers (2) Setting : High income countries (24) Health service characteristics : Hospital / specialist care (7); Primary / Community care (15); Mixed levels (3); Nursing home (1)	12 studies (all cluster randomised) included for meta-regression. Pooled OR : ♦ 1.54 (95% CI, 1.16 to 2.01) (Bayesian analysis); ♦ 1.52 (95% CI, 1.27 to 1.82, p<0.001) (classical analysis)	Interventions tailored to prospectively identify barriers are more likely to improve professional practice than no intervention or dissemination of guidelines.

Authors, Date / Latest search / Para ref	Intervention / Outcomes	Types of studies and profiles of studies	Results	Author conclusion
Estabrooks, Floyd, <i>et al.</i> 2003 February 2001 2.4.6.4	Examine individual characteristics of nurses and how they influence the utilisation of research (Did not test an intervention) Outcome : Identification of individual determinants	20 studies - no restriction on design Health care professionals: Nurses (equivalent of RN) Setting : Not reported; High income countries (7); remainder could not be determined from available information Health service characteristics : Hospital settings - various wards/units (9); clinical and managers (1); public health and agency (1); practice nurses (1); nurse educators (1); unspecified (7)	Methodological problems surfaced in all of the studies. Six categories of potential individual determinants were identified : ♦ beliefs and attitudes ♦ involvement in research activities ♦ information seeking ♦ professional characteristics ♦ education ♦ other socio-economic factors	There was little to suggest that any potential individual determinant influences research use
Flodgren, Parmelli, <i>et al.</i> 2011 4 May 2010 2.4.3.3	Local opinion leaders Outcome : Compliance with desired practice	Cluster RCT 18 trials (13 analysed appropriately at cluster level) (included 296 hospital + 318 primary care practices) Health care professional in charge of patient care : Physicians (14); Nurses (2); Physicians, Nurses and Midwives (2) Setting : High income countries (17); Health service characteristics : Hospital based (14); Primary care (1); Primary and secondary care (2)	Overall the median adjusted RD was +0.12, i.e. 12% absolute increase in compliance in intervention group	Opinion leaders alone or in combination with other interventions may successfully promote EBP but effectiveness varies

Authors, Date / Latest search / Para ref	Intervention / Outcomes	Types of studies and profiles of studies	Results	Author conclusion
Flodgren, Rojas- Reyes, <i>et</i> <i>al.</i> 2012 7 March 2011 2.4.6.5	Organisational infrastructures to promote evidence based nursing practice Outcome measure : Hospital acquired pressure ulcers (HAPU) rate	1 study re-analysed as an interrupted time series. Participants : Health care organisations comprising nurses, midwives and health visitors Setting : High income country Health service characteristics : Hospital	There was no evidence of an intervention effect at 3 months (mean rate per quarter 0.7%; 95% CI 1.7 to 3.3; $r=0.457$). Given the small percentages post intervention it was not statistically possible to extrapolate effects beyond 3 months.	The review question remains unanswered
Forsetlund, Bjorndal, <i>et al.</i> 2009 29 June 2008 2.4.3.5.1	Continuing education meetings and workshops Outcome measures : ♦ Compliance with desired practice ♦ Patient outcomes	RCT 81 trials (involving >11 000 health professionals) Participants : Physicians only (68); Nurses only (5); Combination with nurses or midwives (4); Other health professions only (4) Settings : High income countries (75) Health service characteristics : General practice (43); Hospitals (17); Community (16); Other (5)	Health professional outcomes : ♦ Based on 30 trials (36 comparisons) Compliance with desired practice where educational meetings were a component of the intervention vs no intervention - median adjusted RD = 6% (IQR 1.8% to 15.9%); ♦ Based on 21 trials (19 comparisons) educational meetings alone median adjusted RD = 6% (IQR 2.9% to 15.3%); ♦ Continuous outcomes (5 trials) median adjusted percentage relative to control = 10% (IQR 8% to 32%); Patient outcomes : ♦ (5 trials) Median adjusted RD in achievement of treatment goals = 3.0% (IQR 0.1% to 4.0%)	Educational meetings alone or combined with other interventions can improve professional practice and health care outcomes for the patients. The effect is most likely to be small and similar to other types of continuing medical education such as audit and feedback, and educational outreach visits

Authors, Date / Latest search / Para ref	Intervention / Outcomes	Types of studies and profiles of studies	Results	Author conclusion
Grimshaw, Thomas, <i>et al.</i> 2004 1998 2.4.3 2.4.3.1 2.4.3.2 2.4.3.4	Guideline development, dissemination and implementation. Resource implications and cost/benefit analysis to guide appropriate implementation decisions Outcomes : Objective measures of provider behaviour or patient outcomes	RCT, Controlled clinical trials (CCT), Controlled before and after studies (CBA), Interrupted time series analyses (ITS). 235 studies included with 309 comparisons Participants : Medically qualified health care professionals (physicians or general practitioners only 74 studies) Nurses Settings : High income countries except for 2 studies (Mexico and Thailand) Health service characteristics : Mixed levels (mainly primary) Provider behaviours : prescribing (113); test ordering (73)	The majority of interventions observed modest to moderate improvements in care Median absolute improvement across interventions (cluster randomised) : ♦ reminders 14.1%; ♦ dissemination of educational materials 8.1%; ♦ audit and feedback 7.0%; ♦ multifaceted including educational outreach 6.0% ♦ Economic outcomes - methods employed of poor quality.	There is an imperfect evidence base to support decisions about which guideline implementation and dissemination strategies are likely to be efficient under different circumstances
Horsley, Hyde, <i>et al.</i> 2011 16 June 2011 2.4.3.5.4	Teaching critical appraisal skills in healthcare settings Outcome measures : ♦ Process of care ♦ Patient-related outcomes ♦ Knowledge ♦ Critical appraisal skills	RCT 3 studies Participants : Interns and Physicians Setting : High income countries Health service characteristics : Not relevant - individual practitioners targeted	Impact of teaching critical appraisal on process of care of patients and patient outcomes - none of the included studies reported these. Impact of teaching critical appraisal on knowledge /awareness of health professionals: ♦ Improvement in knowledge reported in 2 studies and was statistically significant; ♦ Improvement in critical appraisal skills was reported in all 3 studies. Statistically significant results found in 2 studies.	Low intensity critical appraisal teaching interventions in health care populations may result in modest gains.

Authors, Date / Latest search / Para ref	Intervention / Outcomes	Types of studies and profiles of studies	Results	Author conclusion
Ivers, Jamtvedt, et al. 2012 15 Sept 2011 2.4.3.2	Audit and Feedback Outcomes ◆ Compliance with desired practice ◆ Patient outcomes	RCT : 140 studies (88 with cluster allocation) Health care professionals : Physicians (121); Nurses (16); Pharmacists (5) Setting : High income countries (136) Health service characteristics : Hospital based (36); Outpatient setting (94)	Health care professional's compliance with desired practice : ◆ 49 studies with dichotomous outcomes - weighted median adjusted RD = 4.3% absolute increase; ◆ 21 studies with continuous outcomes – weighted median adjusted change relative to control = 1.3% Patient outcomes : ◆ 6 studies with dichotomous outcomes – weighted median RD = -0.4%; ◆ 5 studies with continuous outcomes – weighted median percentage = 17%	Audit and feedback generally leads to small but potentially important improvements in professional practice
O'Brien, Oxman, et al. 2007 March 2007 2.4.3.1	Educational outreach visits (EOV) Outcome measures : Effects of EOV on health professional practice or patient outcomes	Randomised trials 69 (with >15 000 health professionals) Health care professionals : Physicians (55); Combination including nurses (4); other combinations and categories (10) Setting : High income countries except for 3 studies (Thailand and Indonesia) Health service characteristics : Primary level (54); Hospital (7); Nursing homes (4); Pharmacies (2); Other (3)	Health care professional's compliance with desired practice: 28 studies (34 comparisons) ◆ Median adjusted risk difference 5.6% (IQR 3.0 to 9.0%); ◆ Prescribing (17 comparisons) median adjusted RD 4.8% (IQR 3.0 to 6.5%); ◆ Other professional performance (17 comparisons) median adjusted RD 6.0% (IQR 3.6% to 16.0%) ◆ Meta-regression : large number of explanatory factors. Did not provide any compelling explanations for observed variation in RDs Patient outcomes : 14 studies For most studies, it was difficult to determine if there was sufficient power to detect an important difference at the patient level.	EOVs alone or combined with other interventions have relatively consistent and small, but potentially important effects on prescribing. Effects on other types of professional performance vary from small to modest improvements, and it is not possible from this review to explain that variation. Compared with audit and feedback, EOVs appeared to be slightly superior

Authors, Date / Latest search / Para ref	Intervention / Outcomes	Types of studies and profiles of studies	Results	Author conclusion
Parmelli, Floodgren <i>et al.</i> 2011 7 Dec 2010 2.4.6.5	Strategies to change organisational culture to improve practice Outcome measures : ◆ Professional performance ◆ Patient outcomes e.g. mortality or functional health status ◆ Organisational performance, e.g. wait times, staff turnover	RCT, Quasi-experimental studies, Controlled clinical trials (CCT), Controlled before and after studies (CBA), Interrupted time series analyses (ITS). None found to meet the quality criteria developed by EPOC. Health service characteristics : Any type of health care organisation in which strategies to change organisational culture to improve health care performance were applied	No results	It is not possible to draw any conclusions.
Pattinson, Kerber, <i>et al.</i> 2011 2009 2.4.3.2	Facility-based perinatal mortality and community audits Outcomes : ◆ Institutional level ◇ Evidence of mortality effect ◇ Experience in perinatal audit process and sustainability ◆ Community level ◆ Implementing and scaling-up	Before and after studies 7 Skilled attendants : 24% - >90% cover Setting : Low and middle income countries, rural and urban Health service characteristics : All levels of care (mobile clinics - central referral / teaching hospital)	Meta-analysis indicated a reduction in perinatal mortality of 30% (95% confidence interval, 21%–38%) after introduction of perinatal audit	The consistency of effect suggests that audit may be a useful tool for decreasing perinatal mortality rates in facilities and improving quality of care. This is dependent on implementing solutions to problems identified, without which audit alone cannot improve quality of care.

Authors, Date / Latest search / Para ref	Intervention / Outcomes	Types of studies and profiles of studies	Results	Author conclusion
Thompson, Estabrooks <i>et al.</i> 2007 2006 2.5	Interventions aimed at increasing research use in nursing Outcome measure : General measure of research use	RCT 3; Controlled Before and after 1 Health care professionals : Nurses (3);multidisciplinary team, including nurses (1) Setting : High income countries Health service characteristics : Hospital based - various specialty areas	Meta-analysis not possible due to heterogeneity, methodological weakness and lack of effect sizes. 6 cohorts (educational meetings (4); educational meeting and local opinion leader (1); multidisciplinary committee for guideline implementation multidisciplinary (1) ♦ Educational meetings of varying content, duration, and frequency cannot be said to be effective research utilisation interventions in Nursing. Require more rigorous investigation. ♦ Educational meeting and local opinion leader - increased research utilisation, but study of low quality and represents inconclusive evidence. ♦ Multidisciplinary committee - effective research utilisation	The evidence to support or refute specific interventions about how to increase research use in Nursing is inconclusive.

Acronyms **CBA** Controlled before-and-after **CCT** Controlled clinical trial **CI** Confidence interval **CRT** Cluster randomised trial
EBP Evidence based practice **EOV** Educational outreach visits **IQR** Interquartile range **ITS** Interrupted time series
OR Odds Ratio **RCT** Randomised controlled trial **RD** Risk difference **RN** Registered (professional) nurse

Appendix 3b Effect of different implementation strategies on professional behaviour change (§)

This appendix presents the data from a range of systematic reviews and illustrates this for each topic. Some comparisons are made with different source studies due to different search terms used for the reviews. In order to assist the reader pick out the trends, the direction of the effects are indicated in different colours - see legend. The majority of studies report effect on desired performance. These are presented first for each strategy. Where patient outcomes are reported, this data follows the performance data.

Legend

Superscript numbers indicate the reference. Those that appear after the heading indicate the index systematic review for that subject, where this is included.

Number in brackets refer to the number of studies and comparisons (in the case of a systematic review), e.g. (12-18) denotes 12 studies and 18 comparisons. Where the number of comparisons is not specified, the number of studies is reflected alone.

Italics : Some studies are reflected in more than one place. This is done to illustrate as full a picture as possible for each strategy. Where a duplicate entry is included, this is italicised.

Overall refers to any study in which the characteristic is included alone or in combination and compared with usual practice / no intervention.

Green text represents any positive change for that characteristic, regardless of the effect size.

Orange text represents varied effects.

Blue text represents any negative change for that characteristic

Black text in relation to effect represents no change, or non-significance.

Abbreviations

A&F : Audit and Feedback	EM : Educational materials	EOV : Educational Outreach Visits	IQR : Inter-quartile range
KT : Knowledge translation	NS : Not significant	RD : Risk difference	VBAC: Vaginal Birth After Caesarean

2.4.3 Single or multifaceted interventions				Conclusion : Multifaceted interventions tend to be more effective than single interventions, but this effect is modest and is dependent on contextual factors.	
Using boxplots, visually there appeared to be no relationship between the number of component interventions and the effects of multifaceted interventions ¹					
Multifaceted interventions tended to be more effective (61) ² Social influence and management support can improve the effect of information transfer				Accompanied by increased cost. ²	
Multi-component KT interventions more effective for self perceived knowledge and practice change than passive dissemination or single KT intervention (12) ³ No effect on attitudes to evidence-based practice				Heterogeneity of data precluded meta-analysis	
2.4.3.1 Educational outreach visits (EOV) ⁴				Conclusion : EOVs alone or in combination show relatively consistent and small effects. Appear to be slightly superior to A and F. ⁴	
Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments		
Overall effect on desired practice ⁴ (28) Small Median adjusted RD 5.6% (IQR 3.0% to 9.0%)	EOV in combination vs EOV alone ⁴ No change Box plot analysis suggest this may be more effective but multivariate analysis showed this was not statistically significant	Any combination including EOV vs any combination including A&F + reminders (8-12) Effect on professional practice : Interventions with EOV appeared to be slightly superior to A&F alone, and to the same degree as no intervention			
Effect alone on professional practice ⁴ Modest to moderate ♦ Dichotomous outcomes (16-18) Median adjusted RD 5% (IQR 3% to 6.2%) ♦ Continuous outcomes (14-15) Median adjusted RD = 23% (IQR 12% to 39%)	Multifaceted including EOV vs no intervention control ¹ Effect on performance Relatively ineffective ♦ Dichotomous outcome measures (5)- median absolute improvement -1.0% (range -7.0% to +3.0%) ♦ Continuous outcome measures (4) – median relative improvement 0% (range -1.4% to 2.7%)	EOV vs A&F ⁵ Effect on professional practice Varied In 2 studies ♦ 1 favoured educational outreach ♦ 1 showed no difference	Where educational outreach was included in the intervention package, the median effect of the combination of two interventions was 4.6%, with three interventions it was 11%, and with 4 interventions it was 0.1%. ¹		

Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments
	EOV, educational meetings and educational materials vs no intervention control ¹ (4) Moderate effect Median absolute improvement in performance 11.0% (range 8.4% to 16.4%)	EOV vs other intervention educational materials ¹ (10-11) absolute improvement in care modest and less than when compared to no intervention control	EOV appeared to be more effective than education materials in 3 out of 5 comparisons and audit and feedback ¹ in two out of five comparisons ¹
Effect of individual visits vs group visits on professional practice ⁴ Varied effect	EOV + educational materials vs no intervention control ¹ (8) Effect on patient outcomes Relatively ineffective Median absolute improvement in performance +1.2% (range -5.6 to 13.1%)	EOV + Educational materials vs educational materials alone ¹⁰ EOV + A&F + educational materials vs educational materials alone ¹⁰ Printed educational material alone might have been as effective as the other two strategies	
Overall effect on patient outcomes ⁴ (14) Undetermined	EOV + educational materials + educational meetings vs no intervention control ¹ (4) Modest to moderate effect Median absolute improvement in performance 11.0% (range 8.4 to 16.4%)		

2.4.3.2 Audit and feedback (A&F) ⁵					Conclusion : Audit and feedback generally leads to small but potentially important improvements in professional practice. When combined with other interventions, the effect size was larger than when A&F used alone. ⁵
Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments		
Overall effect on professional practice ⁵ (70-108) Small weighted mean adjusted RD ◆ 4.3% (dichotomous outcomes) (49-82) ◆ 1.3% (continuous outcomes) (21-26)	A&F as <u>core essential feature</u> vs usual care ⁵ Effect on professional practice Moderate ◆ weighted mean adjusted RD 5.5% (IQR 0.4% to 16%) (dichotomous outcomes) (32-50) ◆ weighted mean adjusted change relative to baseline control 26.1% (IQR 12.7% to 26.1%) (continuous outcomes) (11-12)		Feedback varies : ◆ Type - verbal, written, both ◆ Source - internal or external ◆ Frequency - weekly, monthly, less than monthly ◆ Goals for change and/or Action ^{5,6} plans - present, absent		
Effect alone on professional practice ⁵ Small ◆ weighted median adjusted RD 3% (IQR 1.8% to 7.7%) (dichotomous outcomes) (26-32) ◆ weighted median adjusted RD 1.3% (IQR 1.3% to 11%) (continuous outcomes) (13-14)	A&F as <u>core essential feature</u> + reminders vs A&F alone ⁵ (13) Effect on professional practice Varied (lack of statistical testing, generally positive)	A&F vs reminders ⁵ Effect on professional practice Varied In 8 studies, ◆ 3 favoured reminders ◆ 2 favoured audit and feedback ◆ 3 had no change	Most trials measured professional practice in terms of prescribing or ordering of laboratory tests		
Effect on professional practice ¹ Modest dichotomous <u>process measures</u> (6-6) mean absolute improvement 7.0% (range 1.3% - 16%)	A&F as <u>core essential feature</u> + educational outreach vs A&F alone ⁵ (19-24) Effect on professional practice Varied ◆ weighted median adjusted RD 0.7% (IQR -1.1% to 5.1%) (dichotomous outcomes) ◆ median adjusted change relative to baseline 27% (IQR 0% to 40.5%) (continuous outcomes)	AF vs educational outreach ⁵ Effect on professional practice Varied In 2 studies ◆ 1 favoured educational outreach ◆ 1 showed no difference	In two out of five comparisons A&F appeared to be less effective than EO ¹		

Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments
Effect on patient outcomes ⁵ weighted mean adjusted RD Small to moderate ♦ 0.4% (dichotomous outcomes) (6-12) ♦ 17% (continuous outcomes) (5-8)		A&F vs local opinion leader ⁷ Large effect favours local opinion leader ♦ Trial of scar : 46% increase ($p=0.007$) ♦ VBAC : 85% increase ($p=0.003$)	
Effect on patient outcomes ⁷ ♦ rate of trial of scar ♦ rate of vaginal birth after caesarean section (VBAC) No difference	A&F as <u>core essential feature</u> + other educational interventions vs A&F alone ⁵ (13) Varied depending on targeted behaviour and setting - positive or no change	A&F vs other educational interventions ⁵ ♦ small problem-based learning groups (1) (no p value) ♦ seminar (1) Seminar adjusted change relative to baseline performance =22% ($p=0.03$) A&F less effective ♦ written material (2) (no p value or NS)	
Effect of perinatal audit on perinatal mortality ⁶ (7) Large effect ♦ 32% reduction (95% confidence limit, 21%-38%) in perinatal mortality	A&F + educational intervention ⁸ Effect on patient outcomes Large, sustained over 7 years ♦ 50% reduction in postpartum haemorrhage ♦ partograph recordings improved from 1/3 potentially misleading to 90% clinically helpful in decision-making		♦ Comprehensive feedback and seminars to all staff ♦ One-to-one question and answer sessions ♦ Active re-training of doctors and midwives on management of third stage of labour ♦ Introduction of the use of the partograph ♦ Initiatives supported and encouraged by hospital leadership ♦ Publication of audit reports in hospital newsletter ⁸

Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments
	A&F as core essential feature + case management or organisational interventions vs A&F alone ⁵ (4) Small to large Professional behaviour change ρ not significant or absent	A&F vs case management or organisational interventions ⁵ (2) No effect	
	A&F as core essential feature + financial incentives vs A&F alone ⁵ (2) No change to large adjusted RD -5.4% to 12.7% (no ρ values reported)	A&F vs financial incentives ⁵ In 1 study A&F was more effective adjusted change relative to baseline 41% $\rho=0.05$	
	A&F as core essential feature + patient mediated interventions vs A&F alone ⁵ Only 1 of 5 studies showed a positive change	A&F vs patient mediated interventions ⁵ (3) No difference	
	A&F, Educational Intervention + organisational support ⁹ Large effect defendable or possibly defendable risk of low Apgar scores dropped to 9.25% from a previous level of 32% (from the original baseline of 74%)		Compulsory attendance for training and feedback for ALL relevant staff. Feedback included not only cases where care might have been poor, but also those where care was of a high standard ⁹
		A&F + educational materials vs educational materials alone ¹⁰ A&F + EOVS + educational materials vs educational materials alone ¹⁰ Printed educational material alone might have been as effective as the other two strategies	

2.4.3.3 Local opinion leaders ¹¹				Conclusion : The effect of intervention on compliance with desired practice appears larger when a local opinion leader is included in the intervention. ¹¹	
Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments		
Compliance with desired practice ¹¹ (15-63) Moderate Effect Overall median adjusted RD=0.12 increased compliance			Most studies did not describe in detail the role and activities of the local opinion leader		
Compliance with desired practice ¹¹ (5-37) Moderate Effect Median adjusted RD = +0.09 (dichotomous outcomes)	Opinion leaders part of multiple intervention vs no intervention ¹¹ (7-13) Moderate effect Median adjusted RD = + 0.1	Opinion leaders vs other interventions ¹¹ (2-3) Moderate Effect Median adjusted RD = +0.14			
	Opinion leaders + one or more interventions vs the same one of more interventions ¹¹ (4-10) Moderate Effect Median adjusted RD = +0.1	Local opinion leader vs A&F ⁷ Effect favours local opinion leader ♦ Trial of scar : 46% increase ($p=0.007$) ♦ VBAC : 85% increase ($p=0.003$)			

2.4.3.4 Distribution of educational materials (EM)				Conclusion : Modest improvements where educational materials are the only intervention, may be short lived. They may contribute to an increased effect when used in combination with other interventions. ¹⁰	
Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments		
Provider behaviour outcome (15-15) ¹ Modest effect Dichotomous <u>process</u> measures (5-5) Median absolute improvement 8.1% (Range 3.6% - 17%)			In 3 out of 5 comparisons education materials appeared to be less effective than EO ¹		
	EM vs EM + A&F ¹⁰ Details of results not reported		It appeared that printed educational material alone might have been as effective as the other two strategies, but without a non-intervention arm this was difficult to assess. ¹⁰		
	EM vs EM + EO ¹⁰ Details of results not reported				
	EM vs EM + A&F + EO ¹⁰ Details of results not reported				
Effect on nurses' knowledge, practice and attitudes regarding the benefit of exercise for breast cancer patients. ¹² Statistically significant improvements for each domain					

2.4.3.5 Continuing education meetings¹³ Conclusion : The effect appears to be larger with higher attendance at the educational meetings and with mixed interactive and didactic educational meetings. Educational meetings did not appear to be effective for complex behaviours and they appeared to be less effective for less serious outcomes. ¹³				
Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments	
<u>Overall</u> compliance with desired practice ¹³ (30-36) Small effect Median adjusted RD 6% (IQR 1.8% to 15.9%)			The effect of educational meetings alone on professional practice was the same as for multifaceted interventions that included educational meetings	
Compliance with desired practice ¹³ Small effect ♦ dichotomous data (19-21) median adjusted RD = 6% (IQR 2.9% to 15.3%) ♦ continuous outcomes (5-5) median adjusted relative percentage 10% (IQR 8% - 32%)	Any intervention with educational meetings compared to educational meetings alone ¹³ (1) 12% adjusted relative percentage increase in compliance with desired practice	Educational meetings vs other interventions ¹³ (2) Negative effect Compliance with desired practice adjusted RD -8% and -0.14% respectively.	Characteristics considered : ♦ number of participants ♦ format ♦ source ♦ frequency ♦ total length Intensity determined on above characteristics and categorised as : ♦ intensive ♦ moderately intensive ♦ non-intensive ¹³	
<u>Overall</u> patient outcomes - achievement of treatment goals ¹³ Small effect ♦ dichotomous outcomes (5) median adjusted RD = 3.0% (IQR 8-32%) ♦ continuous outcomes (8-9) median adjusted relative percentage change 4% (IQR 0-11%)		Interactive educational meetings vs didactic (lecture based) educational meetings ¹³ (1) Compliance with desired practice adjusted RD 1.4% not statistically significant		

Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments
Patient outcomes ¹³ Small to moderate effect ♦ dichotomous outcomes (3) median adjusted RD 3% (IQR -0.9% - 4%) ♦ continuous outcomes (6) median adjusted relative percentage change 8% (IQR 0%- 12%)			

2.4.6.3 Tailored interventions to overcome barriers to change ¹⁴				Conclusion : Interventions tailored to prospectively identify barriers are more likely to improve professional practice than no intervention or dissemination of guidelines. ¹⁴	
Overall effect and Effect alone vs usual practice	Effect in combination	Effect vs other strategies	Characteristics / system supports / Comments		
Overall comparison of a tailored intervention to address identified barriers to change vs no intervention or an intervention not tailored to the barriers ¹⁴ (12) Modest effect Pooled OR 1.52 (95% CI, 1.27 to 1.82, p<0.001) (classical analysis)			Methods used to identify barriers and tailor interventions to address them need further development ¹⁴		
Comparison of a tailored intervention vs no intervention ¹⁴ (4) Modest effect OR 1.58 (95% CI, 0.96 to 2.59)		Comparison of a tailored intervention vs a non-tailored intervention ¹⁴ (8) (7 of the 8 studies had dissemination of educational materials as the comparator) Modest effect OR 1.56 (95% CI, 1.27 to 1.90)	Barriers identified : ♦ administrative constraints, e.g. lack of time, staff, facilities ♦ clinical uncertainty ♦ patient expectations ♦ information management ♦ sense of competence ♦ financial disincentives ♦ other, e.g. negative staff attitudes, anxiety about changing practice, perception that clinical issues not a priority, advocacy of certain drugs by pharmaceutical companies ¹⁴		

References (Author lists complete in the Reference list)

1. Grimshaw, J., R. Thomas, *et al.* (2004). "Effectiveness and efficiency of guidelines dissemination and implementation strategies." Health Technology Assessment 8 (6)
2. Wensing, M., T. van der Weijden *et al.* (1998). "Implementing guidelines and innovations in general practice : which interventions are effective?" British Journal of General Practice 48 (427): 991-997.
3. Menon, A., N. Korner-Bitensky, *et al.* (2009). "Strategies for rehabilitation professionals to move evidence based knowledge into practice : a systematic review." Journal of Rehabilitation Medicine 41(13): 1024-1032
4. O'Brien, M.A., S. Rogers, *et al.* (2007) Educational outreach visits : effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews 2007(4):CD000409
5. Ivers, N., G. Jamtvedt *et al.* (2012) Audit and feedback : effects on professional practice and healthcare outcomes (Review)." Cochrane Database of Systematic Reviews 2012(6):CD000259.
6. Pattinson, R., K. Kerber, *et al.* (2009). "Perinatal mortality audit : Counting, accountability, and overcoming challenges in scaling up in low- and middle-income countries". International Journal of Gynecology and Obstetrics 107:S113-122
7. Lomas, J., M. Enkin, *et al.* (1991). "Opinion leaders vs audit and feedback to implement practice guidelines" Journal of the American Medical Association 265(17):2202-2207
8. Mercer, S.W., K. Sevar, *et al.* (2006). "Using clinical audit to improve quality of obstetric care at the Tibetan Delek Hospital in North India." Reproductive Health 3(4)
9. Young, P., R. Hamilton, *et al.* (2001). "Reducing risk by improving standards of intrapartum fetal care." Journal of the Royal Society of Medicine 94:226-31
10. Cheater, F., R. Baker, *et al.* (2006). "Cluster randomized controlled trial of the Effectiveness of Audit and Feedback and Educational Outreach on Improving Nursing Practice and Patient Outcomes." Medical Care 44:543-51
11. Flodgren, G., E. Parmelli, *et al.* (2011). "Local opinion leaders: effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews 2012(2): CD002212
12. Kirshbaum, M. (2008). "Translation to practice : a randomised, controlled study of an evidence based booklet for breast cancer nurses in the United Kingdom." Worldviews on Evidence Based Nursing 5(2):60-74
13. Forsetlund, L., A. Bjorndal *et al.* (2009). "Continuing education meetings and workshops : effects on professional practice and health care outcomes." Cochrane Database of Systematic Reviews 2009(2): CD003030
14. Baker, R., J. Camosso-Stefinovic, *et al.* (2010). "Tailored interventions to overcome barriers to change : effects on professional practice and health care outcomes (Review)". Cochrane Database of Systematic Reviews 2010(3): CD005470.

Appendix 4 Validation documentation

Validation of the partograph utilisation instrument took place over three rounds.

Round One

Judges received version one of the partograph utilisation instrument, the Western Cape partograph, the content validity assessment tool, and an explanatory letter of the process to be followed.

Between each round

When the required level of consensus was achieved on items on the partograph utilisation instrument, changes were accommodated on the instrument as version two or three as applicable. Items where consensus had not been achieved were highlighted, different suggestions (or comments) from the judges captured, and specific responses requested (response document).

Rounds Two and Three

Judges received the relevant version of the partograph utilisation instrument on which to comment, the content validity assessment tool, and the relevant response document.

The Content Validity assessment tool score sheet remained the same throughout, with explanatory notes to assist with clarification. The tool was labelled for each round to ensure that feedback was captured accurately.

As an illustration of the process, the following documents are included in this appendix :

- ◆ Content Validity Assessment tool⁷⁰, Round two (with appendices one to three, see letter, below)
- ◆ Response to judges after round one (Letter dated 19 January 2009)
- ◆ Round two response form (with appendices four to six)

⁷⁰ This is the key document referred to in 3.6.1, hence it is presented first in this appendix. Other documents are supportive of the process.

Content Validity Assessment tool, Round Two

Instruction

Please mark the appropriate box with your score. Using a Likert scale of 1-4, indicate your response, where 1 = least agreement and 4 = very strong agreement.

	1	2	3	4
Development / Modification phase				
1 There is evidence of previous studies on partograph assessment tools				
2 There is reference to already existing tools				
3 There is evidence of expert consultation in the field of intrapartum care				
4 There is a checklist				
Content domain *				
5 Are all the items for instrument construction included in the check list? :				
a) Clinical data				
b) Clinical management information				
c) Medico-legal requirements				
d) Quality assurance measures				
6 The items are unambiguous :				
a) Only one variable per item				
b) Wording is clear				
c) Scoring				
The jury				
7 The panel of judges includes the practicing experts in intrapartum care				
8 The number of judges is adequate to pass acceptable judgement				

Please turn over

* Certain specific responses have been requested to be made on the Response form for Round 2. Any other input (omissions, additions or editing suggestions) for Questions 5 and 6 can be made on this form. Indicate clearly which item/s on the partograph evaluation checklist are being commented on.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Please feel free to comment further on any other question in this content validity assessment tool.

.....

.....

.....

.....

.....

.....

Thank you

c:\mydocs\phd meth\content validity assessment tool-judges Round 2

Appendix 1 : Development and modification (CVI - Questions 1-4)

Some judges indicated difficulty with some of these questions. Some of this is because this checklist was not being developed *de novo*, but is a modification. I left these questions in the CVI for Round 1 in case, for example, there was evidence regarding existing labour record review tools that had been tested, of which I was unaware. If there is any information regarding these 4 statements which needs to be brought to my attention, I would be grateful if you would give me feedback about this.

Most of the studies that I have come across look at the impact of the use of the partograph on maternal and neonatal outcomes. Disclosure about the evaluation tool used on evaluating partographs themselves, as is required by my study, is minimal. One study published in 2000 evaluated the use of the partograph by Angolan midwives in a peripheral delivery unit (Pettersson, Svensson and Christensson, 2000). The objective was to study the impact of an educational intervention on midwives use of the partograph. Ten variables were used to examine the partograph. This is significantly short on detail, and it certainly does not include much of what is in the current Philpott and Voce instrument. So, although this has been used in a published study, there is no information as to the validation and reliability of this as an instrument.

Since starting with the larger research project in 2004, and considering the existing Philpott and Voce Labour record review tool, I have consulted with the following people to ascertain the strengths and limitations in the existing Philpott and Voce Labour record review tool and to inform the development of the checklist currently under modification :

- ◆ Professor Hugh Philpott and Dr Anna Voce.
- ◆ MCWH managers and clinical supervisors of primary and secondary maternity services (metro, rural, region, province), who are responsible for auditing labour records.
- ◆ Members of the Western Cape maternity guidelines group, which include, Professor Ed Coetzee (now retired, UCT - Obstetrics and Gynaecology), Dr Stefan Gebhardt (Western Cape co-ordinating clinician for women's health, and regional obstetrician for West Coast / Winelands region), Professor Cheryl Nikodem (University of the Western Cape – Nursing), Dr Charl Oetlé (Regional obstetrician for Boland / Overberg region), Professor Gerhard Theron (University of Stellenbosch – Obstetrics and Gynaecology).
- ◆ Midwives in leadership positions, which include Mrs Dolly Nyasulu and Ms Zo Mzolo.
- ◆ Professor Bob Pattinson University of Pretoria and Director MRC Unit for Maternal and Infant Care strategies, who is also my co-supervisor.

Appendix 2 : The jury (CVI - Questions 7 and 8)

Lynn's (1986) work indicates that one should have a minimum of three judges, but preferably five to ten. Where there are five or fewer judges, one would require complete consensus, whereas if there are six or more judges, there is some room for minor disagreement. What I'm attempting to achieve during this exercise is to build consensus through this process of consultation and judgement.

My aim therefore was to have a panel of 10 judges. I was not able to give you information about who the other judges were (Q8) as I approached everyone at about the same time to participate and so had not yet obtained their consent to participate. I approached people with a diversity of professional backgrounds and relevant expertise in order to get as rich a set of responses as possible. The people who have agreed to participate as judges are :

Name	Profession / Discipline	Position / Location
Dr Stefan Gebhardt	Obstetrician	Western Cape Co-ordinating Clinician: Women's Health
Ms Liesbeth Mangate	Advanced Midwife	National Department of Health NEPAD MCWH
Ms Mickey Msasa	Advanced Midwife	National Department of Health MCWH / administrator
Ms Zo Mzolo	Advanced Midwife	UKZN Centre for Rural health, formerly educator in the DEPAM programme
Dr Jenny Nash	Medical Officer	Primary health care (E Cape); extensive experience in rural district practice (KZN)
Ms Dolly Nyasulu	Advanced Midwife	SANC member, Programme director PATH, formerly educator and in the DEPAM programme
Prof Bob Pattinson	Obstetrician	Director : MRC Research Unit for Maternal and Infant Care Strategies, University of Pretoria
Prof Hugh Philpott	Obstetrician (retired)	Retired from UKZN, Consultant to Limpopo
Dr Hannes Steinberg	Medical officer	Family Practitioner responsible for MOU in Bloemfontein and MCH outreach in Free State province
Dr Anna Voce	Public health practitioner	UKZN, Consultant to Limpopo

This should assist people in answering questions 7 and 8 on the CVI. Question 7 asks if the jury includes practitioners. One could regard these as the clinical practitioners who do the observations are record these onto the partograph, but it could also mean the practitioner who is the clinical manager/supervisor, whose practice includes audit and quality assurance.

Appendix 3 : Content Domain (CVI – Questions 5 and 6)

This deals with two aspects - the content of the actual variables, and the scoring of these variables. Given that there were 58 items on this instrument, they were relatively few items that elicited significant differences of opinion. What raised the most discussion, was the issue of mark allocation, weighting of specific concepts, and relevant importance of items, and these I will discuss separately.

Some people were uncertain about what factors contributed to Questions 5c (medico-legal requirements) and 5d (quality assurance measures). From a medico-legal point of view, it is always necessary to be able to identify any pre-existing or potential risks and to recall accurately events that have happened, their timing and to whom they happened. Therefore the items which relate to name and age, identification of risk factors, correct use of time, would contribute to enhancing the credibility of the completed partograph in medico-legal terms. As far as quality assurance measures are concerned, the existence of an auditing instrument *per se* is important, as well as establishing its purpose and the way in which it will be used.

It appears that the word 'record' carries some ambiguity as this is used both as a noun (in the case of a document, e.g. partograph and antenatal clinical card) and as a verb (in the case of writing an observation onto the relevant document). This has implications for the definitions associated with the scoring system as well as a number of items themselves, e.g. Item 1. I have tried to correct these where I have seen them, but some might have been missed.

Identification and description of the variables

As a result of responses received, I have made a number of changes which assist in describing the variables more accurately. This has also resulted in a couple more items where previously two or three variables were included in one item, thus introducing uncertainty in interpreting this item. This would probably also lead to a decrease in reliability. All changes are indicated as 'track changes' on the revised partograph checklist. One item (number 40 on the original checklist) has been omitted as there were number of responses indicating this direction. (If you object to this decision, please motivate on the Content Validity Index (CVI) assessment form why it should be retained.) As a result of these additions and deletions, the revised partograph checklist has a new set of numbering.

The specific responses / discussions about the variables can be found in Appendix 4 which is attached to the Round 2 Response form.



University of Cape Town

Division of Nursing and Midwifery

School of Health and Rehabilitation Sciences

Anzio Road, Observatory, 7925

Tel : (021) 406 6449

Fax : (021) 406 6323

eMail : Sheila.Clow@uct.ac.za

19 January 2009

Dear

Labour record review checklist validation exercise - Response to judges after Round 1

Thank you so much for your participation in this research project. I really appreciate all the carefully considered comments that I have received. In many cases there has been quite good degree of alignment in the comments. However, there are some issues that need to be considered further, in order to get a greater degree of consensus, and thus validation of the instrument / checklist.

I know there were some difficulties in completing the documentation that I sent you due to a lack of information or clarity, and for this I apologise. I appreciate people taking the time to engage with me during this process so they would be able to participate. I do feel it is necessary to clarify various aspects of this process to everyone and hope that this will facilitate getting this phase resolved / finalised. The responses required in this round are set out in the "Round 2 response form" with the necessary explanations included in its attached appendices.

Purpose of the checklist

Given our different responsibilities and practice demands, we are likely to have different ways of looking at this checklist and its utility. In terms of this exercise, I need to clarify what I need to achieve, and put it in the context of other potential uses.

The study for which this checklist is required has been in process in the Western Cape since 2004, and therefore the requirements that have been put into the Western Cape partograph have been included, e.g. fetal heart observations before and after a contraction, and 4 hourly comprehensive assessment. Once I started getting responses back from judges I realised that, having not sent a copy of the current partograph in use in the Western Cape, people would have been at a disadvantage. I am therefore including this now for your information. This partograph was re-launched with a full set of guidelines for managing labour developed by the Western Cape maternity guidelines group. This was dated 2004 and so the definitions used are those which were in that guideline and which were in operation at the time of the data being collected in 2004,

2006 and 2007. In 2007 the consolidated national guideline for intrapartum care was distributed, but would not affect the data being collected.

The prime purpose of this checklist is for research purposes, as we currently do not have a checklist to measure the completion of partographs that has been formally validated and tested for reliability. Therefore the focus of this exercise is on the checklist itself and not on the completion of the partograph, although obviously the two are closely linked and should inform each other. Once the checklist has been developed and tested, this will be used to evaluate how well the partograph is completed.

There is an underlying assumption that without a correctly and fully completed partograph, one is at risk of not having the necessary information to make a clinical decision. At present, with poorly completed partographs, it is not necessarily possible to know what has happened without recourse to other clinical notes. A successful outcome of this exercise will be to have a standard, valid, and reliable auditing checklist, that could be used in a constructive way to increase the complete and correct use of the partograph. One would be able to highlight areas of deficit and provide a standardised platform for future action, e.g. educational intervention and practice evaluation.

This is a first step, and other steps can follow once we get this right.

It would be a bonus for this checklist to be simple / clear enough to be used for ongoing auditing purposes by clinical managers and supervisors and self-directed clinicians, but at present this would be a secondary consideration. However the wording of the checklist is such that it should be able to accommodate changes in specifics that might occur from time to time as guidelines change. As one moves towards adopting a standard national document for labour recording, it would be ideal to have a standard national document for auditing labour records. The need for a quality assurance programme has been highlighted in the Saving Mothers' reports over the last number of years (Dept. Health 1999; 2002; 2006). It is hoped that this checklist will contribute towards the quality assurance programme for intrapartum care.

I have used comments received in Round 1 to clarify the purpose of having such a checklist, and have included this with the preliminary comments on the labour record review checklist.

The purpose of having such an instrument is :

- ◆ to stimulate optimal observations and recording of labour
- ◆ to manage intrapartum care according to clinical guidelines
- ◆ to identify areas of weakness and strength in clinical practice and management, in order to inform in-service and other educational programmes
- ◆ to support quality assurance programmes by having an instrument to conduct a clinical audit, and ultimately
- ◆ to improve patient care and outcomes

Research process to establish validity of this checklist

Validity, the extent to which the instrument measures what it is intended to measure, is a crucial factor in the choice or development of an instrument. Content validity is the determination of the content representativeness / content relevance of the elements / items of an instrument by the application of a two-stage (development or judgement) process (Lynn 1986). The developmental stage, when modifying an existing instrument, consists of identification of the full content domain, and the judgement stage relates to the judgement of the content validity of each item as well as the instrument as a whole.

In Round 1 I sent you a Content Validity Index assessment form (CVI), based on the work of Lynn (1986) and Yaghmaie (2003). There are three aspects considered:

- development and modification
- content domain
- the jury

I decided to stick as closely as possible to Lynn's CVI for the validation phase of this project, although I realise that there are different dynamics when modifying an instrument/checklist as opposed to developing one *de novo*.

Development and modification (CVI - Questions 1-4)

The process followed is presented as **Appendix 1** to the Content Validity Index assessment form (attached).

The jury (CVI - Questions 7 and 8)

The process followed is presented as **Appendix 2** to the Content Validity Index assessment form (attached).

Content domain (CVI - Questions 5 and 6)

The process followed is presented as **Appendix 3** to the Content Validity Index assessment form (attached)

Where comments were made and there were no conflicting positions, these changes have been captured directly onto the revised labour record review instrument, which is attached.

As far as finalising *specific variables* on the instrument, in order to try and consolidate the various contributions, I have summarised these in **Appendix 4** attached to the Round 2 Response form. There were some issues raised about the *number of variables* which relates to the overall construction of the instrument. The response and discussion is in **Appendix 5** attached to the Round 2 Response form. The issue that raised the most debate was the *mark allocation and scoring*. I have attempted to distil each of the points that were raised and put forward some proposals for consideration. These responses and discussion points are presented in **Appendix 6** attached to the Round 2 Response form.

In summary (and I apologise for the lengthy documentation, but I hope it will clarify certain points) (!), please give me your responses in the following ways :

1. Complete the attached Round 2 Response form to address specific items raised in this round as set out in the related appendices.
2. Complete the attached Content Validity Index (for the second time), having considered the clarification in the related appendices. There is space in this document to make specific comments on any items which you believe need amending

I would appreciate receiving your response by **12 February 2009**. Please send this both to me and to my research assistant, Cathy de Groot at degroot@babybuy.co.za. If there are any queries, I suggest you email me as I have a very full teaching schedule at the beginning of the year and am unlikely to be at the end of the phone if you call. Let me know what would be a suitable time to return your call.

Many thanks for your contribution thus far. I do hope that this round will proceed more smoothly and perhaps reach a point of sufficient consensus to enable progress.

Wish best wishes for a happy and fulfilling 2009

Kind regards

Sheila

Sheila Clow

References

- ◆ Department of Health 1999 *Saving Mothers : Report on confidential enquiries into maternal deaths in South Africa 1998* Department of Health : Pretoria
- ◆ Department of Health 2002 *Saving Mothers : Second Report on confidential enquiries into maternal deaths in South Africa 1999-2001* Department of Health : Pretoria
- ◆ Department of Health 2006 *Saving Mothers : Third Report on confidential enquiries into maternal deaths in South Africa 2002-2004* Department of Health : Pretoria
- ◆ Lynn MR, 1986 Determination and quantification of content validity. *Nursing Research* Vol 35(6), 382-385.
- ◆ Pettersson KO, Svensson M-L, and Christensson K 2000 Evaluation of an adapted model of the World Health Organisation partograph used by Angolan midwives in the peripheral delivery unit. *Midwifery*, 16, 82-88.
- ◆ Yaghmaie F 2003 Content validity and its estimation. *Journal of Medical Education* Vol3(1) 25-27

Enclosures

- ◆ Content Validity Index assessment form + related appendices 1,2 and 3
- ◆ Round 2 Response form + related appendices 4,5 and 6
- ◆ Western Cape Partograph 2004 + comprehensive assessment form (printed on reverse of partograph)
- ◆ Partograph checklist version 2

Round 2 Response form**Labour record review checklist development****PI : Sheila Clow, UCT**

Judge's name :

Date of response :

Item and checklist construction (see Appendix 4 attached)

No	Item	Response	Decision	Motivation
6	Were the admission findings checked & counter-signed by advanced midwife (or doctor or experienced RM, if no ADM)?		Include Delete	
55	Were these assessments checked by an advanced midwife (or doctor or senior midwife?)		Include Delete	
15	Institution where the monitoring is taking place		Include Delete	
18	Duration of labour entered in each column?		Include Delete	
41	Level of presentation above the pelvic brim (abdominal)		Include Delete	
47	Action to be taken		Include Delete	

1. With the revisions/decisions made, please score the content validity of the entire checklist on a scale of 1 - 4 where

1 = least agreement, and 4 = strong agreement

1	2	3	4

2. Does the ordering need to change in any way?

Yes	No

Give motivation :

.....

.....

For questions 3 and 4, see Appendix 5 attached

3. a) Should the number of variables be reduced?

Yes No

Comment /motivate:

.....

.....

.....

- b) If the number of variables should be reduced, which should be deleted?

Please state item number and description :

.....

.....

.....

4. If you have proposed deleting some items, will this affect the content validity of the entire instrument? Answer this by re-scoring the content validity of the entire instrument using a Likert scale of 1 – 4, where

1 = least agreement, and 4 = strong agreement

1

2

3

4

Any other comments on the content of items can be made directly onto on the Content Validity Index assessment form

- 1 = 1st choice; 2 = 2nd choice; 0 = I do not accept this option

Option 1 Option 2

- b) Propose a more suitable option, if you wish :

$$1 = \dots\dots\dots 1/2 = \dots\dots\dots 0 = \dots\dots\dots$$

- c) Please indicate the level of agreement for the following conditions where “not applicable” would be valid for certain sections of the checklist (denoted in brackets), using a Likert scale of 1 – 4, where 1 = least agreement, and 4 = strong agreement

Mother unbooked (some data from antenatal record)

Intrauterine death (fetal observations)

Admission to time of birth less than 4 hours (comprehensive assessment)

Other conditions – please specify condition and section which would be not applicable :

6. In terms of weighting of items please indicate which of the following two options you choose, by circling the relevant option :

All items carry the same weight

OR

Certain items carry a greater weighting

a) IF the consensus favours weighting certain items, please indicate which items should be weighted more heavily, and by how much

Number	Item description	Weighting, e.g. x2, x3

Thank you!

Appendix 4 : Responses / discussion ensuing from the responses for specific variables.

The information that follows informs the responses requested in the response form above.

[Note : Items with square bullets are the opinions of judges. Where a decision has already been made due to a clear direction from round 1 this is indicated with an arrow. Where I am engaging in the discussion or trying to distil a number of voices, these are labelled alphabetically.]

Items 6 and 55 are similar and contain the same principle, and so are presented together.

Item 6 : Were the admission findings checked and counter-signed by an advanced midwife (or doctor or experienced midwife, if no ADM)? AND

Item 55 : Were these assessments checked by an advanced midwife (or doctor or senior midwife?)

- How well established is this [practice of checking and countersigning] ? This is essential to good practice. Fundamental problem requiring to be addressed nationally.
- In many clinics where antenatal care is conducted and deliveries occur, there is only one registered midwife therefore it is not practical to get counter-signatures on the ANC card or delivery record.
- At night in district hospitals only 1 RN on duty, no senior / advanced midwife available to counter-sign.

There appear to be two approaches to this issue :

- a. The "gold standard" is to have all observations/assessments and decisions to be checked by a senior clinician. The fact that there may be no other clinician on duty, let alone a senior/experienced clinician, in itself should be a matter of audit. Including this in the checklist would enable the systematic gathering of objective data that can be presented to managers and policy makers, indicating the chronic lack of clinical leadership, which impacts on quality care.
- b. The reality is that in many situations there are no other clinicians available to do the checking and therefore it is pointless including this item as it is not achievable

Item 15 : Institution where the monitoring is taking place

- National partogram will be part of maternity case record (booklet) so the information on name/institution will not be entered on the partogram but in the front of the booklet
- Patient might be transferred during AN or IP periods and so the institution should be reflected

Item 18 : Duration of labour entered in each column?

- less important for weighting purposes
- a. Revise / omit? In view of item 13(duration of labour on admission), is this necessary? It can cause confusion when some interpret this as 'from the time of admission' and others 'from the start of labour' based on history given.

Item 41 : Level of presentation above the pelvic brim (abdominal)

- Change to "level of head"
- a. Where the presentation is other than cephalic, referring to 'head' rather than 'presentation' would be incorrect

Item 40 (original) : Station (vaginal)

- I do not regard this as necessary
- less important for weighting purposes
- National guideline - suggest omitting 'station' of the head and only focus on 'head above pelvis'
- OMIT

Item 47 : Action to be taken

- The partograph is for observations only so would not include this question
- a. If it were necessary to do the next assessment before the routinely agreed 4 hours, where would this decision be recorded? without having to go to a separate page where the information about this important decision might be missed

Appendix 5 : Number of variables

The following voice of experience was included in the response to Round 1. "In our experience 25 questions is as many as most practitioners can manage to answer. If you want this to be practicable, 58 questions is too many."

I need to have your opinion about this. In making your response please consider the following :

- a. I believe that form should follow function.
- b. We need to include what we believe is necessary / desirable to meet the objectives of the instrument.
- c. In order to minimise ambiguity and increase reliability, my aim is to ensure that there is only one variable per item.
- d. By reducing ambiguity, where previously more than one variable was included in an item, this might make it easier for auditing purposes.
- e. If one wants to reduce the number of items, this will mean the deletion of an item rather than the combining of two or more variables in a single item.
- f. Can any items be deleted and retain clinical coherence / relevance of the checklist?
- g. Because I recognise that 58 items looks a lot more than 25 items, the next phase of this research (reliability testing) will compare the ease in completing the checklist against the existing 25 point Philpott and Voce instrument determined by time taken to complete these when evaluating the same sample of partographs. The degree of reliability of the two instruments will also be compared.
- h. Once this is formally evaluated, one can decide if simplification is required.

Appendix 6 : Mark allocation and scoring

1. When trying to measure both completeness and correctness, one has to consider a number of different scenarios. For example,
 - Complete : Does it mean that all observations are completely filled in at all time periods as specified in the guidelines? It will be impossible to analyse if every set of observations has to be scored. (When one has capacity and computer power, it is possible to manage a large database but this would put the checklist out of range for the clinical managers.)
 - Incorrect : Defined as using the wrong symbol, but what about when the correct symbol is used but is entered incorrectly?, e.g. time plotted incorrectly
 - Is completeness more important than correctness that the former scores a whole mark?

A couple of suggestions were made and I would value your responses to both of them so we can determine which has the greater acceptability.

Option 1 : most observations were correct = 1
 most incorrect or missing = 1/2
 never done = 0

Option 2 : correct = 1
 incomplete (not, incorrect) = 1/2
 missing = 0

There was a point raised about how a “not applicable” evaluation would affect the score. One could always give guidance on establishing the relevant denominator where there are items which are scored as not applicable. Perhaps one should try and “encourage” evaluators NOT to opt for a “not applicable” finding except in limited situations, e.g.

- (i) If the woman had not booked, the antenatal record would be absent, but history taking could elicit some of the necessary information
 - (ii) One could expect no fetal heart observations being done where there is an intrauterine death, yet the rest of the observations would still be required in terms of the maternal outcome;
 - (iii) One might not find a comprehensive assessment if the duration from admission to giving birth is less than 4 hours.
- Are there other circumstances where any portion of this checklist would be not applicable?

Depending on the responses received the definitions will need to be revised and will need to be explained well.

2. The expectation / interpretation of the scores elicited a number of responses :

- The scoring notes are not very explanatory. They just measure the presence or absence of an entry, but not the intention, cognitive ability, reasoning of management.
- Scoring is not balanced according to importance.

This raises the questions, What is the purpose of a score? Is it to ascribe value, and what does this value mean? Does it signify that something is done or how well something is done.

- a. If the score signifies that something is done, then one can use a criterion referencing approach where each item scores the same. This would be sufficient to identify areas of weaknesses and strengths.
- b. If the score signifies how well something is done, then one needs to look at ascribing value to items that are more important than others. This would require a weighting of some sort.
- c. How will a score be interpreted? Will there be a sense that a score above a certain threshold is acceptable or will guarantee a better outcome? I think that this last point is beyond the scope of the current research. However, once we do have a robust auditing instrument and are able to establish a significant dataset, it might be possible to develop some predictive scores in terms of outcomes.

The following example possibly illustrates the difference that could be made by choosing to apply or not to apply a weighting :

One might obtain a score of 70% while at the same time having a section of the data very poorly completed. If one only then looked at the global score, one would miss such problems.

3. If one were to consider a weighting system, this could give more emphasis to key issues, e.g. correct management, correct interpretation. Some of the comments supporting this approach were :

- Frequency in monitoring has huge implications for quality assessment and assurance. The proposed instrument gives 6 marks for frequency, and 21 for recording observations. A common problem with maternal deaths records is infrequent monitoring and delayed intervention.
- Frequency of observations needs to be weighted more heavily than at present.
- The reference to the antenatal history and identification of risk factors is very important
- One of the judges made comments on a number of the items, with suggestions about their comparative importance / weighting.

Appendix 5 Partograph utilisation checklist and guidelines

Instrument objectives

- ◆ To measure the completeness of the partograph
- ◆ To measure strengths and weaknesses in the use of a partograph

The purpose of having such an instrument is

- ◆ to stimulate optimal observations and recording of labour
- ◆ to manage intrapartum care according to clinical guidelines
- ◆ to identify areas of weakness and strength in clinical practice and management, in order to inform in-service and other educational programmes
- ◆ to support quality assurance programmes by having an instrument to conduct a clinical audit, and ultimately
- ◆ to improve patient care and outcomes

Instruction to evaluators

- ◆ Please circle the scores on the scoring sheet provided
- ◆ Please fill in the administrative and contextual information at the beginning and end of the checklist

Scoring notes

- ◆ 1 = Mostly complete
 - ◇ For hourly or ½ hourly observations more than 75% of expected observations should be entered, i.e. for every 4 expected observations at least 3 should be recorded
 - ◇ For 4 hourly observations all observations should be recorded
- ◆ ½ = Incomplete
- ◆ 0 = Missing / No record
- ◆ N/A for any non-applicable variable (see point 5 below)
- ◆ When scoring individual items, e.g. maternal blood pressure, fetal heart rate, cervical dilatation, one should score the completeness of that item over the entire time period.
- ◆ This evaluation exercise is concentrating on the completeness of the observations (as a proxy for the standard of care) and is not analysing the clinical significance of the findings. Therefore, some findings might be reflected as nil or NAD yet the score will be 1, not 0, because the observation has been made and documented, e.g. 'risk factors nil' or 'urine 100mls and NAD' would result in a score of 1.
- ◆ Referrals / transfers
If a patient is referred to a higher level of care (see point 17 below) or transferred to another facility at the same level of care, the scoring should cease at the time of departure from the sending institution.

Explanatory notes

1. In the introductory section
 - 1.1. If the woman is not in labour on admission, use the definition in point 2 below to determine when labour commenced during this hospitalisation episode.
 - 1.2. If the admission record reflects the cervical finding as 'multips os' or 'tip of finger', reflect this as 1cm
 - 1.3. If there is no partograph, indicate N/R in the introductory section, where the date and time of commencement is required as well as responding 'No' to 6a and N/R to 56c.
 - 1.4. For any other item in the introductory section for which there is no record, indicate this as N/R.
2. Labour is defined as regular contractions AND one of the following : cervical changes, rupture of membranes, 'show'. (National Department of Health, 2006)
3. Latent phase refers to the early stages of labour when the cervical dilatation is 3cm and below (Western Cape Provincial Department of Health, 2002) with incomplete effacement. During the year in which records were collected there was some move to adopt the national guideline of 4cm cervical dilatation as the cut-off.

- 3.1. Where a cervical finding of 4cm is plotted in the latent phase this can be accepted.
- 3.2. Frequency of observations (items 19, 24, 31, 34 and 39). Because different observations have different frequencies, and in order to pinpoint where the strengths and weaknesses are, there are two sets of questions regarding frequency.
 - 3.2.1. Maternal and fetal observations hourly
 - 3.2.2. Vaginal assessments 4 hourly unless otherwise indicated (see also 4.3 below).
 - 3.2.3. If observations are done more frequently than indicated these should score 0, not 1.
4. Active phase refers to cervical dilatation of 4cm to full dilatation of the cervix (Western Cape Provincial Department of Health, 2002) and above.
 - 4.1. Frequency of observations (items 20, 25, 32, 35 and 40). Because different observations have different frequencies, and in order to pinpoint where the strengths and weaknesses are, there are two sets of questions regarding frequency
 - 4.1.1. Maternal and fetal observations ½ hourly
 - 4.1.2. Vaginal assessments 4 hourly, unless otherwise indicated
 - 4.1.3. If observations are done more frequently than indicated these should score 0, not 1.
 - 4.2. It is possible that the score in the active phase is N/A even if there has been a score in the latent phase, e.g. a single vaginal assessment in the active phase (see 5.2.1 below).
 - 4.3. Do not 'double count' in the case of an observation relating to latent phase also being transferred to the active phase of the partograph. Regard this as the completion of the frequency of observations from the latent phase, and only consider it for scoring purposes if a subsequent assessment is done completely and at the correct frequency.
5. 'Not applicable' / N/A.
 - 5.1. In most cases all **variables** are applicable. Evaluators are encouraged to limit the 'not applicable' finding. Under the following circumstances, all or part of the partograph will be considered not applicable :
 - 5.1.1. An intrauterine death (items 16-25, and 52)
 - 5.1.2. Moulding (item 21) and caput (item 23) do not occur in a breech presentation, therefore are not applicable in this situation!
 - 5.2. **Frequency** of observations
 - 5.2.1. Can only be determined when more than one set of observations would have been expected. For example, if only one internal assessment could have been done within the time period the frequency score would be not applicable.
 - 5.2.2. When scoring the items which relate to frequency of observations and adherence to the guideline, one is looking for the intended pattern of observations which either is or is not met. For example, in the active phase the guideline states that ½ hourly observations should be done, yet some centres seem to accommodate hourly at the most. In this situation this would score 0. There is no room for 'not applicable' unless that entire item is not applicable
6. Document - A verb referring to the process of written representation of observations;
whereas
Record - A noun referring to the documentation
7. Time of admission in labour / commencement of partograph (item 13). All records selected for this study indicate that the cervical dilatation was ≤6cm and that this labour resulted in a birth. However, not all women were in labour (see note 2) at the time of admission. For this item if the partograph was only commenced after 6cm cervical dilatation and the time was recorded then this should score ½. Those commenced on the partograph earlier and the time is recorded should score 1.
8. Correct plotting of time (item 14). This means that :
 - 8.1. each block represents a consecutive hour
 - 8.2. the cervical dilatation in the active phase should be commenced on the alert line
9. Observations plotted in the correct section of the partograph (item 15)
This refers to observations relating to the latent phase of labour being plotted in the section

of the partograph relating to the latent phase, and observations relating to active phase of labour being plotted in the section of the partograph relating to the active phase. For example, if a cervical observation of 5cm dilation is made at 09:00 and this (and the related observations at that time point) is plotted in the latent portion of the partograph this would score 0.

10. Fetal heart rate (items 16 and 17). If there is only one observation made, score this as a 'before contraction' observation. Score the 'after contraction' fetal heart observation as missing, i.e. 0.
11. Effacement (item 30). Some might be recorded as percentages. As this instrument is scoring completeness only and not correctness, you should score this item in relation to the appropriate number of observations of effacement recorded. There is another part of this study which measures incorrect representation of effacement.
12. Urine observations (items 41 - 45)
 - 12.1. As these occur at non-standard intervals, assume that there is nothing to record, i.e. 'not applicable'. If there is evidence of some component of the urinary observations present, e.g. if proteinuria is recorded and nothing else, one can assume that urine was passed and not recorded, therefore protein would be scored as 1 and all other components – volume, ketones, blood and glucose would score 0.
 - 12.2. Where 'toilet' is recorded, this should be scored as 0 as this yields no clinically useful data.
13. Action (item 48) – Any appropriate action recorded at any time.
14. Signature (item 49) – If it is present but not legible, it should be regarded as if it is absent.
15. Standard assessment and clinical management form (item 49a). This should be printed as a *pro forma* document on the reverse of the partograph.
 - 15.1. Where there was no such document in the records this is indicated on the reverse of the partograph as 'nil printed on reverse'. In this situation one would record "No" for item 49a, and proceed to item 55a.
 - 15.2. Where there was such a document but not filled in, this is indicated on the reverse of the partograph as 'nil'. In this situation one would record 'Yes' for item 49a, 0 for items 50-53 and N/A for item 54
16. Overall assessment and further management (item 53) – Decisions and actions taken as a result of a comprehensive assessment done at the appropriate time.
17. Higher level of care (item 55)
 - 17.1. In the case of an MOU this refers to a transfer to a district hospital – Hottentots Holland (Helderberg) Hospital or secondary hospital, Eben Dönges Hospital, Worcester
 - 17.2. In the case of a primary level district hospital this refers to a transfer to a secondary or tertiary hospital – Eben Dönges Hospital, Worcester; Paarl Hospital; (New) Somerset Hospital; Tygerberg Hospital
18. Completeness of observations (item 56c)

In a number of instances observations immediately prior to the time of birth are missing, without any explanation. The information for this item is generated by the evaluator.

 - 18.1. If observations are missing and birth/transfer occurs within 1½ hours of the last observation, record 'Yes'
 - 18.2. if observations are missing for a longer period than 1½ hours prior to birth / transfer record 'No'
 - 18.3. if observations are not missing during this period, record 'N/A' (not applicable)
 - 18.4. if there was no partograph used, record 'N/R' (no record)
19. Completeness of observations (item 60)

If only one observation could be expected to be done (in terms of the frequency specified) from the time the partograph was started until the birth had taken place (either within 1 hour if observations commenced in latent phase, or within ½ hour if observations commenced in the active phase) then answer this as 'Yes'.

Partograph utilisation checklist

Evaluator ID : Give the initials of your name :

Date of evaluation :
yy – mm - dd

Partograph ID :

Date of admission :
yy – mm - dd

Time of admission : Cervical dilatation^{1.1} : cm

Date partograph started^{1.2}:
yy – mm - dd

Time partograph started: Cervical dilatation : cm

Date of birth :
yy – mm - dd

Time of birth :

Type of birth recorded

NVD

Forceps

Caesarean Section

Vacuum

Other

Please specify

Partograph 6a Was a partograph used for this patient? Yes ☐ No ☐ **Proceed to item 49a** →

Identification, History and Risk assessment

Is the following information documented⁶ on the partograph?

7	Gravity	1		0	
8	Parity	1		0	
9	Risk Factors	1		0	
10	Time of rupture of membranes	1		0	
11	Duration of labour on admission	1		0	
12	Date of assessment	1		0	
13	Time of admission in labour / commencement of partograph ⁷	1	½	0	
14	Time plotted correctly on the graph? ⁸	1		0	
15	Observations plotted in the correct section of the partograph ⁹	1		0	

Fetal condition

Are the following observations documented on the partograph?

16	Fetal heart rate before a contraction ¹⁰	1	½	0	n/a
17	Fetal heart rate after a contraction ¹⁰	1	½	0	n/a
18	Presence / type of decelerations	1	½	0	n/a

Did the frequency of observations adhere to the guideline for

19	the latent phase? ^{3.2.1}	1		0	n/a
20	the active phase? ^{4.1.1}	1		0	n/a

Are the following observations documented on the partograph?

21	Moulding ^{5.1.2}	1	½	0	n/a
22	State of membranes / liquor	1	½	0	n/a
23	Caput ^{5.1.2}	1	½	0	n/a

Did the frequency of observations adhere to the guideline for

24	the latent phase? ^{3.2.2}	1		0	n/a
25	the active phase? ^{4.1.2}	1		0	n/a

Labour progress

Are the following observations documented on the partograph?

26	Presentation of fetus	1	½	0	n/a
27	Position of fetus	1	½	0	n/a
28	Dilatation of the cervix	1	½	0	n/a
29	Level of head above the pelvic brim (abdominal)	1	½	0	n/a
30	Effacement of the cervix ¹¹	1	½	0	n/a

Did the frequency of observations adhere to the guideline for

31	the latent phase? ^{3.2.2}	1		0	n/a
32	the active phase? ^{4.1.2}	1		0	n/a

Was the following observation documented on the partograph?

33	Number and duration of contractions in 10 minutes	1	½	0	n/a

Did the frequency of observations adhere to the guideline for

34	the latent phase? ^{3.2.1}	1		0	n/a
35	the active phase? ^{4.1.1}	1		0	n/a

Maternal Condition

Are the following observations documented on the partograph?

36	Blood pressure	1	½	0	n/a
37	Pulse	1	½	0	n/a
38	Temperature	1	½	0	n/a

Did the frequency of observations adhere to the guideline for

39	the latent phase? ^{3.2.1}	1		0	n/a
40	the active phase? ^{4.1.1}	1		0	n/a

With each urine specimen provided, are the following documented on the partograph ? ¹²

41	Volume	1	½	0	n/a
42	Protein	1	½	0	n/a
43	Ketones	1	½	0	n/a
44	Blood	1	½	0	n/a
45	Glucose	1	½	0	n/a

Management of labour

Are the following documented on the partograph ?

46	Drugs given	1	½	0	n/a
47	Fluids administered	1	½	0	n/a
48	Action taken ¹³	1	½	0	n/a
49	Legible signature available for each entry ¹⁴	1	½	0	n/a

The 4 (or 2) hourly comprehensive assessment and management form

49a Was there a standard assessment and management form? ¹⁵ Yes ☐ No ☐ **Proceed to item 55a** →

Are conclusions for the following assessments documented? ←

50	Progress of labour	1	½	0	n/a
51	Maternal condition	1	½	0	n/a
52	Fetal condition	1	½	0	n/a
53	Overall assessment and management ¹⁶	1	½	0	n/a
54	Is the decision consistent with the clinical guidelines?	1	½	0	n/a

Contextual information

55a Does the record state if this patient was **referred** to a higher level of the health service? ¹⁷ ←

No	Yes

55b If yes, (i) where?

(ii) reason?

56 If the record was incomplete

56a was there any explanation documented on the partograph?

No	Yes

56b If yes, what? Ward busy

☐

Shortage of staff

☐

Other

☐

Please specify

56c If no, did birth / transfer occur within 1½ hours of last observation? ¹⁸

Yes	No	N/A	N/R

60. If the record was complete, was this because there was time to do only one only observation? ¹⁹

Yes	No	N/A



University of Cape Town

Division of Nursing and Midwifery

School of Health and Rehabilitation Sciences

Anzio Road, Observatory, 7925

Tel : (021) 406 6449

Fax : (021) 406 6323

eMail : sclow@uctgsh1.uct.ac.za

Research project : Improving the standard of intrapartum care by registered midwives in primary level public sector health facilities **Information sheet and consent form for mentors**

Dear Colleague

My name is Sheila Clow and I work in the Division of Nursing and Midwifery at the University of Cape Town. For the past few years I have been a member of the PGWC Maternity guidelines group which tries to address factors which contribute to maternal and fetal deaths in the Western Cape.

While we in the Western Cape are proud of the fact that we have the lowest levels in the country, we also know from the Saving Mothers and Saving Babies reports that there are many avoidable factors identified. In fact it has been shown that in the rural areas, intrapartum asphyxia and birth trauma are responsible for over 50% of all deaths of babies with birthweights over 2500g. A key recommendation of both these reports is to ensure that each site conducting births has the necessary equipment and protocols and that health care providers are appropriately trained to manage labour and are especially trained in the use of the partogram.

With the support of the Maternal Child and Women's Health sub-directorate, I aim to support this key national recommendation. I plan to undertake a study on improving the standard of intrapartum care by registered midwives. Two interventions are planned. One is to have a focussed update on training of registered midwives in relation to intrapartum care, which I will conduct. However, we know that improving knowledge and skills in themselves does not necessarily lead to a change in practice and an improvement in health outcomes. Something else in the system is required.

The second intervention therefore is planned as a mentoring type of programme, the training for which I will conduct. This would equip senior clinicians/nurse managers to support, guide and strengthen the clinical practice of the colleagues in the intrapartum setting, and to develop skills for continuous quality assurance through the use of clinical audit. Two registered nurse/midwives have been selected from each of the sites which have been randomly selected to receive these interventions.

What does this mean for you?

You would undertake a 4 day (on-duty time) training programme which will

- ◆ give you an update on maternal, fetal and labour assessment
- ◆ introduce you and equip you for mentoring
- ◆ introduce you to clinical review

Thereafter you would be expected (along with your mentor colleague from your workplace) to introduce a mentoring system in your workplace.

You will not be tested formally on any aspect of this training programme. However the expectation is that the effect of this mentoring will improve the standard of care and clinical decision making of the registered midwives at your site, and this will be evaluated periodically in terms of auditing of partographs and analysis of maternal and perinatal data, in the year following the training.

Eight other sites (distributed in your region and one other rural region) will also have this project running. However eight sites in the two regions will not have this programme. You are requested therefore not to share material from this project with colleagues in those sites, until the data collection period is completed (approximately 1 year).

If you have any questions or aspects that need clarification, please do not hesitate to contact me at (021) 406 6449 (voicemail facility) or 083 659 5266. If you have any queries about the research you may contact my supervisor, Prof George Swingler at (021) 658 5324 or the chair of the UCT Faculty of Health Sciences Human Research Ethics committee, Dr Marc Blockman (021) 406 6492.

I trust that this will be a valuable experience for you, and will enable you to more effectively meet the quality assurance requirements of your post. These activities will be recognised as part of your performance review as they form part of your responsibilities.

Thank you

Yours sincerely

Sheila Clow

Consent form

I, ... have read the information sheet relating to the study, "Improving the standard of intrapartum care by registered midwives in primary level public sector health facilities". I have had the study explained to me and my role in it. I have had the opportunity to ask any questions I had, and these have been answered to my satisfaction. I freely agree to participate.

Print name : Sign :

Witness name : Sign :

Date :

Appendix 7 Philpott and Voce (2005) Labour record review tool

Examine 20 consecutive labour records. Do this after labour has been completed.

Use the following scoring system for each patient's record:

For each 'yes' answer, score 1 point on the summary sheet provided.

You can give half points where the information is incomplete.

Admission assessment form

1. Is there evidence that the health worker has reviewed and summarised the ANC record and listed the maternal and fetal risk factors?
2. Check the items on the admission form. Are all completed?
3. At the end of the form, is there a decision on diagnosis and management?
4. Were the admission findings checked and counter-signed by an Advanced Midwife (or doctor or experienced midwife if no ADM available)?

Labour graph

5. Is the list of risk factors recorded at the top of the labour graph?
6. Has the fetal heart rate been recorded half-hourly?
7. Has the state of the liquor (as recognised by a pad check) been recorded at least 4-hourly?
8. Has the degree of moulding been recorded when a P.V. has been done?
9. Have the contractions been recorded half-hourly?
10. Has the cervical dilatation been recorded at least 4-hourly during the Latent Phase and at least two-hourly in the Active Phase.
11. Has the cervical dilatation been plotted in relation to the lines drawn for the Latent and Active Phases, and for the Alert and Action Lines?
12. Has the level of the head in relation to the brim of the pelvis been recorded at least 4-hourly since admission?
13. Have the maternal BP and pulse been recorded at least hourly?
14. Have the maternal temperature and urinary output been recorded at least 4-hourly?
15. Is there a record of drugs and IV fluids given?

Management of Labour Form (On a page separate from the Labour Graph)

16. Is this recorded after doing each vaginal examination, or at least 4-hourly?
17. Is the summary of fetal condition recorded?
18. Is the summary of labour progress recorded?
19. Is the summary of maternal condition recorded?
20. Is the decision on further action recorded?
21. Is the time of next intended review stated?
22. Were these assessments checked 4-hourly by an ADM (or doctor or senior midwife)?

The assessment of the newborn

23. Has this form been completed?

Final Summary of Labour

24. Has this form been completed?
25. In the third stage of labour, is there a record that active management was carried out?

Appendix 8

Intervention package

Appendix 8 supplements the information contained in 3.7.4.1 (Appendix 8a) and 3.7.4.2 (Appendix 8b).

Appendix 8 a : The first table (2 pages) outlines knowledge, attitudes and skills required for each identified outcome, and the planned activity to address this. The second table (2 pages) reflects the training timetable planning grid over the four day period.

Outcome	Knowledge	Attitude	Skills	Activity
Establish and maintain a learning environment	Understand ♦ Stages of change ♦ different learning styles ♦ System issues	♦ Partnership ♦ Co-operation	♦ Encourage reflection and critical thinking in the labour ward staff ♦ Establishing / negotiating collegial relationships ♦ Reduce threat ♦ Set agenda	♦ Roleplay : Assessing readiness / motivation (importance/confidence) ♦ Reflective exercise
Understand the purpose and role definition of a clinical facilitator	Role clarification –clinical facilitator vs line manager	Partnership	Develop this role personally / 'de-role'	♦ Assessment of statements that might be made by either role player ♦ Reflective exercise
Be able to manage individual and group feedback	♦ Styles of communication ♦ Reducing resistance ♦ Information exchange	♦ Positive regard ♦ Valuing of people	Interpersonal and facilitation skills ♦ Establish rapport ♦ Develop open communication ♦ Negotiation ♦ Enhance motivation ♦ Being encouraging ♦ Reduce resistance ♦ Information exchange ♦ Reduce threat	♦ Develop a strategy for clinical feedback / reporting ♦ Roleplay ♦ Reflective exercise

Outcome	Knowledge	Attitude	Skills	Activity
Be able to conduct a clinical review meeting	◆ Scope of meeting ◆ Styles of communication	◆ Inclusivity ◆ Non-judgmental ◆ Encouraging	Facilitation skills (as above)	Roleplay
Become familiar with the guidelines for intrapartum care	◆ Guidelines for practice ◆ legislation ◆ clinical guidelines	Commitment to adhering to clinical guidelines	Assertiveness	◆ Roleplay ◆ Clinical case scenarios
Supervise and monitor intrapartum care	Normal progress of labour	Non-judgemental Encouraging	Clinical assessment on ◆ risk status ◆ progress of labour ◆ fetal assessment	◆ Case scenarios / partograph ◆ Clinical simulations ◆ Reflective exercises (Learning style inventory)
Conduct a clinical audit of intrapartum care using an appropriate tool	Implications of recording of labour progress and clinical management decisions	Attention to detail	◆ Utilisation of partograph ◆ Use of Philpott and Voce instrument⁷¹	◆ Partograph audit utilisation exercise (pairs)
A programme of ongoing support of clinical facilitators	Understanding sustainable systems	Co-operation	◆ Negotiation	Jointly develop an acceptable and sustainable strategy **

⁷¹ For the purposes of the audit and feedback aspect of clinical facilitation, the existing Philpott and Voce tool for auditing care in labour was used. It was recognised that this had limitations from a scientific point of view, but it was adequate for this purpose. It was anticipated that once the development and testing of the partograph utilisation tool was completed, it would not be difficult to introduce the revised tool for future auditing and quality assurance.

Clinical facilitator training timetable planning grid

	MONDAY	TUESDAY	WEDNESDAY	THURSDAY
08h00 - 10h00	1. Content Beginnings – introductions, expectations, ground rules ----- Process Creating a positive, safe, democratic, learning environment ----- Activity : Icebreaker Experiential	5. Content Process and progress of labour Partograph Risk status ----- Process Group dynamics Thinking skills ----- Activity : Case discussion Groupwork	9. Content Fetal monitoring – Applied Anatomy and Physiology Risk status ----- Process Being unfamiliar Reduce threat/resistance Problem solving ----- Activity : Groupwork Case discussion CTG tracings	12. Content Clinical review meeting Scope; Risk status Group dynamics / facilitation ----- Process Group dynamics, difficult scenarios Partnership Facilitation ----- Activity : Simulated clinical review meeting
10h30 - 12h30	2. Content Intrapartum challenges Introduction to partograph Stages of change ----- Process Readiness for change ----- Activity : Roleplay	6. Content Quality of partograph recording Audit tool Risk status ----- Process Interpersonal skills – giving feedback, guidance, non-judgmental, reducing resistance, encouraging ----- Activity : Roleplay Reflective journal	10. Content a) Fetal monitoring – technical setup b) Partograph utilisation audit ----- Process Technical skills Clinical thinking / application ----- Activity : a) Setting up CTG + trouble shooting; Use of doptone b) Groupwork : Case study	13. Content Implementation and sustainability strategies ----- Process Information exchange Reduce threat / resistance Increase safety ----- Activity : Site pairs / teams or else group work to plan; Reflective journal
	12h30 - 13h30	LUNCH	LUNCH	LUNCH

	MONDAY	TUESDAY	WEDNESDAY	THURSDAY
13h30 – 15h15	3. Content What is a mentor / clinical facilitator? Tasks/attributes of mentors Power relations ----- Process Communication skills : listening, non-judgmental, styles of questioning, reflecting, summarising De-role "boss" ----- Activity : Role play Reflective journal	7. Content Clinical assessment Applied Anatomy and Physiology ----- Process Clinical application / thinking Self-assessment ----- Activity : Clinical assessment exercise using models with clinical assessment tools	11. Content Systems issues Guidelines/legislation ----- Process Group dynamics Assertiveness Information exchange ----- Activity : Groupwork Reflective journal Climate assessment	14. Content Follow up strategy Evaluation Conclusion ----- Process Partnership, negotiation ----- Activity : Implementation strategy presentations Farewell exercise
	15h15 - 15h30	TEA	TEA	TEA
15h30 - 18h00	4. Content Learning styles Reflection ----- Process Reflection Self-assessment ----- Activity : Ref. journal Learning styles inventory Climate assessment	8. Content Clinical management Clinical guidelines Legislative framework ----- Process Thinking processes ----- Activity : Case discussion Use of intrapartum algorithm Climate assessment	CHILL / PREP / WALK etc..	HOME!!
	Blockword		Celebratory dinner!	

Implementation and sustainability strategies **

Part of the clinical facilitator training programme was to plan for the way forward where systems issues and sustainability were addressed. There was a session devoted to the mentors planning for this and the one region made a firm plan to meet with each other quarterly and to be in telephonic contact in between to support each other. The regional co-ordinator started by calling the PI before a planned quarterly meeting for insights or suggestions for discussion but the PI did not ask for any formal feedback from these sessions (in keeping with the pragmatic nature of this study).

The PI had contact with the mentors as part of the project management especially in regard to midwife testing. However when visiting for midwife testing 3 months after the mentor training, she did try to connect with them. This did not always occur due to logistical reasons. The contact with the regional co-ordinators was either at their own instigation or in relation to the project management. The intention was to try and have as realistic a process as possible, i.e. without an external resource injection that would then be withdrawn at the end of the study.

One of the requests from the mentors was for the test memoranda which could be used as a resource for workshopping. Once all the tests were completed in both regions for a specific timepoint, a copy of the test and memorandum was forwarded to the regional co-ordinators and mentors. There was no feedback about the overall performance within sites or regions during the 12 months, as the tests were only marked after all tests were completed in month 12 in region 2.

Different sites implemented their activities differently. One site had a small monthly partograph audit which was done by a different person each time; another had developed key messages that were displayed in the labour ward / room and referred to; another did individual feedback to staff members.

Appendix 8b Midwife Training Programme - Session Overview

The detail regarding teaching and learning approaches as well as the content, skills and attitudes are presented in Appendix 8a. The relevant portions were extracted for the midwife training programme and are reflected here.

	Day 1	Day 2
08h00 - 10h00	Session 1 ◆ Beginnings – introductions, expectations, ground rules ◆ Project overview ◆ Test	Session 5 Fetal monitoring + Risk status
	Tea	Tea
10h30 – 12h30	Session 2 ◆ Expectations, ground rules ◆ Process and progress of labour ◆ Partograph	Session 6 ◆ Fetal monitoring – technical setup ◆ Quality of partograph recording and risk status
	Lunch	Lunch
13h30 – 15h15	Session 3 ◆ Clinical assessment ◆ Clinical management	Session 7 ◆ Clinical management and clinical review meetings ◆ Systems issues
	Tea	Tea
15h30 - 16h30	Session 4 ◆ Quality of partograph recording + risk status ◆ Group dynamics ◆ Climate assessment	Session 8 ◆ Evaluation ◆ Conclusion ◆ Farewell

Appendix 9 Midwife knowledge and analysis evaluation

For each test there should be

- a) short questions on basic labour information including anatomy and physiology (15)
- b) a partograph record with questions relating to the case and clinical decision-making (25)
- c) a CTG record with questions relating to description and recognition of major patterns. (Interpretation not required) (5)
- d) Clinical evaluation (OSCE) with station/s measuring the following
 - a. Vaginal assessment + record as text
 - b. Abdominal assessment + anticipate vaginal findings
 - c. Fetal heart recognition and recording of rate (15)

With the following breakdown (as close as possible)

- ◆ Fetal (25)
- ◆ Maternal (20)
- ◆ Labour + progress (15)
- ◆ Risk assessment is incorporated in each of the above sections.
This will be able to be extracted for data analysis purposes

Note :

1. Mark weighting for all sections of test = ½ mark per point
2. For statistical analysis all ½ marks were converted into full units

Test	Section	Fetal (25)	Maternal (20)	Labour Progress (15)	(Risk)	Totals
Test 1	Short questions	5½	5	4½	(3)	15
	Partograph	6½	9½	9	(6)	25
	CTG	3	1	1		5
	OSCE	10 ½		4½		15
	Totals	25 ½	15 ½	19	(9)	60
Test 2	Short questions	1½	5	8½	(4)	15
	Partograph	7	14	4	(4)	25
	CTG	4	1		(1)	5
	OSCE	10½		4½		15
	Totals	23	20	17	(9)	60
Test 3	Short questions	5	8	2	(2½)	15
	Partograph	4	10½	10½	(1½)	25
	CTG	3	1	1	(1½)	5
	OSCE	10½		4 ½		15
	Totals	22½	19½	18	(5½)	60
	Section	Knowledge and Comprehension		Application, Analysis and Synthesis		
Test 1	Short questions	7½		7½		
	Partograph	8		17		
	CTG	2½		2½		
	OSCE	5½		9½		
	Totals	23½		36 ½		
Test 2	Short questions	7½		7½		
	Partograph	14		11		
	CTG	2½		2½		
	OSCE	5½		9½		
	Totals	29½		30½		
Test 3	Short questions	5½		9½		
	Partograph	8 ½		16½		
	CTG	2½		2½		
	OSCE	5½		9½		
	Totals	22		38		

Intrapartum project : Test 1

Notes : Mark allocation : ½ mark per point
Total marks : 45

Section A

Question 1 True / False

- a) The descent of the head should be established abdominally (½)
b) The average rate of cervical dilatation during labour for all women is 1cm per hour (½)

Question 2 Choose the correct answer/s

Maternal hypotension can be caused by

- A. The woman lying on her back
B. The woman lying on her side
C. Giving Ringer's Lactate intravenous infusion
D. Giving Neprosol intravenous infusion (1)

Question 3

Give 3 features of foetal stress (1½)

Question 4

What signs indicate that a woman is in established labour? (2)

Question 5

What 5 urinary observations should be done (in labour)? (2½)

Question 6

If a woman has symptoms of blurred vision, headache, epigastric pain

- a) what diagnosis should be suspected? (1½)
b) what observations should be done? (1)

Question 7

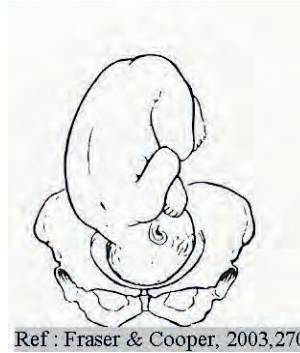
How many plusses of moulding are there when the foetal skull bones overlap but can be separated with digital pressure? (½)

Question 8

Match the presentation (alphabetical letter) with the presenting diameter (number), e.g. G6 (2)

A. Vertex	1. Bitrochanteric, 10cms
B. Cephalic, deflexed	2. Submento bregmatic, 9.5 cms
C. Face	3. Biparietal, 9.5 cms
D. Complete Breech	4. Occipitofrontal, 11.5 cms

Question 9



In relation to this diagram,

- a) Indicate the following
- i. Lie (½)
 - ii. Presentation (½)
 - iii. Position (½)
 - iv. Attitude (½)
- b) How do you expect this labour to progress, and why? (1)

[15]

Section B - Partograph

See the attached partograph.⁷²

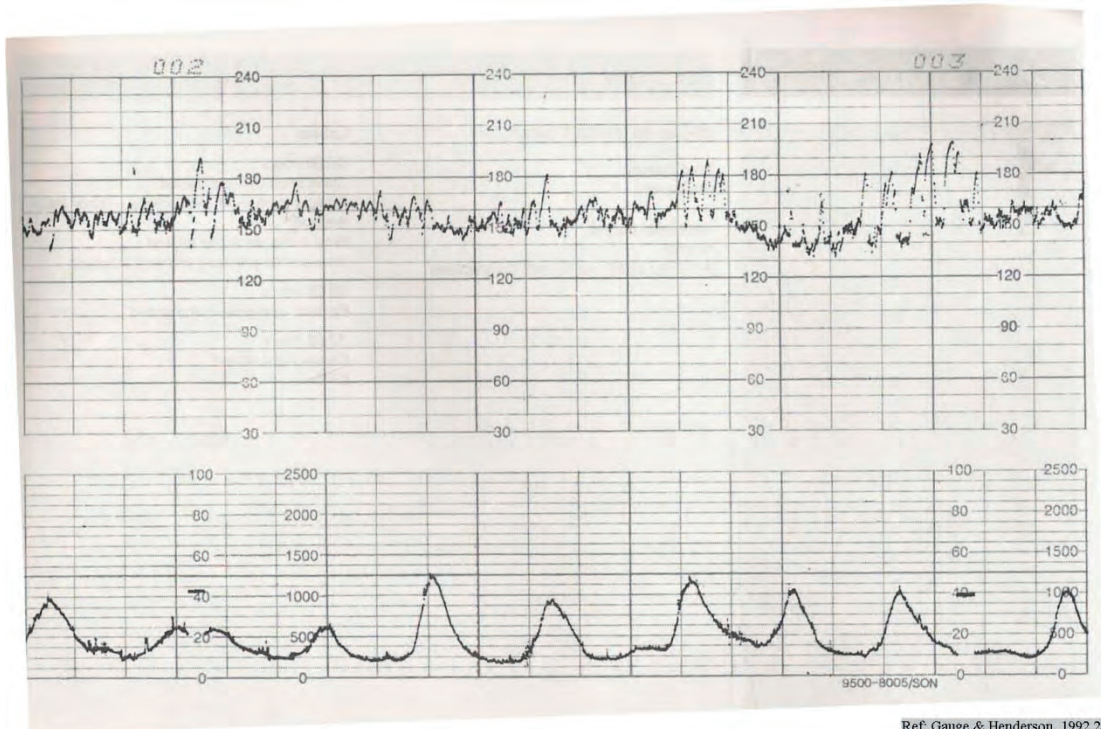
- (i) Make a comprehensive assessment of this labour. (16)
- (ii) Comment on management thus far. Would you have done anything differently and why? (3)
- (iii) Indicate appropriate management for the remainder of the 1st stage of labour. (4)
- (iv) What delivery planning should occur? (2)

[25]

⁷² The partographs used in the three tests did not use actual patient records but were simulated. Names were recorded on the 'patient data' as a precaution to minimise the possibility of an incorrect 'case study' being attached to a particular test.

Section C –CTG

Describe this CTG tracing and give your conclusion as to the status of this labour. (5)



Section D - Clinical skills

1. Abdominal assessment

- a) State what you will assess and how you will do this. (8)
- b) Using the models provided, record your clinical findings. (2 ½)

2. Vaginal / internal examination

- a) Indicate what factors can be determined during a vaginal examination of a woman in labour. (3 ½)
- b) Using the model provided, record the cervical dilatation and effacement. (1)

[15]

PARTOGRAM

Name: MRS JACOBS Risk Factors: NIL Spontaneous Labour ☒ Induced Labour ☐
 Grav.: 3 Parity: 2 Time of ROM: 12:30 10.06.06
 Order No.: 3394 Duration of Labour (OIA):
 Situation:

LATENT PHASE						ACTIVE PHASE					
TIME	6:00	6:00	6:00	6:00	6:00	11:00	12:00	13:00	14:00	15:00	16:00
DURATION IN HOURS	1	2	3	4	5	5	6	7	8	9	10
FETAL HEART RATE											
DECELERATIONS*	E/L/N					E/L/N					
MOULDING*	E/L/N					E/L/N					
LIQUOR*	E/L/N					E/L/N					
PRESENTING PART*	Vx					Vx					
CAPUT	-2					-1					
STATION	-2					-1					
APPLICATION	X					X					
CONTRACTIONS											
OXYTOCIN	UNITS: 1, 2, 2, 2, 4 RATE: 1, 2, 2, 2, 4					UNITS: 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4 RATE: 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4					
URINE	Volume: 200 Protein: 1 Ketones: 1 Blood: 1 Glucose: 1					Volume: 150 Protein: 1 Ketones: 1 Blood: 1 Glucose: 1					
TEMP - °C	36.1					36.5					
MANAGEMENT	Cervix postero-central					Cervix central - moderately well applied to it					
MEDICATION	Ringer's Lactate 1000ml with 21.0 Syntocinon					Spontaneous rupture of membranes					
IV FLUID											
SIGNATURE	PN					PN					
DESIGNATION	S/ABC					S/XYZ					

* (OIA) - ON ARRIVAL

* MOULDING:
OPM - OCCIPITAL PARIETAL
PM - PARIETAL MOULDING

* PRESENTING PART:
BRECH - B:
VERTEX - Vx

* LIQUOR:
I - INTACT
C - CLEAR
M - MECONIUM STAINED

* DECELERATIONS: E/L/N
E - Early
L - Late
N - No

Intrapartum project : Test 2

Notes : Mark allocation : ½ mark per point
Total marks : 45 + 20 (Clinical skills)

Section A**Question 1 True / False**

A maternal blood pressure of 150/90 is normal (½)

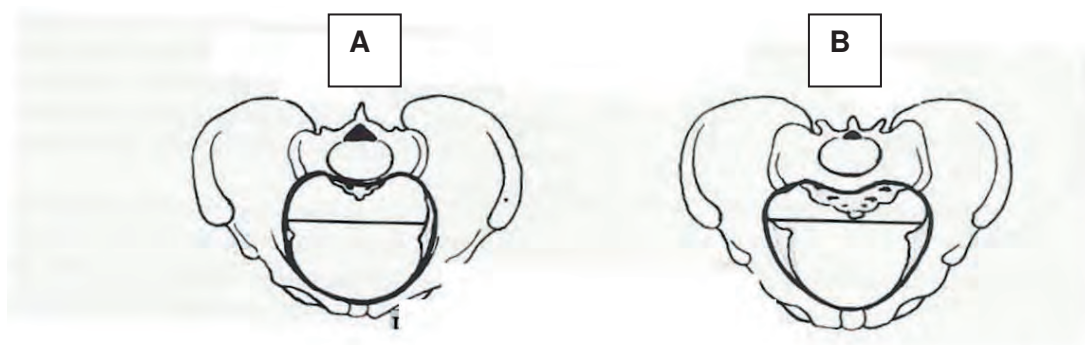
Question 2 Choose the correct answer/s

The presenting part is engaged when

- A. it enters the pelvic brim
- B. 2/5 of it is palpable above the brim
- C. the largest diameter has entered the brim (1)

Question 3

- a) Identify these two different types of pelves (1)



Ref : Fraser & Cooper, 2003,106

- b) Which is more likely to present with difficulties in labour, and why? (1)

Question 4

State 3 observations that can be made in relation to liquor. (1½)

Question 5

What signs indicate that a woman is in labour? (2)

Question 6

- a) What symptoms may a woman with Gestational proteinuric hypertension have? (1½)
and
b) What observations should be done? (1)

Question 7

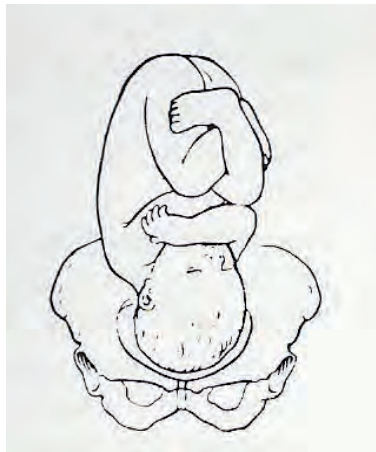
What is the significance of finding the following on successive vaginal examinations in labour? (1)

Occipito posterior, occipito lateral, occipito anterior

Question 8

What risk/s is/are associated with rupture of membranes longer than 12 hours? (1)

Question 9



In relation to this diagram,

- a) Indicate the following
- | | | |
|-------|-------------|-----|
| (i) | Denominator | (½) |
| (ii) | Position | (½) |
| (iii) | Attitude | (½) |
- b) Using the diagram on your answer sheet draw how this position would be felt on vaginal examination. (2)

[15]

Section B Partograph

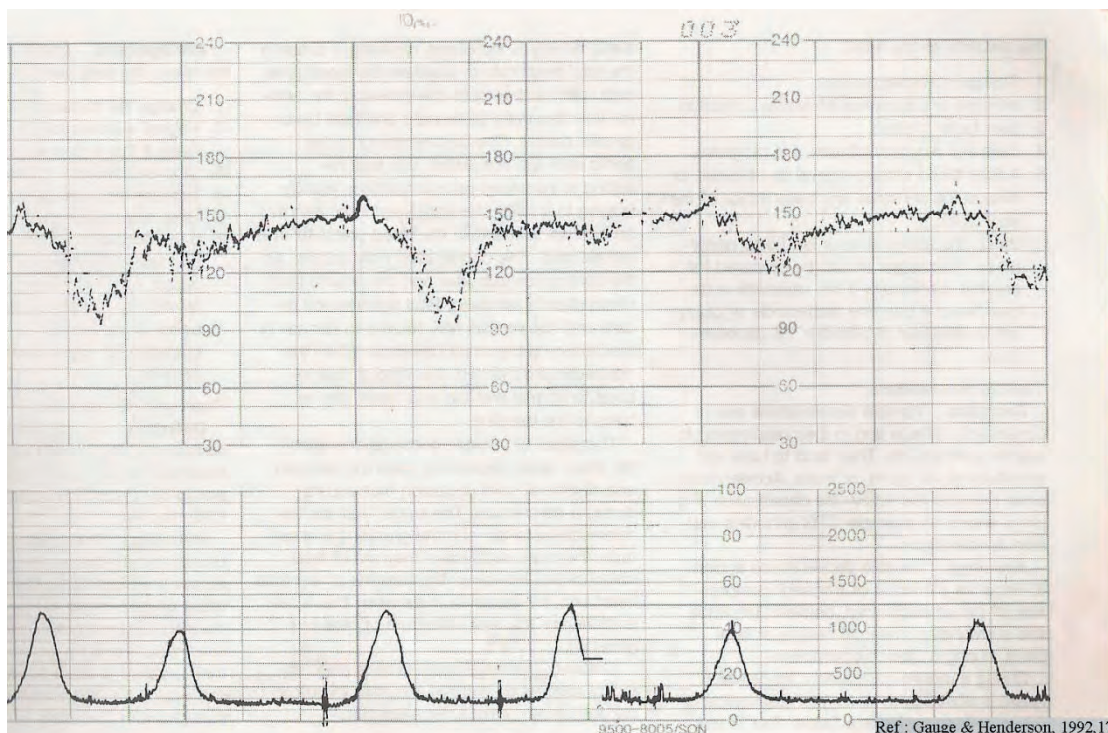
Mrs Adams, a 30 year old gravida 2 para 1 lady is admitted to the maternity unit at 33 weeks gestation with a raised blood pressure and 4+ protein in her urine. After two days, due to her deteriorating health status it was decided to induce her labour. Study the attached partograph and then answer the questions that follow.

1. Comment on the progress of labour and make an assessment of the maternal and fetal health status. (16)
2. Indicate what would be appropriate medical and midwifery management. (3)
3.
 - a) For what indication is Magnesium Sulphate prescribed, and what is its action? (1)
 - b) State the dosage that is given. (3)
 - c) What observations should be done while Magnesium Sulphate is being administered? (2)

[25]

Section C - CTG

Describe this CTG tracing and give your conclusion as to the status of this labour. [5]



Section D – Clinical skills

See answer sheet [Same questions as for test 1 and 3 - author note]

PARTOGRAM

Name: <u>MRS ADAMS</u>		Gravity: <u>G2</u>		Parity: <u>P1</u>		Spont. Labour		Induced Labour <u>X</u>	
Age: <u>30</u>		Risk Factors: <u>33/40</u>		Time of ROM: <u>08 00</u>					
Folder No.: <u>00802006</u>		<u>↑ BP 4+ proteinuria</u>		Duration of Labour (O/A)*:					
Institution:									

DATE		LATENT PHASE		DATE		ACTIVE PHASE	
16.10.06		16.10.06		16.10.06		16.10.06	
TIME		TIME		TIME		TIME	
DURATION IN HOURS		DURATION IN HOURS		DURATION IN HOURS		DURATION IN HOURS	
FETAL HEART RATE		FETAL HEART RATE		FETAL HEART RATE		FETAL HEART RATE	
Before Contraction		Before Contraction		Before Contraction		Before Contraction	
After Contraction		After Contraction		After Contraction		After Contraction	
DECCELERATIONS* EL/V/N		DECCELERATIONS* EL/V/N		DECCELERATIONS* EL/V/N		DECCELERATIONS* EL/V/N	
MOULDING*		MOULDING*		MOULDING*		MOULDING*	
LIQUOR*		LIQUOR*		LIQUOR*		LIQUOR*	
PRESENTING PART*		PRESENTING PART*		PRESENTING PART*		PRESENTING PART*	
CAPUT*		CAPUT*		CAPUT*		CAPUT*	
STATION		STATION		STATION		STATION	
APPLICATION*		APPLICATION*		APPLICATION*		APPLICATION*	
PROGRESS OF LABOUR		PROGRESS OF LABOUR		PROGRESS OF LABOUR		PROGRESS OF LABOUR	
CONTRACTIONS		CONTRACTIONS		CONTRACTIONS		CONTRACTIONS	
OXYTOCIN		OXYTOCIN		OXYTOCIN		OXYTOCIN	
BLOOD PRESSURE AND PULSE		BLOOD PRESSURE AND PULSE		BLOOD PRESSURE AND PULSE		BLOOD PRESSURE AND PULSE	
URINE		URINE		URINE		URINE	
TEMP - °C		TEMP - °C		TEMP - °C		TEMP - °C	
MANAGEMENT MEDICATION IV FLUID		MANAGEMENT MEDICATION IV FLUID		MANAGEMENT MEDICATION IV FLUID		MANAGEMENT MEDICATION IV FLUID	
SIGNATURE & DESIGNATION		SIGNATURE & DESIGNATION		SIGNATURE & DESIGNATION		SIGNATURE & DESIGNATION	

*O/A - ON ARRIVAL	*MOULDING: 0 1+ 2+ 3+	CAPUT: 0 1+ 2+ 3+	*LIQUOR: I - INTACT C - CLEAR M - MECONIUM STAINED B - BLOOD STAINED	*PRESENTING PART: BREECH - Br VERTEX - Vx FACE - F BROW - Bw	*APPLICATION O - GOOD P - POOR	*DECCELERATIONS: EL/V/N E - Early L - Late V - Variable N - None

Intrapartum project : Test 3

Notes : Mark allocation : ½ mark per point

Total marks : 45

Section A**Question 1 True / false**

- a) A maternal blood pressure of 150/90 is abnormal (½)
- b) Early fetal heart decelerations are more serious than late decelerations. (½)

Question 2

Choose the correct answer/s

The fetal heart rate may be higher when :

- A. The mother has a pyrexia
- B. The fetus is postterm
- C. The fetus is preterm
- D. The mother has hypertension (1)

Question 3

Give 3 features of fetal stress (1½)

Question 4

What signs indicate that a woman is in labour? (2)

Question 5

- a) What is the significance of a women having ketonurea during labour? (½)
- b) What action is required? (½)

Question 6

If a fetus enters the pelvic brim but is unable to proceed beyond the mid-cavity, what type of pelvis might the woman have? Give a reason for your answer. (1½)

Question 7

Which of the following symptoms are associated with abruptio placentae? (1)

- A. very tender abdomen
- B. bright red blood
- C. dark red blood with clots
- D. high presentation

Question 8

Which of the following symptoms are associated with placenta praevia? (1)

- A. painlessness
- B. clotting disorder (disseminated intravascular coagulopathy – DIC)
- C. gestational proteinuric hypertension
- D. abnormal presentation

Question 9

If a woman has symptoms of blurred vision, headache, epigastric pain

- c) what diagnosis should be suspected? (1½)
- d) what observations should be done? (1)

Question 10

Ref : Fraser & Cooper, 2003,558

In relation to this diagram, indicate the following :

- a) Lie
- b) (Presentation (not presenting part))
- c) Position
- d) Attitude (2)

Question 11

How many plusses of moulding are there when the foetal skull bones overlap and cannot be separated with digital pressure? (½)

Section B - Partograph

Mrs Cloete is a gravida 2 para 1 lady who is now at term. Her previous baby born at term per vaginam had a birth weight of 3900g.

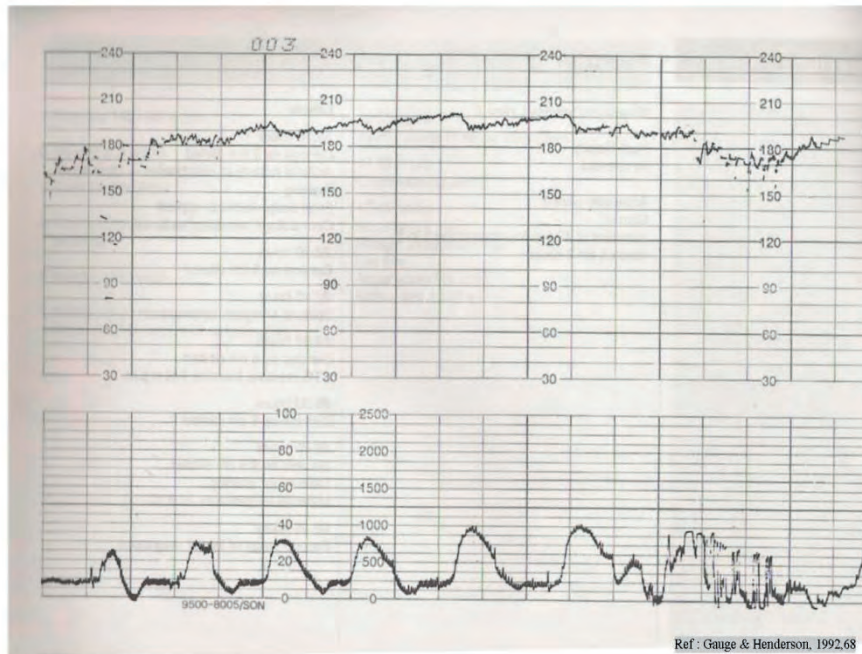
Study the attached partograph and answer the questions that follow.

- (a) Describe this labour and make a full assessment. (15½)
- (b) At 12:30 she had a spontaneous rupture of membranes. How would this affect the assessment and management of this labour? (1½)
- (c)
 - (i) Indicate how she would have been experiencing this labour at 15:00. (2)
 - (ii) In relation to the treatment she was receiving at 15:00, what specific observations should be made? (1)
 - (iii) What were her needs at this time and what interventions would have been appropriate? (4)
- (d) Explain the significance of the cervical dilation at 16:00 (1)

[25]

Section C – CTG

Describe this CTG tracing and give your conclusion as to the status of this labour. (5)



Section D - Clinical skills

1. Abdominal assessment

- State what you will assess and how you will do this. (8)
- Using the models provided, record your clinical findings. (2 ½)

2. Vaginal / internal examination

- Indicate what factors can be determined during a vaginal examination of a woman in labour. (3 ½)
- Using the model provided, record the cervical dilatation and effacement. (1)

[15]

392

Appendix 11a Midwife tests' data entry template - Test 1

Marking legend : 1 = correct 0 = incorrect blank = missing ☐ = no option

Study number

ID	Question	Description					
1	A1a	True/False	True				
2	A1b	Rate of Cx dilatation	False				
3	A2	MCQ	A	D			
4	A3	Fetal stress	Brady / tachycardia	Late decelerations		Meconium stained liquor	
5	A4	Labour signs	Regular contractions		ROM		
			Cervical changes		Show		
6	A5	Urinary obs	Volume	Protein	Glucose	Ketones	Blood
7	A6a	GPH diagnosis	GPH				
8	A6b	GPH obs	Blood Pressure	Urinary protein			
9	A7	Moulding	2+				
10	A8a	Diameters	A3				
11	A8b	Diameters	B4				
12	A8c	Diameters	C2				
13	A8d	Diameters	D1				
14	A9a i	Lie	Longitudinal				
15	A9a ii	Presentation	Cephalic / vertex				
16	A9a iii	Position	ROA				
17	A9a iv	Attitude	Flexed				
18	A9b	Conclusion	Normal OA delivery			Well flexed - should fit normal pelvis	
19	Bi	Maternal observations	Pulse rate has been climbing slowly			BP between 100/55 and 120/75	
			She has been nil per mouth			Low urine output (350 ml in 7 hours)	
			in spite of IV fluids			At 14:00 has ++ ketonurea	
			She is tired			Probably in some pain	
20	Bi	Fetal observations	Initially a Grade 1 foetal heart (120 - 160 bpm, no decels)		SROM at 11:00 - clear liquor		14:00 MSL
			From 12:30 the foetal heart pattern has been worsening (Grade 2 – 100 - 120 bpm or >160 bpm with no decels)		then Grade 3B (baseline 100 - 120 bpm or >160 bpm with early decels)		Indicates that the foetus is experiencing some stress
			15:00 caput +		15:00 no moulding		PP descending satis.
			Foetus is stressed		Req. careful monitoring		
21	Bi	Progress of labour	In early labour on admission		admitted 9 hours ago		Multip's os
			cervix was 1 cm long		presenting part was 5/5		2 IU Synto @ 15 dpm
			Synto inc 15 dpm ½ hrly till 11:30 (60 dpm)		Maintained at 60 dpm for past 3½ hours		Contractions developed quickly in frequency
			Contractions developed quickly in strength – overstimulation		Cx dilatation progressing reasonably		Cx dilatation not as quick as anticipated in a Grav 3

Marking legend : 1 = correct 0 = incorrect blank = missing

 = no option

ID	Question	Description			
22	Bii	Management	Evidence of fetal stress x 3 hours	Maternal fatigue	
			long and frequent contractions	Earlier intervention appropriate	
			associated with oxytocin usage	decrease / stop oxytocin	
23	Biii	Maternal management	Reduce / switch off oxytocin	give oxygen	
			Position on side	refer to obstetrician / medical officer	
			give calories	re-examine in 2 hours	
			comfort e.g. position, epidural, N ₂ O ₂	No Narcotic due to foetal condition	
			empty bladder	Give fluids	
24	Biii	Fetal management	careful monitoring	Possible foetal scalp sample for pH if condition worsens	
25	Biv	Delivery planning	may require assistance	inform nursery	
			will require infant resuscitation	call medical assistance timeously	
26	C	Baseline rate	Baseline rate 150-155 bpm	Normal	
27		Variability	Variability 5-15 bpm	Normal	
28		Pattern	Pattern no decels, some accels	Acceptable	
29		Contractions	Uterine contractions 4:10 >40s	Acceptable	
30		Conclusion	Normal pattern	Acceptable	
31	D1a	Gestational age	Gestational age, + 1 of the following	S-F measurement	Fundal grip
				History	Head hardness
			Fetal size		Landmarks
32		Lie	Lie	Combined grip	
33		Presentation	Presentation	Pawlick's grip	
34		Position	Position	Lateral grip	
35		Attitude	Attitude	Deep pelvic grip	
36		Level	Level	Pawlick's grip OR Deep pelvic grip	
37		Fet. movement	Fetal movement	Observe	Question mother
38		Fetal heart auscultation	Fetal heart auscultation, + 1 of the following	Pinard's stethoscope	Doptone
				CTG	Doppler
39	D1b	Lie	Longitudinal		
40		Presentation	Vertex / cephalic		
41		Position	LOA		
42		Level	4/5		
43		Attitude	Flexed		
44	D2a	Internal exam	Cervical Position	Station of presentation	Moulding
			Cervical Dilatation	Application of PP	Caput
			Cervical Effacement	Position of fetus	Condition of spines
			Presenting part	Attitude	Status of membranes
45	D2b i	Cx Dilatation	5cm		
46	D2b ii	Cx Effacement	5mm Rim / Full	Expressed as percentage	

Appendix 11b Midwife tests' data entry template - Test 2

Marking legend : 1 = correct 0 = incorrect blank = missing ☐ = no option

Study number

ID	Question	Description				
1	A1	True/False	False			
2	A2	MCQ	B	C		
3	A3a	Pelves	A = Gynaecoid	B = Android		
4	A3b	Labour implications	Android	Narrower outlet than inlet / DTA		
5	A4	Liquor obs	Clear	Meconium stained	Blood stained	
6	A5	Labour signs	Regular contractions AND one of the following			
			Cervical changes	ROM	Show	
7	A6a	GPH symptoms	Blurred vision	Headache	Epigastric pain	
8	A6b	GPH obs	Blood Pressure	Urinary protein		
9	A7	Labour PV	Long rotation of OP position			
10	A8	Prol ROM	Chorio-amnionitis	Infection – mother / baby		
11	A9a i	Denominator	Occiput			
12	A9a ii	Position	ROP			
13	A9a iii	Attitude	Flexion			
14	A9b	PV ROP	Sagittal suture in rt oblique diam.	Posterior fontanelle in RP	Attitude flexed	Anat. correct orientation
15	B1a	Maternal observations	BP a bit high (130/100)	BP reasonably controlled		Temp within acceptable range
			Pulse increasing	persistent proteinurea		ketonurea at 16:00
			Urine volume adequate	possibility of eclampsia developing		Indicative of maternal fatigue
16	B1b	Progress of labour	oxytocic induction	Contractions increased to 3-4:10 >40s since 12:30		oxytocics stopped at 14:30 due to evidence of fetal stress
			Cervical effacement has occurred	Cx dilatation progressing 1cm/hour		Some rotation from ROP to ROL at 14:00
17	B1c	Fetal observations	FHR range 110 – 124 bpm	13:00 - evidence of early decelerations		one occasion of a late deceleration at 14:30
			Liquor was clear from SROM at 07:30	13:00 meconium liquor observed		Presentation has descended and engaged
			No moulding	no caput		
18	B1d	Labour conclusion	The mother is coping well although	BP and proteinurea to be watched carefully		Appropriate medications required
			Exhibiting some fatigue	The fetus is showing signs of stress		There are early and late decelerations
			FHR in a preterm fetus would be expected to be higher	This rate might be considered as relatively bradycardic		accompanied by meconium stained liquor
			No descent of presentation in the past 2 hours	Should not be a problem as this should be a small head (preterm fetus)		

Marking legend : 1 = correct 0 = incorrect blank = missing

 = no option

ID	Question	Description				
19	B2a	Medical Management	Appropriate supportive drug therapy	CNS depressant	hypotensive agent	
			Consider an assisted delivery	so as not to increase her blood pressure	associated with prolonged pushing	
			Need for a paed. / neonatal nurse	to assist with resuscitation of infant	Because of meconium stained liquor	
20	B2b	Midwifery Management	Give calories	Some analgesia, e.g. Entonox	Maintain supportive care	
21	B3a	MgSO4 indic.	Presence of proteinuria		CNS depressant	
22	B3b	MgSO4 dosage	Loading dose – 4G in 200ml over 20-30 minutes		Maintenance dose – 4G in 200 ml at 50 ml/hr	
23	B3c	MgSO4 obs	Urine volume	Proteinuria	Respiration Rate	Reflexes
24	C	Baseline rate	145-150 bpm		Normal	
25		Variability	5-10 bpm		Normal	
26		Pattern	Late decels lasting approx 60 secs		Indicative of distress	
27		Ut. contractions	4:10 >40s, moderate strength with resting periods		Acceptable	
28		Conclusion	Fetal distress		Intervention required	
29	D1a	Gestational age	Gestational age, + 1 of the following :	Fundal grip	SF measurement	
				Head hardness	History	
				Fetal size	Landmarks	
30		Lie	Lie		Combined grip	
31		Presentation	Presentation		Pawlick's grip	
32		Position	Position		Lateral grip	
33		Attitude	Attitude		Deep pelvic grip	
34		Level	Level	Pawlick's grip OR Deep pelvic grip		
35		Fet. movement	Fetal movement	Observe	Question mother	
36		Fetal heart auscultation	Fetal heart auscultation, + 1 of the following	Pinard's stethoscope	Doptone	
				CTG	Doppler	
37	D1b	Lie	Longitudinal			
38		Presentation	Vertex / cephalic			
39		Position	ROA			
40		Level	2/5			
41		Attitude	Flexed			
42	D2a	Internal exam	Cervical Position	Station of presentation	Moulding	
			Cervical Dilatation	Application of PP	Caput	
			Cervical Effacement	Position of fetus	Condition of spines	
			Presenting part	Attitude	Status of membranes	
43	D2b i	Cx Dilatation	7 cm			
44	D2b ii	Cx Effacement	Rim / Full		Expressed as percentage	
45	D3a	FH rate			Not available	
46	D3b	FH pattern			Not available	
47	D3c	Nature of decel			Not available / applicable	

Appendix 11c Midwife tests' data entry template - Test 3

Marking legend : 1 = correct 0 = incorrect blank = missing ☐ = no option

Study number

ID	Question	Description			
1	A1a	Mat BP	True		
2	A1b	Decelerations	False		
3	A2	MCQ	A	C	
4	A3	Fetal stress	Brady / tachycardia	Late decelerations	Meconium stained liquor
5	A4	Labour signs	Regular contractions AND one of the following		
			Cervical changes	ROM	Show
6	A5a	Ketonurea	Maternal fatigue		
7	A5b	Ketonurea Rx	Glucose / calories / energy		
8	A6	Pelvis	Android	Narrower mid-cavity and outlet than inlet	
9	A7	MCQ	A	C	
10	A8	MCQ	A	D	
11	A9a	GPH diag	GPH		
12	A9b	GPH obs	BP	Urinary protein	
13	A10a	Lie	Longit		
14	A10b	Presentation	Cephalic		
15	A10c	Position	RML		
16	A10d	Attitude	Extended		
17	Ba i	Maternal observations	Latent phase : maternal condition good		
			Active phase : vital signs are normal		pulse is rising and will need to be watched
			urine output is poor		suggests dehydration
			At 15:30 there was + ketonurea		indicative of fatigue
			Will need to watch her nutrition		Will need to monitor fluid intake
18	Ba ii	Fetal observations	Latent phase : fetal condition good		FH within normal range
			Active phase : fetal heart rate good until 16:00		16:00 early decelerations noted (also at 17:00)
			16:00 +caput		no other signs of fetal distress despite strong contractions
19	Ba iii	Progress of labour	Latent phase : Slow dilatation of cervix		Contractions increasing slightly from 2:10 minutes < 20 s, to 2:10 minutes 20 - 40 s
			Active phase : cervix dilating <1 cm / hour		and Oxytocic augmentation
			potential overstimulation of uterus		rotating forward as evidenced by PV findings at 16:00
			Will need to watch if this progresses or arrests		Some descent has taken place
			Effacement of cervix		
			in the presence of good contractions		
			Presenting part - head - posterior position		
			SROM @ 12:30 – clear liquor		

Marking legend : 1 = correct 0 = incorrect blank = missing

 = no option

ID	Question	Description			
20	Ba iv	Labour conclusion	slightly slow progress especially for a G2 woman	problems could be developing and will need to be watched carefully	Propose next full assessment in 2 hours
21	Bb	>12:30 management	This occurred at a time when the head was still 4/5↑	there could be a risk of a prolapsed cord	Immediate internal assessment to determine fetal condition
22	Bc i	15:00 experience	Augmentation of labour at 14:00	Contractions have become a lot stronger	
			she is probably experiencing a lot of pain from them	Ability to accommodate the changes not as effective as for a normally developing labour	
23	Bc ii	15:00 treatment obs	Reaction of the foetus to the uterine activity	Relaxation of uterus between contractions	
24	Bc iii	15:00 needs and intervention	Need for physical comfort	relaxation, breathing	
			positional changes if possible	back rubbing, epidural anaesthesia	
			Need for emotional support	she is probably quite tired	
			Nourishment	IV therapy Dextrose 5%	
25	Bd ii	Cervical dilatation	Recording on right of alert line	Progress < 1 cm/hour – slower than average	
26	C	Baseline rate	190 - 200 bpm	Abnormal, ? preterm or infection	
27		Variability	<5 bpm	Not acceptable ? hypoxia	
28		Pattern	No decelerations	Acceptable	
29		Uterine contractions	4:10 varying strength/duration (most >40 s) resting periods	Acceptable	
30		Conclusion	Some abnormal readings	Req further investig for cause + mgt	
31	D1a	Gestational age	Gestational age, + 1 of the following :	Fundal grip	SF measurement
				Head hardness	History
			Fetal size		Landmarks
32		Lie	Lie	Combined grip	
33		Presentation	Presentation	Pawlick's grip	
34		Position	Position	Lateral grip	
35		Attitude	Attitude	Deep pelvic grip	
36		Level	Level	Pawlick's grip OR	Deep pelvic grip
37		Fet. movement	Fetal movement	Observe	Question mother
38		Fetal heart auscultation	Fetal heart auscultation, + 1 of the following	Pinard's stethoscope	Doptone
				CTG	Doppler
39	D1b	Lie	Longit		
40		Presentation	Cephalic		
41		Position	LOA		
42		Level	2/5		
43		Attitude	Flexed		
44	D2a	Internal exam	Cervical Position	Station of presentation	Moulding
			Cervical Dilatation	Application of PP	Caput
			Cervical Effacement	Position of fetus	Condition of spines
			Presenting part	Attitude	Status of membranes
45	D2b i	Cx Dilatation	7 cm		
46	D2b ii	Cx Effacement	Rim / Full		Expressed as percentage

Appendix 12 Alignment of reporting with CONSORT Statement requirements

The original Consolidated Standards of Reporting Trials (CONSORT) statement developed in 1996 was an initiative to promote accurate, complete, clear and transparent reporting of methodology and findings of randomised controlled trials which reflected actual trial design and conduct. In addition this aids in appraisal of published trial reports (Schulz, Altman *et al.* 2010) and facilitates comparisons for systematic reviews. (Moher, Hopewell *et al.* 2010) Over the years, in addition to the two major updates in the guidelines in 2001 and 2010, various extensions have been developed to address particular study designs or research issues, e.g. for non-inferiority and equivalence randomised trials, blinding in randomised trials, better reporting of harms. The two CONSORT extensions relevant to this study are those for cluster randomised trials and for pragmatic trials. The three sets of criteria have been consolidated into a single table (below) and the relevant paragraph reference in the study given. With the 2010 update of the standard CONSORT statement, other extensions of this statement are in the process of being reviewed and updated. This has been completed for the cluster randomised trials (Campbell, Piaggio *et al.* 2012), but no update has been published for pragmatic trials. In order to align the content in the table below, pragmatic extension items 3 ('participants') and 4 ('interventions') are reflected appropriately alongside the so-named items 4 and 5 respectively, on the standard and cluster extension descriptions.

Section	Item	Standard CONSORT description ¹ <i>with extension for a cluster randomised trial</i> [in italics] ²	(A) (B)	Extension for pragmatic trials ³ (C)	Study description / paragraph reference
Title and abstract					
	1a	Identification as a <i>cluster</i> randomised trial in the title			Title
	1b	Structured summary of trial design, methods, results and conclusions (for specific guidance see CONSORT for abstracts)			Abstract

Section	Item	Standard CONSORT description ¹ <i>with extension for a cluster randomised trial</i> <i>[in italics]</i> ²	(A) (B)	Extension for pragmatic trials ³ (C)	Study description / paragraph reference
Introduction					
Background and objectives	2a	Scientific background and explanation of rationale <i>Rationale for using a cluster design</i>		Describe the health or health service problem that the intervention is intended to address and other interventions that may commonly be aimed at this problem	A 3.2 B 3.2.1.2.1 C 1.3.2, 1.8, 1.9
	2b	Specific objectives or hypotheses <i>Whether they pertain to the cluster level, the individual participant level, or both</i>			A 2.6.2 2.6.3 B 2.6.2
Methods					
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio. <i>Definition of cluster and description of how the design features apply to the clusters</i>			A 3.7.1 B 3.7.1, 3.7.3
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons			Not applicable
Participants	4a	Eligibility criteria for participants <i>Eligibility criteria for clusters</i>		Eligibility criteria should be explicitly framed to show the degree to which they include typical participants and/or, where applicable, typical providers (e.g. nurses), institutions (e.g. hospitals), communities (or localities e.g. towns) and settings of care (e.g. different health care financing systems)	A 3.7.3.1, 3.7.3.2 B 3.7.2 C 1.4, 1.5, 1.6, 3.7.2

Section	Item	Standard CONSORT description ¹ <i>with extension for a cluster randomised trial</i> [in italics] ²	(A) (B)	Extension for pragmatic trials ³ (C)	Study description / paragraph reference
	4b	Settings and locations where the data were collected			A 3.7.2
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered <i>Whether interventions pertain to the cluster level, the individual participant level, or both</i>		Describe extra resources added to (or resources removed from) usual settings in order to implement intervention. Indicate if efforts were made to standardise the intervention or if the intervention and its delivery were allowed to vary between participants, practitioners, or study sites. Describe the comparator in similar detail to the intervention	A 3.7.4, 3.7.4 B 3.7.4 C 3.7.3.4, 3.7.4
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures including how and when they were assessed. <i>Whether outcome measures pertain to the cluster level, the individual participant level, or both</i>		Explain why the chosen outcomes and, when relevant, the length of follow up are considered important to those who will use the results of the trial	A 3.7.5 B 3.7.5 C 3.7.5
	6b	Any changes to trial outcomes after the trial commenced, with reasons			Not applicable

Section	Item	Standard CONSORT description ¹ <i>with extension for a cluster randomised trial</i> <i>[in italics]</i> ²	(A) (B)	Extension for pragmatic trials ³ (C)	Study description / paragraph reference
Sample size	7a	How sample size was determined <i>Method of calculation, number of cluster(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intra cluster correlation (ICC or k), and an indication of its uncertainty</i>		If calculated using the smallest difference considered important by the target decision maker audience (the minimally important difference) then report where this difference was obtained	A 3.7.3.1, 3.7.3.2 B 3.7.3.1, 3.7.3.2 C 3.7.3.1, 3.7.3.2
	7b	When applicable, explanation of any interim analyses and stopping guidelines			Not applicable. Presented in 3.7.3.1 and 3.7.3.2
Randomisation					
Sequence generation	8a	Method used to generate the random allocation sequence			A 3.7.3.1, 3.7.3.2
	8b	Type of randomisation; details of any restriction (such as blocking and block size) <i>Details of stratification or matching if used</i>			A 3.7.3.1, 3.7.3.2 B 3.7.3.3
Allocation concealment mechanism	9	Method used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned. <i>Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level, or both</i>			A 3.7.3.1, 3.7.3.2 B 3.7.3.3

Section	Item	Standard CONSORT description ¹ <i>with extension for a cluster randomised trial</i> <i>[in italics]</i> ²	(A) (B)	Extension for pragmatic trials ³ (C)	Study description / paragraph reference
Implementation	10	Who generated the allocation sequence, who enrolled the participants, and who assigned participants to their groups			A 3.7.3.1, 3.7.3.2
	10a	Who generated the random allocation sequence, who enrolled the clusters, and who assigned clusters to interventions			B 3.7.3.3
	10b	Mechanism by which individual participants were included in clusters for the purposes of the trial (such as complete enumeration, random sampling)			B 3.7.3.1, 3.7.3.2
	10c	From whom consent was sought (representatives of the cluster, or individual cluster members, or both) and whether consent was sought before or after randomisation			B 3.8.4, 3.8.5
Blinding	11a	If done, who was blinded after assignment to interventions (for example participants, care providers, those assessing outcomes), and how		If blinding was not done, or was not possible, explain why	A 3.7.3.3, 3.7.6.1.4, 3.7.6.2.2, 3.7.6.2.4, 3.7.9 C 3.7.3.3
	11b	If relevant, description of the similarity of interventions			Not relevant
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes <i>How clustering was taken into account</i>			A 3.7.8.1, 3.7.8.2 B 3.7.8.1, 3.7.8.2
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses			A 3.7.8.1

Section	Item	Standard CONSORT description ¹ <i>with extension for a cluster randomised trial</i> [in italics] ²	(A) (B)	Extension for pragmatic trials ³ (C)	Study description / paragraph reference
Results					
Participant flow (a diagram is strongly recommended)	13a	For each group, the number of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome. <i>For each group, the number of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome</i>		The number of participants or units approached to take part in the trial, the number which were eligible, and reasons for non-participation should be reported	A 4.3.1, 4.4.1 B 4.3.1 C 4.3.1, 4.4.1
	13b	For each group, losses and exclusions after randomisation, together with reasons <i>For each group, losses and exclusions for both clusters and individual cluster members</i>			A 4.3.1 B 4.3.1, 4.4.1
Recruitment	14a	Dates defining the periods of recruitment and follow up			A 3.7.4.3, Table 3.3, 4.2.1
	14b	Why the trial ended or was stopped			A Ended as specified in the protocol
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group <i>Baseline characteristics for the individual and cluster levels as applicable for each group</i>			A 4.2.2, 4.2.3 B 4.2.1
	16	For each group, number of participants (denominator) in each group included in each analysis and whether the analysis was by original assigned groups <i>For each group, number of clusters included in each analysis</i>			A 4.3.2, 4.4.2 B 4.3.1, 4.4.2

Section	Item	Standard CONSORT description ¹ <i>with extension for a cluster randomised trial</i> <i>[in italics]</i> ²	(A) (B)	Extension for pragmatic trials ³ (C)	Study description / paragraph reference
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group and the estimated effect size and its precision (such as 95% confidence interval) <i>Results at the individual cluster level as applicable, and a coefficient of intracluster correlation (ICC or κ) for each primary outcome</i>			A 4.3.2, Table 4.3.3, and 4.4.2.1, Tables 4.4.5 and 4.4.7
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended			B 4.3.2, Table 4.3.3
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory			A 4.3.2, 4.3.2.1
Harms	19	All important harms or unintended effects in each group			A 5.2.1.3, 5.3.1.3
Discussion					
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and if relevant, multiplicity of analyses			A 3.7.9, 5.2.1.4, 5.3.1.4, 5.4.2
Generalisability	21	Generalisability (external validity, applicability) of the trial findings <i>Generalisability to clusters and / or individual participants (as relevant)</i>		Describe key aspects of the setting which determined the trial results. Discuss possible differences in other settings where clinical traditions, health service organisation, staffing, or resources may vary from those of the trial	A 5.5 B 5.5 C 5.2.3.4, 5.2.3.5, 5.2.3.6, 5.4.2.5, 5.4.2.6, 5.5.1

Section	Item	Standard CONSORT description ¹ <i>with extension for a cluster randomised trial</i> [in italics] ²	(A) (B)	Extension for pragmatic trials ³ (C)	Study description / paragraph reference
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence			A 5.2.1, 5.2.2, 5.2.3, 5.3.1, 5.3.2, 5.3.3
Other information					
Registration	23	Registration number and name of trial registry			Not applicable. South African registry for pharmaceutical trials only, at commencement of this study
Protocol	24	Where the full trial protocol can be accessed, if available			A 3.8.3
Funding 3.8	25	Sources of funding and other support (such as supply of drugs), role of funders			A Acknowledgements

References

- Schulz KF, Altman DG, Moher D, for the CONSORT group (2010). "CONSORT 2010 statement : updated guidelines for reporting parallel-group randomised trials." BMC Medicine 8:18
- Campbell MK, Piaggio G, Elbourne DR, Altman DG for the CONSORT group (2012). "CONSORT 2010 statement : extension to cluster randomised trials." British Medical Journal 334:e5661
- Zwarenstein M, Treweek S, Gagnier J, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D, for the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups (2008) "Improving the reporting of pragmatic trials : an extension of the CONSORT statement." British Medical Journal 337:a2390

Appendix 13 Ethics approval, University of Cape Town



UNIVERSITY OF CAPE TOWN

Health Sciences Faculty
Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Telephone [021] 406 6338 • Facsimile [021] 406 6411
e-mail: preaward@curic.uct.ac.za

07 July 2006

REC REF: 311/2004

/Prof S Clow
Nursing & Midwifery
Health & Rehabilitation Sciences

Dear A/Prof Clow

PROJECT TITLE: IMPROVING THE QUALITY OF CLINICAL MANAGEMENT IN LABOUR BY MIDWIVES AT PRIMARY AND SECONDARY LEVELS OF THE PUBLIC SECTOR HEALTH SERVICE IN THE WESTERN CAPE, THROUGH THE USE OF THE PARTOGRAPH

Thank you for your letter to the Research Ethics Committee dated 25 June 2006.

Thank you for confirming your intention to proceed with a doctoral degree based on your findings from an earlier study. In particular, the committee appreciates your thorough and comprehensive analysis of anticipated ethical issues. We look forward to receiving your finalised proposal.

Please quote the REC. REF in all your correspondence.

Yours sincerely

signature removed

DR. M. BLOCKMAN
CHAIRPERSON, HSF HUMAN ETHICS

lemjedi



UNIVERSITY OF CAPE TOWN

Health Sciences Faculty
Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Telephone [021] 406 6338 • Facsimile [021] 406 6411
e-mail: lemjcdi@curie.uct.ac.za

31 May 2007

REC REF: 244/2007

A/Professor Sheila Clow
Health and Rehabilitation Sciences
Old Main Building
Groote Schuur Hospital

Dear Professor Clow

**IMPROVING THE QUALITY OF CLINICAL MANAGEMENT IN LABOUR BY REGISTERED
MIDWIVES AT THE PRIMARY LEVEL OF THE PUBLIC SECTOR HEALTH SERVICE
THROUGH TRAINING AND MENTORING PROGRAMME**

Thank you for submitting your study to the Research Ethics Committee for review.

It is a pleasure to inform you that the Ethics Committee has **formally approved** the above-mentioned study.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the REC. REF in all your correspondence.

Yours sincerely

signature removed

PA A/PROF. M. BLOCKMAN
CHAIRPERSON, HSF HUMAN ETHICS

lhenley

Appendix 14 Permission from Provincial Government

Navrae
Enquiries
Imibuzo

Ms L Olivier
Ms E Arends

Telefoon
Telephone
Ifowuni

(021) 483-2686
(021) 483-2680

Verwysing
Reference
Isalathiso

Datum
Date
Umhla

10 December 2004



To: **Professor S Clow**
School of Health & Rehabilitation Services
Division of Nursing & Midwifery
University of Cape Town
Anzio Road
OBSERVATORY 7925

Dear Prof Clow

PERMISSION TO CONDUCT RESEARCH STUDY IN MATERNAL AND WOMEN'S HEALTH

Thank you for submitting your research proposal on **"Improving quality of clinical management in labour by midwives at primary and secondary levels of the public sector health services of the Western Cape through use of the partogram"**.

The provincial Maternal Child and Women's Health (MCWH) sub-directorate, is in full support of this research and would have no objection to your conducting the research within the Western Cape health facilities as indicated in the proposal. The MCWH sub-directorate agrees to providing the costs for the training workshops in the four health regions and assistance with statistical analysis in the form of the services of Mr Rauf Sayed at UCT.

We would like to thank you for your valuable contributions to the maternal services as member of the provincial Maternal Guidelines Reference group and are looking forward to the outcome of your research as this could make a great impact on future maternal services in the province.

4 Dorp Street
P.O. Box 2060
Cape Town
8000
Fax No: (021) 483 4345

Dorpstraat 4
Posbus 2060
Kaapstad
8000
Faks: (021) 483 4345

We wish you every success with your research study.

Kind regards

signature removed

MS L OLIVIER
DEPUTY DIRECTOR
MATERNAL CHILD & WOMEN'S HEALTH SUB-DIRECTORATE

signature removed

MS E ARENDS
ASSISTANT DIRECTOR
MATERNAL CHILD & WOMEN'S HEALTH SUB-DIRECTORATE

4 Dorp Street
P.O. Box 2060
Cape Town
8000
Fax No: (021) 483 4345

Dorpsstraat 4
Posbus 2060
Kaapstad
8000
Faks: (021) 483 4345



Verwysing
Reference
Isalathiso

Navrae
Enquiries
Imibuzo

Dr T Naledi

Telefoon
Telephone
Ifowuni

021-483 9901

Departement van Gesondheid
Department of Health
iSebe lezeMpilo

Prof S. Clow
Division of Nursing and Midwifery
School of Health and Rehabilitation Sciences
Anzio Road
Observatory
7925

Dear Prof Clow

Improving the quality of clinical management in labour by midwives at primary, secondary levels of the Public Sector Health services in the Western Cape through the use of the partogram

Thank you for submitting your proposal to undertake the above-mentioned study.

We are pleased to inform you that your research proposal has been approved. Please contact Dr Frans Krige the Director for Boland/Overberg Region on email : ikrige@pgwc.gov.za or 023 348 1401 and Ms Carine Bester the Director of the Westcoast/Winelands on email: cwbester@pgwc.gov.za or tel 023 487 9906 to make the necessary arrangements for your study.

We would however like to be informed of the results of your study and would like you to please inform us in writing when the report would be available.

We look forward to hearing from you.

Yours sincerely

signature removed

DR J CUPIDO
ACTING DEPUTY-DIRECTOR GENERAL
DISTRICT HEALTH SERVICES AND PROGRAMMES
DATE: 21/9/2006

CC: Dr Frans Krige Director: Boland/Overberg
Ms Carine Bester Director: Westcoast/Winelands Region

Dorpstraat 4
Posbus 2060
KAAPSTAD
8000

4 Dorp Street
PO Box 2060
CAPE TOWN
8000



Verwysing
Reference
Isalathiso

19/18/RP244/2007

Navrae
Enquiries
Imibuzo

Dr T. Naledi

Telefoon
Telephone
ifowuni

021 483 9901

Departement van Gesondheid
Department of Health
Isaka IsakaMali

Associate Professor S. Clow
Division of Nursing and Midwifery
University of Cape Town
Faculty of Health Sciences
Anzio Road
Observatory
7925

Fax to 021 4086323

Dear Professor Clow

Improving the quality of clinical management in labour by registered midwives at the primary level of the public sector health service through training and mentoring programme

Thank you for requesting permission to access client records. Kindly contact the following members of staff to assist you with access to the facility:

1) Dr E. Erasmus at elerasmus@gwc.gov.za tel: 021 8511170 (Helderberg CHC)

We look forward to hearing from you.

Yours sincerely

signature removed

DR JCUPISO
DEPUTY DIRECTOR GENERAL
DISTRICT HEALTH SERVICES AND PROGRAMMES
DATE: 26/5/2008

CC: DR K. GLOETE
DR V. ZWIGENTHAL

CD: MDHS

Dorpstraat 4
Posbus 2060
KAAPSTAD
8000

4 Dorp Street
PO Box 2060
CAPE TOWN
8000



University of Cape Town

Division of Nursing and Midwifery

School of Health and Rehabilitation Sciences

Anzio Road, Observatory, 7925

Tel : (021) 406 6449

Fax : (021) 406 6323

eMail : sclow@uctgsh1.uct.ac.za

Research project : Improving the standard of intrapartum care by registered midwives in primary level public sector health facilities
Information sheet for registered midwives (intervention sites)

Dear Colleague

My name is Sheila Clow and I work in the Division of Nursing and Midwifery at the University of Cape Town. For the past few years I have been a member of the PGWC Maternity guidelines group which tries to address factors which contribute to maternal and fetal deaths in the Western Cape.

While we in the Western Cape are proud of the fact that we have the lowest levels in the country, we also know from the Saving Mothers and Saving Babies reports that there are many avoidable factors which have been identified. In fact it has been shown that in the rural areas, intrapartum asphyxia and birth trauma are responsible for over 50% of all deaths of babies with birthweights over 2500g. A key recommendation of both these reports is to “ensure that each site conducting births has the necessary equipment and protocols and that health care providers are appropriately trained to manage labour and are especially trained in the use of the partogram.”

With the support of the Maternal Child and Women’s Health sub-directorate, I aim to support this key national recommendation. I plan to undertake a study on improving the standard of intrapartum care by registered midwives. Two interventions are planned. One is to have a focussed update on training of registered midwives in relation to intrapartum care, which I will conduct. The second intervention is planned as a mentoring type of programme to equip senior clinicians/nurse managers to support, guide and strengthen the clinical practice of the colleagues in the intrapartum setting, and to develop skills for continuous quality assurance through the use of clinical audit.

What does this mean for you?

You would undertake 2 days training one month apart which will

- ◆ give you an update on maternal, fetal and labour assessment
- ◆ update clinical skills
- ◆ introduce you to the cardiotocograph and its use

You will be tested formally on your working knowledge and skills on 3 occasions :

- ◆ at the beginning of the project
- ◆ two months after the commencement of the project
- ◆ at the end of 12 months

The effect of the programme will be evaluated periodically in terms of auditing of partographs and analysis of maternal and perinatal data, in the year following the training. Your evaluations will be coded and no feedback will be given about individual performance. No-one else will see this, and your results will not be shared with ANYONE else. This means that this information will be confidential. You will also be asked to evaluate the training programme. The feedback will guide any revisions that are necessary. These opinions will also be confidential. Training and evaluation time will be regarded as on-duty time. You have the right not to participate, but I do hope that you will make the choice to be involved.

Eight other sites (randomly distributed in your region and one other rural region) will also have this project running. However eight sites in the two regions will not have this programme. You are requested therefore not to share material from this project with colleagues in those sites, until the data collection period is completed (approximately 1 year). At the end of the research project a dissertation will be written. I also expect to present conference papers and publish in a journal. At no time will you or your results be identified.

If you have any questions or aspects that need clarification, please do not hesitate to contact me at (021) 406 6449 (voicemail facility) or 083 659 5266. If you have any queries about the research you may contact my supervisor, Prof George Swingler at (021) 658 5324 or the chair of the UCT Faculty of Health Sciences Human Research Ethics committee, Dr Marc Blockman (021) 406 6492.

I trust that this will be a valuable experience for you, and will enable you to more effectively meet the clinical expectations of you.

Thank you

Yours sincerely

Prof Sheila Clow

Consent form

I, ... have read the information sheet relating to the study, "Improving the standard of intrapartum care by registered midwives in primary level public sector health facilities". I have had the study explained to me and my role in it. I have had the opportunity to ask any questions I had, and these have been answered to my satisfaction. I freely agree to participate.

Print name : Sign :

Witness name : Sign :

Date :



University of Cape Town

Division of Nursing and Midwifery

School of Health and Rehabilitation Sciences

Anzio Road, Observatory, 7925

Tel : (021) 406 6449

Fax : (021) 406 6323

eMail : sclow@uctgsh1.uct.ac.za

Research project : Improving the standard of intrapartum care by registered midwives in primary level public sector health facilities

Information sheet for registered midwives (control sites)

Dear Colleague

My name is Sheila Clow and I work in the Division of Nursing and Midwifery at the University of Cape Town. For the past few years I have been a member of the PGWC Maternity guidelines group which tries to address factors which contribute to maternal and fetal deaths in the Western Cape.

While we in the Western Cape are proud of the fact that we have the lowest levels in the country, we also know from the Saving Mothers and Saving Babies reports that there are many avoidable factors identified. In fact it has been shown that in the rural areas, intrapartum asphyxia and birth trauma are responsible for over 50% of all deaths of babies with a birthweight over 2500g. A key recommendation of both these reports is to “ensure that each site conducting births has the necessary equipment and protocols and that health care providers are appropriately trained to manage labour and are especially trained in the use of the partogram.”

With the support of the Maternal Child and Women’s Health sub-directorate, I aim to support this key national recommendation. I have undertaken a study on improving the standard of intrapartum care by registered midwives. Certain sites have been randomly selected as intervention sites, and your workplace has been randomly selected as a control site.

It is important to be able to evaluate both control and intervention sites to determine the effectiveness of the interventions.

What does this mean for you?

- ◆ You will be tested formally on your working knowledge and skills. You are not required to do any swotting as this relates to your current operating knowledge

and skills. This will take approximately 1 hour on two occasions (second half of 2006 and approximately 1 year later.)

- ◆ Your evaluations will be coded and no feedback will be given about individual performance. This will not be able to be used for or against you in any work related capacity. No-one else will see this, and your results will not be shared with ANYONE else. This means that this information will be confidential.

Does this carry any benefit?

- ◆ It will contribute to the understanding of how to influence and support clinical management in labour.
- ◆ Training in clinical management of labour and fetal monitoring will be made available after the data collection period is complete.
- ◆ A tool to support clinical decision making in labour will be developed during this project, and you will get the opportunity to be introduced to it and have this available in your labour ward.

Evaluation time will be regarded as on-duty time. You have the right to choose not to participate, and there will be no negative consequences for you. However I do hope that you will choose to participate.

At the end of the research project a dissertation will be written. I also expect to present conference papers and publish in a journal. At no time will you or your results be identified.

If you have any questions or aspects that need clarification, please do not hesitate to contact me at (021) 406 6449 (voicemail facility) or 083 659 5266. If you have any queries about the research you may contact my supervisor, Prof George Swingler at (021) 658 5324 or the chair of the UCT Faculty of Health Sciences Human Research Ethics committee, Dr Marc Blockman (021) 406 6492.

Thank you
Yours sincerely

Prof Sheila Clow

Consent form

I, ... have read the information sheet relating to the study, "Improving the standard of intrapartum care by registered midwives in primary level public sector health facilities". I have had the study explained to me and my role in it. I have had the opportunity to ask any questions I had, and these have been answered to my satisfaction. I freely agree to participate.

Print name : Sign :

Witness name : Sign :

Date.....



University of Cape Town

Division of Nursing and Midwifery

School of Health and Rehabilitation Sciences

Anzio Road, Observatory, 7925

Tel : (021) 406 6449

Fax : (021) 406 6323

eMail : Sheila.Clow@uct.ac.za

Research project : Improving the standard of intrapartum care by registered midwives in primary level public sector health facilities

Information sheet for registered midwives (all sites after month 3)

Dear Colleague

My name is Sheila Clow and I work in the Division of Nursing and Midwifery at the University of Cape Town. For the past few years I have been a member of the PGWC Maternity guidelines group which tries to address factors which contribute to maternal and fetal deaths in the Western Cape.

While we in the Western Cape are proud of the fact that we have the lowest levels in the country, we also know from the Saving Mothers and Saving Babies reports that there are many avoidable factors identified. In fact it has been shown that in the rural areas, intrapartum asphyxia and birth trauma are responsible for over 50% of all deaths of babies with a birthweight over 2500g. A key recommendation of both these reports is to “ensure that each site conducting births has the necessary equipment and protocols and that health care providers are appropriately trained to manage labour and are especially trained in the use of the partogram.”

With the support of the Maternal Child and Women’s Health sub-directorate, I aim to support this key national recommendation. I have undertaken a study on improving the standard of intrapartum care by registered midwives. Certain sites have been randomly selected as intervention sites and some as control sites. It is important to be able to evaluate both control and intervention sites to determine the effectiveness of the interventions.

What does this mean for you?

- ◆ You will be tested formally on your working knowledge and skills. You are not required to do any swotting as this relates to your current operating knowledge and skills. This will take approximately 1 hour.
- ◆ Your evaluation will be coded and no feedback will be given about individual performance. This will not be able to be used for or against you in any work related capacity. No-one else will see this, and your results will not be shared with ANYONE else. This means that this information will be confidential.

Does this carry any benefit?

- ◆ It will contribute to the understanding of how to influence and support clinical management in labour.
- ◆ Training in clinical management of labour and fetal monitoring will be made available after the data collection period is complete.

Evaluation time will be regarded as on-duty time. You have the right to choose not to participate, and there will be no negative consequences for you. However I do hope that you will choose to participate.

At the end of the research project a dissertation will be written. I also expect to present conference papers and publish in a journal. At no time will you or your results be identified.

If you have any questions or aspects that need clarification, please do not hesitate to contact me at (021) 671 1838 (voicemail facility) or 083 659 5266. If you have any queries about the research you may contact my supervisor, Prof George Swingler at (021) 658 5324 or the chair of the UCT Faculty of Health Sciences Human Research Ethics committee, Dr Marc Blockman (021) 406 6492.

Thank you

Yours sincerely

Prof Sheila Clow

Consent form

I, ... have read the information sheet relating to the study, "Improving the standard of intrapartum care by registered midwives in primary level public sector health facilities". I have had the study explained to me and my role in it. I have had the opportunity to ask any questions I had, and these have been answered to my satisfaction. I freely agree to participate.

Print name : Sign :

Witness name : Sign :

Date :

Appendix 16 Feedback from the clinical facilitators at the end of their training programme

The feedback from the 3 groups of clinical facilitators was combined (translated from Afrikaans where necessary and précised)

1. *How important is this learning for me (motivation)?*

Response averaged across whole group : **10/10**

(Likert scale, where 10 = exceptionally motivated)

- ◆ Very important to give better intrapartum care and to identify problems early
- ◆ I can identify where weaknesses are in our work and that we can do something about these
- ◆ Quality care is so important. Feel much more confident to also look at other projects
- ◆ Very important. Correct completion of the partograph and the correct interpretation and appropriate action can be lifesaving/or the other way round.
- ◆ Partogram – without this instrument you can't get the whole picture of the progress of labour; the correct use can assist in interpretation so that there are better outcomes for the baby and mother
- ◆ It was very good to refresh my clinical knowledge and also learn how to become a mentor
- ◆ Very important to improve standard and quality of care; to increase clinical skills

2. *How confident am I to implement this?*

Response averaged across whole group : **8/10**

(Likert scale where 10 = exceptionally confident)

- ◆ I will try hard and with time I will grow in self-confidence. This is a new approach but there are guidelines to help.
- ◆ I will motivate myself to ensure that each time I give the best that I can to complete the observations and recording, and to be there for my colleagues

- ◆ I would really like to see the difference and improvement according to labour before and after this session
- ◆ This has been a good experience for me to develop more self confidence
- ◆ I am very motivated and look forward to this opportunity. I am a bit anxious about what my colleagues' attitudes and receptivity will be like.
- ◆ I realize that I cannot do this on my own and need the co-operation of my colleagues and I need to know what I am talking about.
- ◆ Very self confident BUT anxious about transmitting this enthusiasm to my colleagues; Am prepared to take this on even though I expect some resistance
- ◆ I am excited to get going
- ◆ Need a little bit more practise in clinical skills – for more confidence

3. One thing I have learned from this and one thing I am going to do differently

- ◆ The importance and value of the abdominal assessment
- ◆ Empower my colleagues by stressing the value of the partogram and frequent fetal heart monitoring.
- ◆ I realize the importance of the correct and adequate utilisation of the partograph and recording findings correctly in order so that the correct diagnosis can be made and action taken in time.
- ◆ I will definitely pay more attention to my work and to address problems, including better record keeping
- ◆ CTG tracing, recording of actions that might impact the CTG trace, and interpretation – what an “eye opener”!
- ◆ Develop as a mentor – be non-judgemental / less critical and be more supportive for those who need this, and to instil confidence. Also try and get them to share their problems and their fears.
- ◆ The importance of the clinical audit. This we plan to implement asap! I really enjoyed this time – it was of great value – thanx.
- ◆ Everyone must be on board, ready for change, has got different styles. Try to motivate and implement immediately

