Towards Rheumatic Heart Disease vaccine development: Defining host immune responses to Group A Streptococcal infection in Cape Town

Mogamat Taariq Salie

A thesis presented in fulfilment of the requirement for

the degree of Master of Science (Med) in the

Department of Medicine, University of Cape Town

January 2023

Supervisors
Prof Mark E Engel
Prof Jim B Dale

DECLARATION

I, Mogamat Taariq Salie, 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 a part of it has been, is being, or it to be submitted for another degree in this or any other university.

I empower the University of Cape Town to reproduce for the purpose of research either the whole or any portion of the contents in any whatsoever.

Signed by candidate

Signature

03 September 2023

Date

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

ABBREVIATIONS

LIST OF FIGURES AND TABLES

ABSTRACT

PUBLICATIONS

Chapter One: Introduction

Background

Rationale for this research

Chapter Two: Study 1. Systematic review of African emm coverage [Paper]

Systematic review and meta-analysis of the prevalence of Group A Streptococcal *emm* clusters in Africa to inform vaccine development.

Chapter Three: Study 2. Longitudinal epidemiology of GAS in Cape Town

The longitudinal assessment of Group A Streptococcal *emm* types in Cape Town, South Africa.

Chapter Four: Study 3. ARF systematic review [Paper]

Systematic Review: Utility of human immune responses to GAS antigens as a diagnostic indicator for ARF

Chapter Five: Study 4. Serum immune responses to GAS antigens

Serum immune responses to Group A Streptococcal antigens following new pharyngeal acquisitions.

Chapter Six: Highlights and Conclusions

ACKNOWLEDGMENTS

This thesis was supported in part by funds from the American Heart Association, National Research Fund and University of Cape Town.

My sincerest appreciation and thanks to the following persons:

Professor Mark Engel (Supervisor): It goes without saying how influential you have been in my career and personal life. Thank you for your continuous support, encouragement and guidance throughout this journey. You are the definition of a mentor, leader and friend.

Professor James Dale (Co-supervisor): An honour to be associated with your field of research. Thank you for your expert guidance and recommendations throughout the study.

Dr. Tom Penfound: The laboratory skills and Strep A knowledge I have gained from your 2-week stay in Cape Town was invaluable to this thesis. Thank you for your assistance with laboratory-related questions and continuous feedback regarding the data.

Professor Liezl Zühlke: Thank you for your invaluable input, guidance, and support throughout this thesis.

The AFRO*Strep* research team - Kélin Engel, Kimona Rampersadh, Rezeen Daniels, Lwazi Mhlanti: Thank you for your hard work, commitment, and dedication on the project, for the words of encouragement and continuous support throughout.

Mr Richard Meldau (Lab manager): Thank you Richard for your willingness to assist whenever I required it. For your invaluable training and usage of laboratory equipment and overall, guidance to laboratory protocols.

Contributing authors - Kélin Engel, Kimona Rampersadh, Rezeen Daniels, Annesinah Moloi, Babu Muhamed: Thank you for your hard work and contribution, it goes without saying that this thesis would not have been completed without you. Thank you for your patience and understanding when emails were sent days later and overall, your willingness to assist with feedback and input throughout.

Dr. Babu Muhamed and Dr. Dylan Barth – Thank you for your mentorship and guidance throughout this thesis. For your words of encouragement, teaching laboratory techniques, and being a role model.

The staff of the Department of Medicine for creating a pleasant environment to work in.

I acknowledge the relentless support, prayers, and encouragement of my family: parents, Fauzia and Faghree; my brothers, Natheer (& wife, Ghaleemah) and Ra'eez; and mother-inlaw, Tania Bagley.

To my extended family – Thank you for your kind words, prayers and encouragement.

Finally, to my wife, Thanaa Salie – Words cannot describe how grateful I am for your patience and support throughout this thesis. For motivating me when procedures failed, for encouraging me despite having imposter syndrome, and for always believing in me.

ABBREVIATIONS

GAS – Group A Streptococcus

RHD – Rheumatic heart disease

ARF – Acute rheumatic fever

SLO – Streptolysin O

DNase B – Deoxyribonuclease B

ULN – Upper limit of normal

SSE – Serine esterase

SpyAD – GAS adhesion and division protein

Mrp – M related peptides

STSS – Streptococcal toxic shock syndrome

PGSN – Post-streptococcal autoimmune sequelae such as post-streptococcal glomerulonephritis

APCs – Antigen-presenting cells

MLST – Multi-locus sequence typing

CDC – Centers for Disease Control and Prevention

HCW – Health care worker

ST – Sequence type

AHA – American Heart Association

IDSA – The Infectious Diseases Society of America

BPG – Benzathine penicillin G

ELISA – Enzyme-linked immunosorbent assay

SCPA – C5a peptidase

SLS – Streptolysin S

NETs – Neutrophil extracellular traps

MeSH – Medical Subject Headings

PRISMA – Preferred Reporting Items for Systematic reviews and Meta-Analysis

95% CI – 95% confidence interval

NINV – Non-invasive

INV – Invasive

Prot – Protected

N/P – Not protected

N/D – Not differentiated

NA – None

NCS – Not clearly stated

CRF – Case report form

NHLS – National Health Laboratory Services

LTFU – Lost to follow-up

βHS – ***β***-haemolytic streptococci

ASO – Anti-streptolysin O

NOS – Newcastle-Ottawa Scale

ASK – Anti-streptokinase

OR – Odd ratio

GAC – GAS carbohydrate

F/T – Full transcript

IU – International units

LIST OF FIGURES AND TABLES

Chapter 1

Figure 1 - Illustrating the role of GAS virulence factors played in pathogenesis. Adopted from Tart, Walker [2].

Figure 2 - Association of proteins with the cell surface of GAS. Adopted from Fischetti [3].

Table 1 - Summary of studies reporting the distribution of *emm* types from regions across the world.

Table 2 - *Emm*-clusters and their corresponding *emm*-types. Adopted from Sanderson-Smith, De Oliveira [4].

Figure 3 - Prevention strategies for combating GAS infections. Adopted from Carapetis [1].

Table 3 - GAS antigens proposed for this study.

Chapter 2

Table 1 – The *emm* clusters and their corresponding *emm* types. Adopted from Sanderson-Smith, De Oliveira [4].

Figure 1 – Schematic PRISMA flow diagram of the literature search. Figure is modelled after Moher, Liberati [5].

Table 2 – Characteristics of included studies.

Figure 2 – The five countries included in the review, representing the most abundant *emm* clusters.

Table 3 – The *emm* cluster distribution, representing the five countries included into the review and their respective isolate count.

Figure 3 – Forest plots showing the combined prevalence estimates of the four most abundant emm clusters of all included studies. (A) E6 emm cluster; (B) E3 emm cluster; (C) E4 emm cluster; (D) D4 emm cluster. Studies are represented by author, year of publication, effect size (ES) (proportion [95% CI]). n, number of isolates in cluster; N, total number of isolates in study.

Table 4 – Summary of the meta-analyses completed.

Figure 4 – Count of isolates within emm clusters from African studies included in this systematic review. Green bars indicate emm clusters represented in the 30-valent vaccine. Blue bars represent emm clusters not included in the vaccine. Gray bars represent isolates unassigned to a cluster or were "untypeable" according to the authors' report. Numbers represent the count of isolates across all studies.

Table 5 – Summary of risk of bias assessment.

Chapter 3

Table 1 – Primer sets used in *emm* typing procedure.

Table 2 – Representation of samples collected across all participants.

Table 3 – The number of participants according to classification groups

Table 4 – Interesting cases that could have been grouped in more than one category and those participants in the recurrent infection category with corresponding symptoms and antibiotics

Table 5 – Characterization of *emm* types of GAS isolates

Figure 1 – Graphical representation of emm type distribution with corresponding emm cluster and clade. Green bars represent the emm types included in the 30-valent vaccine (n=42, 63.2% coverage), with cross-protection afforded to emm8 and emm94 (n=48, 71.2% coverage). Un emm type, unknown emm types that includes: untypable (no emm sequence) = 9, failed cultures = 4, missing sample = 2; No emm cluster, has not been classified according to Sanderson-Smith (2014).

Chapter 4

Figure 1 – Schematic PRISMA flow diagram of the literature search. Figure is modelled after Moher, Liberati [5].

Table 1 – Characteristics of included studies.

Figure 2 – Forest plot evaluating the odds of association between Anti-SLO titres and ARF. Subgrouping based on whether a guideline was used to diagnose ARF.

Figure 3 – Forest plot evaluating the odds of association between Anti-DNase B titres and ARF. Subgrouping based on whether a guideline was used to diagnose ARF cases.

Figure 4 – Forest plot evaluating the odds of association between Anti-DNase B titres and ARF. Subgrouping based on antigen.

Chapter 5

Table 1. Group A streptococcus Antigens Evaluated

Table 2 – Grouping clinical events according to throat culture results and immune responses.

Figure 1 – Examples of antibody responses for clinical event groups 1 (A), and 3 (B).

Table 2 – Throat culture results and immune responses to shared GAS antigens and M peptides within three defined clinical event groups.

Table 3 – New antibody responses and pre-existing high antibody levels against GAS antigens according to clinical event group.

Figure 2 – Comparison of clinical events exhibiting pre-existing high levels of antibodies (open bars) against shared GAS antigens (SSE, SpyAD, SCPA, SLO and SpyCEP) versus those without pre-existing high levels of antibodies (grey bars). Grp 1 vs Grp 2: P=8.6E-07, Grps 1 and 3 vs Grp 2: P=4.1E-07, Grp 3 vs Grp 2: P=5.3E-07, Chi Square.

PUBLICATIONS

Peer reviewed

Systematic review GAS Africa (Chapter 2 – Study I)

Salie, T., Engel, K., Moloi, A., Muhamed, B., Dale, J.B. & Engel, M.E., 2020. Systematic Review and Meta-analysis of the Prevalence of Group A Streptococcal emm Clusters in Africa To Inform Vaccine Development. *Msphere*, 5(4), pp.e00429-20.

Systematic review ARF (Chapter 4 – Study III)

Salie, M. T., Rampersadh, K., Muhamed, B., Engel, K. C., Zühlke, L. J., Dale, J. B & Engel, M. E. 2021. Utility of Human Immune Responses to GAS Antigens as a Diagnostic Indicator for ARF: A Systematic Review. *Frontiers in Cardiovascular Medicine*, 8.

Chapter 5 – Study IV

Conference presentations

Salie T, Engel M. Systematic review of the worldwide prevalence of group A Streptococcus emm-clusters: Vaccine development to reduce rheumatic heart disease in Africa. 19th Annual SA Heart Conference, Sun City, North-West Province, South Africa. October 4-7, 2018.

Salie MT, Moloi H, Muhamed B, Engel ME. Reducing the Burden of Rheumatic Heart Disease: Prevalence of Group a Streptococcus emm Clusters in Africa Suggests Adequate Coverage by 30-Valent Vaccine. American Heart Association 2018, Chicago, USA. 05-08 November 2018.

Mogamat Salie, Kelin Engel, Kimona Rampersadh, Dylan Barth, Babu Muhamed, Liesl Zühlke, James Dale, Mark Engel. The AFROStrep registry: Report on the prevalence of pharyngeal group A streptococcal emm-types in Cape Town, South Africa. 14th Pascar Congress, in association with 20th Annual Congress of the SA Heart Association, Sandton Convention Centre, Johannesburg, South Africa. 31 October – 3 November 2019.

Salie T, Engel K, Rampersadh K, Barth D, Muhammed B, Zühlke LJ, Dale JB, Engel ME. The AFROStrep Registry: Report of the prevalence of pharyngeal group A streptococcal emm-types in Cape Town, South Africa. 12th International Meeting on Microbial Epidemiological Markers (IMMEM XII), Dubrovnik, Croatia. 18-21 September 2019.

Salie MT, Rampersadh K, Muhamed B, Engel K, Zhulke L, Dale J, Engel ME. Systematic Review: Utility of Human Immune Responses To Group A Streptococcus Antigens As A Diagnostic Indicator For Acute Rheumatic Fever. 15th Pascar Congress, in association with Kenya Cardiac Society, Mombassa, Kenya. 22–25 November 2021. Cardiovasc J Afr. 2021; Submission No 84: Suppl:p16.; American Heart Association's Scientific Sessions 2021, Boston, USA. 13-15 November 2021. Circulation 144 (Suppl 1), A13788-A13788.

Non-peer reviewed

Rapid GAS global review

Salie, Taariq & Engel, Mark E. 2020. Rapid review of Global Strep A emm types. University of Cape Town. Dataset. https://doi.org/10.25375/uct.13056074.v1

ABSTRACT

Group A streptococcus (GAS) vaccines, as a primary prevention strategy, have the potential to prevent the spread of the bacteria, thus limiting post-sequelae diseases such as acute rheumatic fever (ARF) and rheumatic heart disease (RHD). Due to the limited supply of penicillin, adherence issues and problems associated with the diagnosis of post-sequelae diseases, a vaccine is thought to be the way forward on reducing the morbidity and mortality of ARF and RHD in low-resourced areas across the world. The latest vaccine formulation, a 30-valent M protein-based vaccine, currently in clinical trials is, however, challenged by the sheer volume of *emm* types (>230). This, along with the fact that significant *emm* types change over time, results in many gaps in protective efficacy. This body of work sought to contribute much-needed evidence in the quest for GAS vaccine development.

This thesis provides a comprehensive understanding of the bacteria and its complexity regarding its *emm* type epidemiology and underlying immune effects within a population in Cape Town. This information could be used to make added improvements for the early detection and diagnosis of GAS infections, and further inform vaccine formulations, hopefully to ultimately reduce the burden of RHD in high-risk populations.

The Masters' thesis comprises four linked studies: a systematic review summarising the molecular epidemiology of GAS on the African continent, a longitudinal study, over a 2-year period of the molecular epidemiology of GAS, a systematic review assessing the association of GAS antigens with ARF and, an ELISA-based assessment of the human immune response to common GAS antigens. Chapter 1 provides the background and rationale for undertaking the study. Chapters 2, 3, 4, 5 report the respective studies described above. Finally, Chapter 6 serves to provide a context for the work with recommendations for further research and implications for clinical practice.

Study I, published in *mSphere 2020*, highlights the dearth of epidemiological data found across the African content despite the high burden of GAS infections and post-sequalae diseases. Data, only available from five countries, indicated noticeable gaps in current vaccine formulations; for example, the 30-valent M protein vaccine (StrepAnova vaccine), only affords 58.22% protective coverage.

Study II serves to update an earlier study undertaken in Cape Town, South Africa, showing a continued significant prevalence of GAS in 256 children presenting with sore throat at community clinics. Among the 83 GAS strains isolated, characterisation through the *emm* typing procedure, documents a potential vaccine coverage of around 72%.

Study III, published in *Frontiers in Cardiovascular Medicine 2021*, provides evidence confirming that a recent GAS infection is associated with increased levels of antibody titres to GAS antigens, SLO and DNase B in cases of ARF in comparison to controls. The study further indicates caution in the use of the upper limit of normal (ULN) when testing antibody responses and, rather, recommends the use of sequential sampling and testing sera against a panel of GAS antigens.

Study IV provided evidence-based support for the utility of GAS-shared antigens and M peptides in documenting the characteristics of human immune responses following GAS infection. This study concludes that the array of GAS-specific antibody responses to GAS infection is broad, individuals demonstrate GAS-specific antibody responses in the absence of symptomatic GAS infection, GAS-negative individuals exhibit pre-existing antibody levels and lastly, having a panel consisting of five shared GAS antigens, increased the overall sensitivity of predicting a preceding GAS infection to 73.7%.

In conclusion, this thesis provides valuable insights into GAS pathobiology, serving to aid in the development of effective, safe and affordable GAS vaccines for global deployment. It also serves to emphasize the need for population-based studies in Africa, endemic to GAS infection with a high burden of ARF/RHD. Lastly, this study provides empirical support for the value of sequential sampling in ascertaining recent infection, suggesting a panel of at least five antigens in diagnostic assessment.

Chapter one

Background

1. Streptococcus pyogenes (Group A Streptococcus)

1.1 Bacteriology of GAS

Streptococcus pyogenes, also referred to as Group A Streptococcus (GAS or Strep A), belongs to the genus, *Streptococcus* which comprises of Gram-positive, fastidious, coccoid-shaped cells that could be found in pairs or long chains [6]. GAS is a facultative anaerobe with complex growth requirements, specifically medium supplemented with blood for optimum growth of the bacterium. GAS is a strict human pathogen that plays a role in the development of a variety of human infections and diseases such as pharyngitis (sore throat), scarlet fever, impetigo (skin infection), to more severe invasive diseases such as necrotizing fasciitis, Streptococcal toxic shock syndrome (STSS) and post-streptococcal autoimmune sequelae such as post-streptococcal glomerulonephritis (PGSN), acute rheumatic fever (ARF) and rheumatic heart disease (RHD) [7].

1.2 History of GAS

Theodor Billroth first described Streptococcal infections in 1874, isolated from erysipelas and skin infections. It was described to be small coccoid organisms found in pairs or chains [8]. However, the official entry of Streptococci infections was identified by Louis Pasteur, in which the bacterium was isolated from puerperal fever patients and demonstrated that the organism was an etiological agent that caused the highest burden of mortality in women and newborns [9]. It was in 1884 that the organism was named *Streptococcus pyogenes* by Friedrich Julius Rosenbach, whom examined suppurative lesions [10]. Rebecca Lancefield however, performed a series of identification tests using surface antigens in order to differentiate the Streptococci into various subgroups using letters, A through X [11]. Group A Streptoccoci or *Streptococcus pyogenes* was then further divided into various strains based on the surface protein called the M protein and identified over 50 different *emm*-types [12]. Since then, there has been approximately more than 230 *emm*-types identified by a series of molecular typing methods [13]. Lastly, Lancefield provided evidence that the M protein plays a major role in the bacterium's virulence due it its antiphagocytic properties, allowing the organism to evade the immune system [12].

1.3 GAS pathogenesis

In order for GAS to be the causative agent of such a wide array of diseases in the human host, the bacterium must be able to adapt to a diverse range of physiological conditions. To infect the host and successfully colonise the oropharynx, the bacterium must overcome the first line of defence, saliva. If an invasive strain of GAS infects the host, the organism must be able to disseminate from the original

site of infection to survive in blood. Thus, GAS has evolved, implementing evasive strategies and complex operational mechanisms to overcome the host defence systems. GAS has a wide arsenal of virulence factors contributing to the host-pathogen interactions, for the survival of the organism and ultimately, the cause of infection (Figure 1) [2].

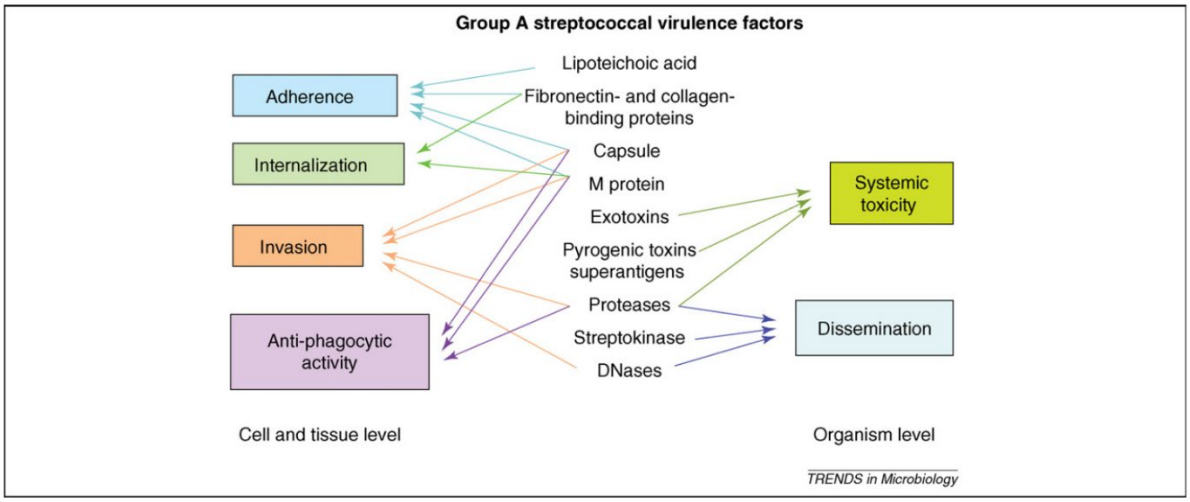


Figure 1 - Illustrating the role of GAS virulence factors played in pathogenesis. Adopted from Tart, Walker [2]

Lipoteichoic acid is an amphipathic molecule which acts as an adhesin to buccal epithelial cells that binds to fibronectin, an epithelial binding receptor [14]. Fibronectin- and collagen-binding proteins play a role in both adhesion and invasion, which is comprised of protein F, SfbI, FbaA and FbaB, the latter being associated with invasive isolates [15]. It has been shown that GAS strains are able to express more than one binding protein; the more binding proteins the strain is able to express, is directly proportional to its virulence [16]. The hyaluronic acid capsule acts as a protective covering that resists phagocytosis and also assists in adherence to CD44 on epithelial cells [17]. Streptococcal pyrogenic exotoxins comprises of superantigens that are mitogenic for certain T cell subsets and does not require processing by antigen-presenting cells (APC's). Superantigens has the capacity to bind the major histocompatibility complex II molecule, stimulating the influx of a large number of T cells, resulting in the secretion of cytokines and thus, associated with invasive infections [18]. Whereas, proteases, streptokinases and DNases are secreted by the bacterium contributing to virulence by damaging the membranes of cells within the host, disrupting the complement system from recruiting phagocytic cells [19] and avoiding destruction by neutrophils [20].

The most important virulent factor would be the cell-surface M protein which provides GAS with the ability to persist in these infected tissues by resisting phagocytosis and other immune-evasive actions. The M protein binds to protein factor H or fibrinogen, resulting in the production of plasmin, that degrades polymorphonuclear leukocytes [21].

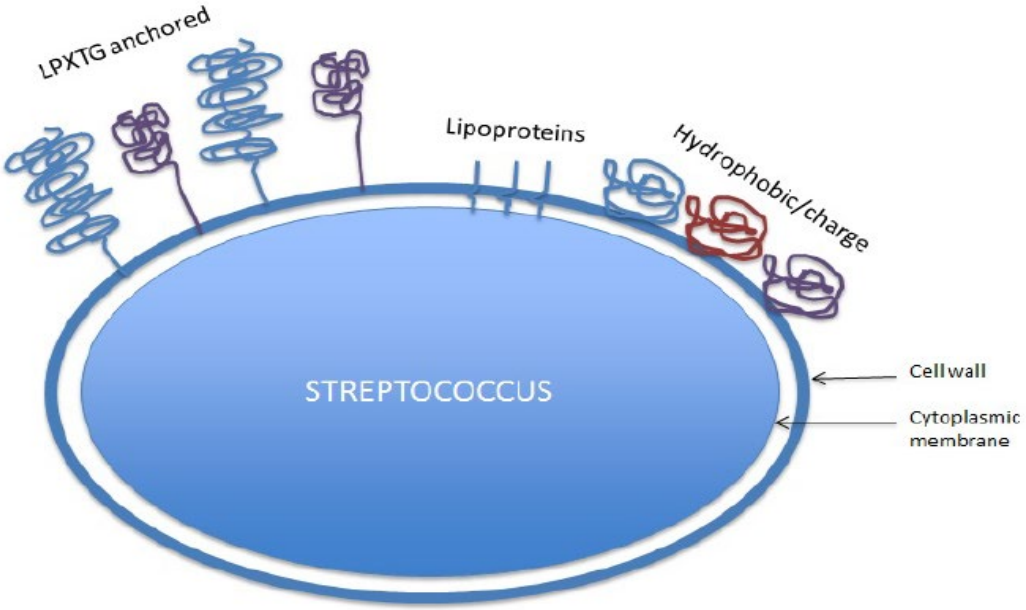


Figure 2 - Association of proteins with the cell surface of GAS. Adopted from Fischetti [3]

The M protein is an elongated α-helical coiled-coil structure that is anchored to the cell wall via the C-terminal LPXTG motif (Figure 2). The size of the M protein varies from 750bp-1200bp (41-80kDa), due to extensive repeat regions at the DNA and protein level [22]. Whereas, the N-terminal region also referred to as the A repeat region confers *emm* type specificity. This region is comprised of around 50 amino acids differentiating to a total of over 230 *emm* types reported in a systematic review completed by Steer, Law [13]. However, there is no information or data providing insights as to how the bacteria allows for changes to occur within this region [23].

2. Molecular epidemiology of GAS

2.1 Serology of GAS

Serological classification of GAS into subtypes are performed for epidemiological studies or outbreak situations, providing information about the evolutionary relatedness between strains of the same type and could be used in the development of a vaccine [24]. The conventional typing schemes of GAS are based upon the determination of antigen specificity of surface-exposed proteins, namely T and M proteins. The T protein is trypsin-resistant that forms part of the pilus structure that could be subgrouped into T-types. A single T-type pattern may represent a group of strains of different *emm*-types [25]. Falugi, Zingaretti [26] described a T typing method by characterizing the *tee* gene which is used as the orthodox method. M protein serotyping of GAS is considered by many as the gold standard typing scheme as it classifies strains into subgroups based on the differences located within the N-terminal region of the M protein [27, 28]. This typing scheme was first conducted by Lancefield whom used a precipitation method, utilising M-protein specific antisera [29].

A study completed by Beall, Facklam [27] adapted the M typing scheme from Lancefield to a more conventional molecular method, established based on the amplification of the *emm* gene with a specific primer set. Following PCR amplification, there would be differences in the size of the PCR products due to sequence variation within the N-terminal region. The products are sequenced and assigned an *emm* type with comparison to a large dataset of *emm* sequenced genes also including various other β-haemolytic Streptococcal species (Groups C, G and L Streptococci) available on the CDC website (http://www2a.cdc.gov/ncidod/biotech/strepblast.asp) [30].

Multi-locus sequence typing (MLST), an alternative typing scheme developed for characterizing GAS into groups to provide relatedness amongst strains and their pathogenetic capabilities. MLST is used to provide further insights to the genetic relationships between organisms of a bacterial species [31]. It involves the detection of housekeeping genes as they are vital for bacterial function. Clones can therefore be described as those identified with having common alleles at each of the respective housekeeping genes (n=7), further distinguished into a sequence type (ST) [32].

2.2 Prevalence of *emm* types

The characterization of prevalent *emm* types of GAS in a specific region, forms the groundwork for tracking the epidemiology of the organism. Furthermore, gaining insight to the pathogenesis, transmission and the ability of the bacterium to survive in the host, would assist the prevention and ultimately, eradication of the harmful pathogen [33].

Over the last decade, many epidemiological studies of GAS have been undertaken, with the majority coming from the USA, Canada and Western Europe. However, based on these studies, the distribution amongst *emm* types differ significantly based on region and the standard of living within the particular region (i.e. developed vs developing countries). A systematic review completed by Steer, Law [13] provided evidence in reporting the global distribution of *emm* types from developed and developing countries, noting the 25 most prevalent *emm* types within each region (High-income countries, Africa and the Pacific region), illustrating the variability amongst GAS strains. The heterogeneity amongst prevalent *emm* types across the globe is summarized in Table 1, depicting the top three prevalent *emm* types within each specific country, with the African continent producing the least amount of data on the surveillance of GAS.

Table 1 - Summary of studies reporting the distribution of *emm* types from regions across the world

Continent	Year	Country	No. of isolates	No. of *emm* types	Prevalent *emm* types	Reference
Africa	2005	Ethiopia	90	44	3, 5, 29/208	Abdissa [34]
	2013	Mali	255	67	18, 65, 25	Dale, Penfound [35]
	2012	South Africa	143	35	48, 12, 4	Muhamed [36]

	2016	Kenya	357	88	44, 65, 18	Seale, Davies [37]
Asia &	2006	Taiwan	677	44	12, 4, 1	Chiang-Ni, Zheng [38]
Pacific	2009	Taiwan	1218	23	12, 4, 1	Chiou, Liao [39]
	2018	India	348	45	28, 49, 63	Jose, Brahmadathan [40]
	2010	Korea	299	11	4, 12/75, 89	Koh and Kim [41]
	2011	Japan	150	15	12, 1, 28	Ogawa, Terao [42]
Oceania	2014	New Caledonia	318	47	95/76, 25, 1	Baroux, D'Ortenzio [43]
	2007	Australia	340	39	55, 11, 71	McDonald, Towers [44]
	2010	Australia	635	73	55, 11, 91	Richardson, Towers [45]
	2017	New Zealand	376	36	12, 1, 41	Mhlanga, Sharp [46]
	2015	New Zealand	2861	111	1, 49, 81	Williamson, Morgan [47]
North	2009	Canada	1434	33	12, 1, 28	Shulman, Tanz [48]
America &	2016	USA	9557	99	1, 12, 28	Nelson, Pondo [49]
Europe	2016	Alaska	422	51	1, 82, 49	Rudolph, Bruce [50]
	2012	Spain	806	21	4, 1, 28	Rubio-López, Valdezate [51]
	2014	France	637	87	4, 1, 28	Plainvert, Dinis [52]
	2015	Greece	1282	35	1, 12, 77	Koutouzi, Tsakris [53]
	2017	Germany	719	46	1, 28, 89	Imohl, Fitzner [54]

2.3 Importance of M typing

As stated above, the surveillance of GAS and the characterization of prevalent *emm* types within a population forms the foundation for the prevention of GAS infections [33]. It is also important to note that the distribution of GAS *emm* types in a particular region has an impact on vaccine development, as it will assess the potential coverage of the putative 30-multivalent M protein-based vaccine [55]. The *emm* typing procedure also allows for classifying the complex bacterium into groups, that could possibly be accountable for a particular trait or disease-causing effect [56]. Thus, following epidemiological studies, conclusions could be made in identifying patterns in the spread of disease through the emergence of a new strain or clone [57] and, to determine which particular *emm* type is capable of causing a specific disease following infection [58].

Epidemiological studies of GAS completed by Anthony, Kaplan [59], Bessen, Sotir [60] and a report by Cunningham [61], provided evidence that there are certain *emm* types which have a strong inclination towards pharyngitis but have not been shown to cause impetigo and vice versa. Therefore, strains can be categorized based on the clinical phenotypes caused. It has also been stated by previous reports, that climate affects disease manifestations, where GAS pharyngitis is found in temperate regions, peaking in the winter season. Whereas, GAS impetigo is dominant in tropical areas, peaking in summer [62, 63].

As stated above, the M typing technique is useful in identifying the causal *emm* type strain responsible for the outbreak of a disease. For instance, in Southern Taiwan, patients from a common hospital with symptoms of STSS were included into a study completed by Lin, Chang [64] whom after M typing, identified the most prominent *emm* type for causing this invasive disease was *emm*102. According to a study completed by Lacy and Horn [65], identified that a health care worker (HCW) was infected with an invasive strain of GAS which resulted in the development of STSS, after caring for a patient infected with GAS. In this case, the M typing procedure determined the clonality between the strains in both the HCW and the patient.

The *emm* types obtained from the M typing procedure could also be connected to other typing systems such as MLST and the *emm* cluster system, to measure the diversity amongst isolates within a given region [66]. A study completed by Bowen, Harris [67] determined the GAS transmission within a remote community, whether GAS was being transmitted from within the household and/or throughout the community. They performed M typing and MLST procedures in tandem, resulting in a particular *emm* type with a corresponding ST. This allowed the authors to conclude that strains of the same *emm* type carrying the same ST were actually clones and transmitted from one individual to the next. The *emm* cluster system devised by Sanderson-Smith, De Oliveira [4] is directly correlated with the *emm* pattern typing, that distinguishes *emm* types into three main groups. These *emm* patterns are determined by the arrangement of *emm* genes and *emm like* genes, resulting in patterns, A-C, D and E [66]. It has been shown that the *emm* patterns display tissue tropism, where patterns A-C has a high tropism for pharyngitis, pattern D for impetigo and pattern E could be isolated from both tissue sites [62].

2.4 *emm* cluster system

The prevention strategies of GAS infections and post-sequelae diseases have been difficult to implement in low-resource settings as overcrowding remains a prevalent issue as well as the diversity of *emm* types within these regions [68]. Thus, the *emm* cluster system has been developed to serve as a functional classification scheme for the ~230 *emm* types into groups that could support the development of a GAS vaccine. The study obtained over 1000 isolates from thirty-one countries, resulting in 175 heterologous *emm* types. A single sequence of the entire M protein of each of the isolates were included for phylogenetic analysis which resulted in the formation of two clades, clade X and Y, representing 85 and 84 proteins respectively with two outlier proteins. Sanderson and colleagues focused on these heterologous M proteins and classified them into 48 *emm* clusters based on similar binding properties to host proteins; where 26 clusters were made up of a singular *emm* type. It was clear that the two respective clades, had distinct binding-functional profiles; clade X was restricted to the binding of immunoglobulin and C4BP, whilst clade Y isolates were restricted to binding to plasminogen and fibrinogen host proteins. They then focused on specific binding motifs within each clade to further distinguish the M proteins into clusters, A-C, D and E (Table 2). It has been stated that the immunity to

GAS infection is *emm* type specific [69, 70] and thus, the new *emm* clustering system complements the M typing system, serving as a cross-protection framework, for better characterization of GAS.

Table 2 – The *emm* clusters and their corresponding *emm* types. Adopted from Sanderson-Smith, De Oliveira [4]

***emm* types**	***emm* cluster**	**Clade**
4, 60, 78, 165 (st11014), 176 (st213)	E1	Clade
13, 27, 50 (50/62), 66, 68, 76, 90, 92, 96, 104, 106, 110, 117, 166 (st1207), 168 (st1389)	E2	X
9, 15, 25, 44 (44/61), 49, 58, 79, 82, 87, 103, 107, 113, 118, 144 (stknb1), 180 (st2460), 183 (st2904), 209 (st6735), 219 (st9505), 231 (stNS292)	E3	
2, 8, 22, 28, 73, 77, 84, 88, 89, 102, 109, 112, 114, 124, 169 (st1731), 175 (st212), 232 (stNS554)	E4	
34, 51, 134 (st2105), 137 (st465), 170 (st1815), 174 (st211), 205 (st5282)	E5	
11, 42, 48, 59, 63, 65 (65/69), 67, 75, 81, 85, 94, 99, 139 (st7323), 158 (stxh1), 172 (st2037), 177 (st2147), 182 (st2861UK), 191 (st369)	E6	
164 (st106), 185 (st2917), 211 (st7406), 236 (sts104)	Single type *emm* clusters	
36, 54, 207 (st6030)	D1	Clade
32, 71, 100, 115, 213 (st7700)	D2	Y
123, 217 (st809)	D3	
33, 41, 43, 52, 53, 56, 56.2 (st3850), 64, 70, 72, 80, 83, 86, 91, 93, 98, 101, 108, 116, 119, 120, 121, 178 (st22), 186 (st2940), 192 (st3757), 194 (st38), 208 (st62), 223 (stD432), 224 (stD631), 225 (stD633), 230 (stNS1033), 242 (st2926)	D4	
97, 157 (stn165), 184 (st2911)	D5	
46, 142 (st818)	A-C1	
30, 197 (st4119)	A-C2	
1, 163 (st412), 227 (stil103), 238, 239	A-C3	
12, 39, 193 (st3765), 228 (stil62), 229 (stmd216)	A-C4	
3, 31, 133 (st1692)	A-C5	
5, 6, 14, 17, 18, 19, 23, 24, 26, 29, 37, 38 (38/40), 47, 57, 74, 105, 122, 140 (st7395), 179 (st221), 218 (st854), 233 (stNS90), 234 (stpa57)	Single type *emm* clusters	
55, 95, 111, 215 (st804), 221 (stCK249), 222 (stCK401)	Single type *emm* clusters	Outliers

504

3. Streptococcal pharyngitis

GAS pharyngitis is limited to the severity it may cause. However, current understanding is that due to the molecular mimicry of the bacteria, following repeated episodes of pharyngitis, may drive the progression to the autoimmune diseases (ARF and RHD) [71]; the post-sequelae diseases will be discussed later in the review. GAS pharyngitis transmission occurs through nasal and salivary secretions from an infected individual to others in the same vicinity [72], attributing to the prevalence of disease cases. It is also difficult to document the global burden of GAS pharyngitis infections as not all cases are reported to clinical practitioners or hospitals [73].

3.1 Clinical manifestations

According to an annual report in the USA, physicians primarily attend to patients with upper respiratory infections, inclusive of acute pharyngitis [74]. The infection could be caused by either a virus or bacteria; with GAS the leading cause of bacterial pharyngitis [75]. However, making a clinical diagnosis between GAS pharyngitis and viral pharyngitis is difficult based on physical symptoms alone and thus, epidemiological testing for the physical presence of GAS in the pharynx should be completed for confirmation [76]. Pharyngitis also referred to as "sore throat" is a self-limited infection with physical symptoms that could be described as the inflammation of the pharynx with anterior cervical lymphadenitis, exudative tonsillopharyngitis, abdominal pains and general malaise of fever [77]. Whereas, a viral pharyngeal infection may include symptoms of rhinorrhoea, hoarseness, cough, conjunctival irritation and/or even diarrhoea.

3.2 Epidemiology of GAS pharyngitis

It has been reported by numerous studies that the peak incidence age group for GAS pharyngitis is between 5-15 years old [75, 77], affecting ~20% of individuals within Cape Town, South Africa [78]. GAS pharyngitis is more likely to occur in regions of temperate climates during spring and winter, with outbreaks prone to occur in places with close human-to-human contact, especially schools and low socio-economic households [79]. The abundance of epidemiological data on GAS pharyngitis is derived from first world countries, whilst data from developing countries are lacking. This is a major concern as stated by previous reports, that the epidemiology of GAS is vast across different geographical areas [62]. It is therefore important for areas, where the burden of GAS diseases are rampant, to report on the epidemiological data, as this will assist vaccine development programmes.

A study completed by Engel, Muhamed [78], reported on the epidemiology of GAS strains obtained from patients presenting to the clinical practitioner with symptoms of sore throat located in Cape Town, South Africa. They enrolled 742 patients, aged 3-15 years between May 2008 - September 2011. Throat swabs were obtained and resulted in a 21.6% (n=160 GAS strains) GAS culture-positive rate, of which,

there was a total of 26 heterologous *emm* types recovered. The top 5 prevalent *emm* types recovered were as follows, *emm*48, *emm*89, *emm*4, *emm*12 and *emm*75.

A study completed in Bamako, Mali by Tapia, Sow [80], studied the epidemiology of GAS amongst school children, 5-16 years of age between May 2006 - September 2009. All children experiencing symptoms of sore throat were advised to visit the on-site study clinicians, which resulted in the enrolment of 1,418 students in total and contributed to 449 GAS-positive cultures. Seventy heterologous *emm* types were represented, of which the top 5 prevalent *emm* types were as follows, *emm*65, *emm*18, *emm*55, *emm*42 and *emm*81. Dale, Penfound [35], also completed a study in Bamako, Mali, in which they tested the protective coverage of the 30-valent M protein vaccine within a population at high risk of GAS disease, specifically RHD. This was completed by obtaining pharyngeal isolates from school children and subjecting it to rabbits that has been immunized with the 30-valent vaccine to evaluate the bactericidal activity. They recovered 372 GAS strains, representing 67 heterologous *emm* types, of which only eighteen were covered within the vaccine. However, following the bactericidal activity assay, they witnessed cross-protection by the antisera, resulting in a potential coverage of 84% across all isolates of GAS within the region.

Abdissa [34] studied healthy children residing in Ethiopia during November 2004 - January 2005 to understand the effects of GAS carrier strains within the population. They obtained data from 937 children aged between 6-14 years old, of which 29% reported that they had an episode of pharyngitis within the last month and notably, most did not seek any professional medical attention. They recovered 167 β-haemolytic streptococci isolates, of which, 90 belonged to GAS, Group G (n=30), Group C (n=21) and the remainder belonged to Group D-G.

These epidemiology studies within the developing world are important as it provides the surveillance of GAS strains recovered from children presenting symptoms of sore throat within the region, forming the foundation of developing prevention strategies.

4. Acute rheumatic fever and rheumatic heart disease

As stated above, repeated GAS exposure may trigger ARF due to an autoimmune response and repeated episodes of ARF, may result in the final presentation of damaged cardiac tissue, RHD [1]. This autoimmune response thought to be primarily due to the molecular mimicry exhibited by the bacteria's M protein with muscle tissue located in the heart. More specifically, the M protein shares an α-helical structure with that of myosin, a cardiac protein [81]. The M protein stimulates an immune response for the production of opsonizing antibodies, that cross-react with cardiac tissue, resulting in heart valves deterioration [82].

Symptoms of ARF could occur within 2-6 weeks of the initial non-invasive infection, resulting in the onset of fever, arthritis, carditis and chorea [83]. Arthritis mainly affects large joints such as the knees and

elbows, with a mysterious relocation or migratory action. ARF causes a pancarditis, affecting the pericardium, epicardium, myocardium and endocardium, the latter is where the mitral and aortic valves are situated [84]. Chorea, a symptom that only occurs in 30% of ARF patients, in which involuntary movements of the body occurs, affecting the face and limbs [85]. These point to the persistent uncertainty revolving around the pathogenesis of ARF, thus contributing to the absence of a clear-cut diagnostic assessment for ARF. Notwithstanding, patients are assessed based on clinical criteria referred to as the Jones criteria, that has been adjusted to suit all types of populations [83].

RHD patients experience the effects after long-term damage caused by the abovementioned symptoms, specifically that of carditis [86]. Carditis results in tissue lesions specifically in the aortic and mitral valves, the latter being the most common. Thus, the diagnosis of RHD is through evident damage of heart tissue which can be seen through echocardiography, the gold standard assessment [87]. Mitral regurgitation is the most common presentation of RHD, portraying valvular thickening, evident on 2D echocardiography, whereas aortic regurgitation according to previous reports [88] are less common, occurring in only 4.5% of the total RHD patients within the community.

4.1 Epidemiology of ARF and RHD

The disease burden of ARF and in turn, RHD is widespread with the bulk of cases found in resource poor regions. However, due to inadequate surveillance systems in and around the developing world, incidence rates are poorly documented. This could be related to healthcare systems lacking in infrastructure, facilities, equipment and training of health professionals [75]. The peak incidence age of ARF is between 5-15 years of age and estimated that if untreated, 60% of individuals with confirmed ARF will go on to develop RHD later in life [89], with a peak incidence between 20-29 years of age [90]. Global disease estimates according to Vos, Barber [91], indicated that approximately 33 million individuals suffer from the disease with 275,000 deaths occurring yearly which increased to 40.5 million in 2019 with 306,000 deaths yearly [90]. Cases of RHD varies by region with Africa [88] and the Pacific regions [92] leading the prevalence charts. Whereas, in the USA and Western Europe, there has been a decline in the incidence rates of ARF and RHD, particularly due to uncrowded living conditions, better health care and mainstay supply of antibiotics [93].

A study completed by Zuhlke, Engel [88] conducted a systematic review providing evidence of the incidence cases of RHD reported throughout South Africa from 1994-2014. They identified that the overall symptomatic RHD incidence rate was 24.3 per 100,000 adult individuals (>13 years of age). Whereas, the asymptomatic RHD incidence rate recorded from schoolchildren was 20.2 per 1000 individuals per annum. They also identified that there is a high case fatality rate with RHD patients experiencing acute heart failure. The prevalence of RHD is still high within the region. However, they have shown a decrease in mortality rates over the past decade. With that, they have identified a combination of risk factors such as overcrowding households, poor infrastructure within the healthcare

system and poor diagnosis of disease. These risk factors play a pivotal role in the development of disease, beginning with the initial GAS infection, through to ARF and later on, RHD.

4.2 Diagnosis of ARF and RHD

It is of utmost importance to establish the true burden of ARF and RHD, to launch preventative methods for control, which lies in the detection through proper diagnosis. In order to increase the sensitivity of diagnosis, the community as a whole needs to be aware of the manifestations and diagnostic procedures. In South Africa, there are limiting factors resulting in insufficient notifications as to the awareness by the community and especially primary healthcare workers, reducing the efficacy of capturing the proper diagnosis of ARF, and thus incidence rates [94-96]. Of note, high-risk communities within developed country settings with proper investment infrastructure experience similar challenges [97].

Other issues with underdiagnoses of ARF in developing countries would be that most clinical infrastructure and resources are dedicated to other diseases such as HIV/AIDS, malaria and tuberculosis (TB) [98]. However, in recent times there have been progress in closing the gap between diseases, with the hope of combating all diseases in a combined approach benefiting the developing world. Another impediment is the inadequate training of medical personnel for the proper diagnosis of ARF [99]. Thus, awareness with appropriate guidelines and adequate training, will be effective in the control of ARF and RHD.

ARF has a number of clinical manifestations as listed above, and due to the differences associated with disease symptoms and various combinations thereof in different population groups, there is no one clear diagnostic measure. This has led to the development of the Jones criteria, with guidelines as to appropriately diagnose ARF and RHD [100], that has been modified by the American Heart Association (AHA) in recent times [101], made easier by the recent advancements in medical technology such as 2D echocardiography and Doppler flow assessments [83].

5. GAS Carrier State

GAS is also able to colonize the pharynx of an individual with no signs of infection. When a positive GAS culture is obtained from a healthy individual, they are referred to as a carrier of GAS [102]. However, serological evidence of the immune response is required to ensure that there is no rise in antibodies specific to GAS antigens. Measuring antibody titres also presents certain difficulties with analysis, as a previous study has shown that GAS carriers too, has an increase of GAS-specific antibodies [103]. This may be due to the fact that a symptomatic infection usually occurs before entering the carrier state, resulting in the elevated levels of the antibodies [104]. With that, it has been stated that a more accurate definition of a GAS carrier, would be an individual who does not present any clinical symptoms of infection, has received a certain dosage of antibiotics, after which, still provides a positive GAS culture [105].

The incidence rates of GAS carriers vary between populations which could be due to the living conditions and medical infrastructure. Proper antibiotic treatment of GAS nonsuppurative infections decreases the incidence rate of carriers. Another factor affecting the incidence rate of carriers, are the lack of surveillance systems, in which follow-ups are not scheduled with patients, thus no throat culture is obtained [105].

A study completed by Pichichero, Marsocci [106] studied the occurrence of GAS-positive throat cultures from paediatric patients presenting with sore throat, patients that were treated with penicillin and healthy children. They identified that children presenting with sore throat were 4.4% positive and those treated with penicillin were 11.3% positive for GAS. Whereas GAS only occurred in 2.5% of healthy children. In another study [107], with a longitudinal design aimed to characterize the carrier state of healthy school children over a 4-year period. Enrolled participants were instructed to provide a throat culture at 2-monthly intervals and when presenting symptoms of sore throat. They discovered that at any given time, that 16% of children located in the community were in fact, carriers of GAS, with an overall total of 40% GAS carriers over the 4-year period. When referring to studies that focused on microbiological failures, in which antibiotics cleared the symptoms however, failed to eradicate the bacteria; a 5-25% GAS carriage rate was described by Kaplan and Johnson [108] and Casey, Kahn [109].

6. Treatment and control of GAS diseases

The Infectious Diseases Society of America (IDSA) describes guidelines [110] which has been updated by Shulman, Bisno [111] for the control and management of GAS pharyngitis. According to the report, treatment recommendations involve the use of antibiotics at the correct dosage for an appropriate duration to eradicate the bacteria completely. The most common, cost-effective antibiotic for the treatment of GAS pharyngitis is penicillin or amoxicillin. However, for patients with allergies towards the agents mentioned, they would require cephalosporin, clarithromycin or azithromycin [112]. Usually, the duration for the prescribed antibiotics ranges from 5-10 days, dependent on dosage and severity of infection.

The treatment of ARF involves penicillin prophylaxis due to its cost and safety [113]. Depending on the severity of ARF, a specific dosage of a penicillin injection will result in the clearance of the GAS bacteria, reducing the risk of obtaining further GAS infections [114]. For further joint complications such as arthritis and/or arthralgia, anti-inflammatory medication is prescribed, such as ibuprofen and naproxen; the latter prescribed for older patients [115].

As for RHD, there is only treatment options available for the complications associated with the disease. Patients with damaged cardiac tissue may require β-blockers or angiotensin-converting-enzyme inhibitor therapies to avoid further complications. Patients with atrial fibrillation requires warfarin with anticoagulation and rhythm control effects [116]. While patients with critical valvular lesions may require

interventional treatment strategies such as surgery or cardiac catheterism, in which the valves are replaced either mechanically or bioprosthetically [117].

The control and management of GAS diseases involve four discrete strategies, primordial, primary, secondary and tertiary prevention (Figure 3). The primordial prevention mechanism is based on the organism's transmission method by secretory droplets. Therefore, this arm of prevention aims at

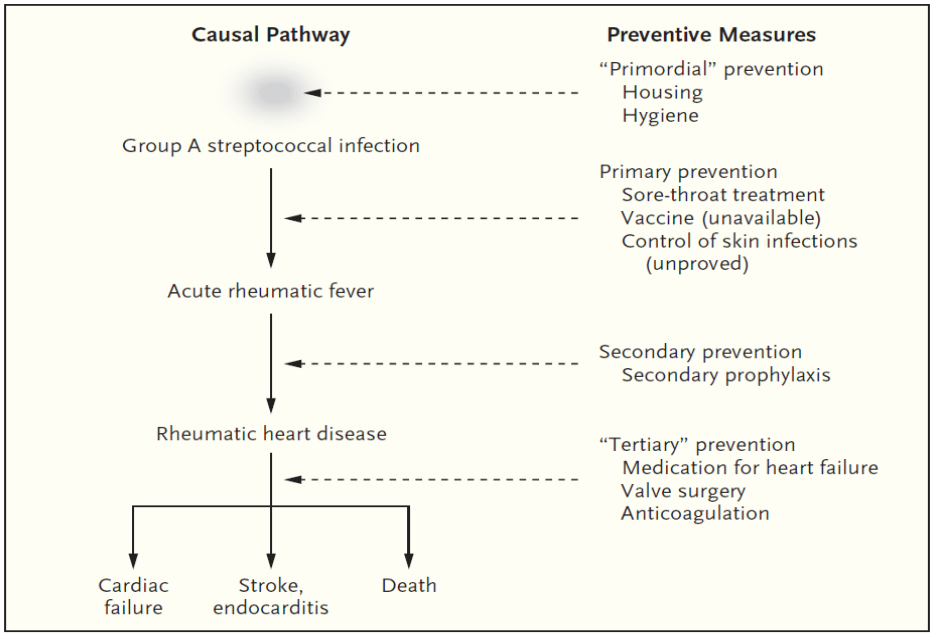


Figure 3 - Prevention strategies for combating GAS infections. Adopted from Carapetis [1]

reducing impoverished living conditions such as household crowding [118, 119]. Current primary prevention strategies aim to inhibit GAS colonization as well as the spread of the organism with the administration of antibiotics, to eradicate the bacteria before it can cause any autoimmune diseases [120]. Examples of countries implementing successful primary prevention programmes include two French Caribbean islands, Martinique and Guadeloupe, which saw a ~75% reduction in ARF [173] and Cuba which reported a 5-fold reduction in ARF fever incidence following GAS infection treatment[172]. Secondary prevention involves the prohibition of recurrent episodes of ARF, with the implementation of secondary prophylaxis, making use of benzathine penicillin G (BPG). BPG is administered monthly, intramuscular for a minimum period of 10 years [114]. Tertiary prevention involves the use of surgical interventions as a means to prevent loss of life due to the effects of post-sequelae diseases [121].

7. GAS Vaccine development

The global burden of GAS diseases are substantial, and excess mortality from ARF, RHD and invasive infections are significant [86]. Thus, the need for a safe, effective, and affordable GAS vaccine. There has been evidence provided that protective immunity may occur following a natural GAS infection and this concept could be implemented into the vaccine. However, vaccine development strategies are negatively impacted by the complexity of the bacteria, based on the number of *emm* types, differences

in prevalent *emm* types across geographical regions and the wide array of clinical manifestations of GAS infections [13]. Another factor is the use of autoimmune epitopes in the vaccine, that could stimulate the disease that it was intended to prevent.

Protective immunity towards GAS infections has two major mechanisms, 1) organisms entering the host can be blocked from attachment to mucosal surfaces by IgA specific to the C-repeat region of the M proteins and 2) the bacteria, once it has entered the host tissues, is effectively eliminated by opsonization with type-specific antibody and complement, with subsequent phagocytosis and killing [12]. One mechanism prevents colonization, while the other mechanism prevents multiplication in the host and elimination of the bacterium in host tissues or blood. M protein-based vaccine strategies have taken advantage of the type-specific epitopes conferring these two modes of protection, stimulating the production of opsonic antibodies (MW, Cunningham., 2000). However, the multivalent vaccine developed by Hu, Stroop [69] is impeded by the diversity of serotypes (>230 *emm* types) and only able to provide coverage for prevalent *emm* types from developed countries. This results in suboptimal protection of the multivalent M protein-based vaccine against GAS infections globally [13]. A solution to this problem is the development of vaccines that contain common protective antigens that are shared by many or all serotypes of GAS, in addition to multivalent M peptides [68]. Indirect evidence suggests that repeated exposure to multiple serotypes of GAS results in relative resistance to GAS infections in older individuals, that are conferred by immunity to shared antigens (Dale et al., 2016). The overall hypothesis to be tested is that at least some antigens of GAS are immunogenic following natural infection and that these antigens may be associated with subsequent protection against new GAS acquisitions. Antigens shown to be highly immunogenic could be included in next-generation vaccines.

However, little is known about the human immune response following natural infection with GAS or gene expression profiles in GAS isolates during asymptomatic carriage vs active infection. Thus, efforts should now focus on putative target antigens that could produce antibodies capable of providing some sort of immunity. This requires the detailed knowledge of the molecular mechanisms and antigens utilized by GAS to adhere to host cells and those involved in evading the immune response [122].

8.1 Human immunological responses through ELISA

Enzyme-linked immunosorbent assay, also referred to as ELISA, is used for the detection and quantification of proteins, peptides, hormones and antibodies. The technique involves the immobilization and binding of reagents with compliment reactants all within a well-plate [123]. The process encompasses for 4 main steps namely; coating, blocking, detection and analysis, with numerous washing steps in-between to rid the unwanted material. There are different types of ELISA formats, either direct, using a labelled primary antibody or indirect, using a labelled secondary antibody, the latter being the approach taken in this thesis. Indirect ELISA may be used for the detection of antibodies. Briefly, it involves the adsorption of antigens to the walls of the microtiter plate [124]. The test serum is

731 then added and if the complement antibody is present, it will bind to the adsorbed antigen and the
732 unbound-nonspecific material is washed away. The attached complex (antigen-antibody) can be
733 detected with the use of a chromagen; a colourless substrate that reacts with the enzyme conjugate
734 attached to the secondary antibody. The latter binding only if the primary antibody is bound. The
735 substrate is then added to the complex, which if the indicator-antibody (secondary antibody-conjugate
736 enzyme) is bound, will react to the enzyme attached, resulting in a colour change. The optical density
737 (OD) readings of each well is determined, at the corresponding peak wavelength depending on the
738 substrate used [125].

739 ELISA assays will be conducted to determine the antibody production towards a select panel of putative
740 vaccine candidates; antigens shared by many or most GAS serotypes to explore the antigen specificity
741 and kinetics following the acquisition of GAS (Table 3).

Table 3 - GAS antigens proposed for this study

Antigen	Bacterial Location	Function	References
Type-Specific M peptides (estimated 25 synthetic peptides)	Cell surface	Opsonic epitopes	Hu, Stroop [69]
M-related peptides (Mrp) Groups I, II, III	Cell surface	Opsonic epitopes	Dale, Niedermeyer [126]
GAS adhesion and division protein (SpyAD)	Cell Surface	Cell Division and Adhesion	Gallotta, Gancitano [127]
C5a peptidase (SCPA)	Cell surface and Secreted	Cleaves C5a	Cleary, Prahbu [128]
Serine esterase (SSE)	Secreted	Tissue Invasion	Liu, Zhu [129]
Streptolysin O (SLO)	Secreted	Hemolysin	Johnson, Kurlan [130]
Deoxyribonuclease B (DNase B)	Secreted	Degrades NETs	Johnson, Kurlan [130]

742

743 ## 8.2 GAS cell-surface antigens

744 ### 8.2.1 Type-Specific M Peptides

745 The type specificity of GAS isolates is largely determined by the epitopes located in the amino-terminal
746 end of the M protein, 40 to 50 amino acid residues [13]. Antibodies reacting to these regions of the M
747 protein has shown the greatest bactericidal activity [23]. Thus, the approach taken in the multivalent
748 vaccine was to combine the predominant serotypes of GAS, specifically the M peptides that would elicit
749 phagocytic antibodies [131]. Hu, Stroop [69], constructed a 26-valent M protein-based vaccine using
750 recombinant technology, to produce each of the selected serotypic-specific M peptides. The vaccine

was developed in such a way that there are four fusion proteins containing a link of six or seven M peptides, serving as their own carrier. Serotype selection was based on the epidemiological data that included isolates causing frequent uncomplicated pharyngitis, isolates frequently recovered from normally sterile sites and those associated with ARF/RHD. They also included the *Spa* peptide-fragment, which has been shown to elicit protective immunity [132]. The vaccine was examined by performing indirect immunofluorescent assays, which showed no cross-reactivity with host tissues. However, a study completed by Engel, Muhamed [78], demonstrated that in a population where the burden of RHD is highest, only 60% of the predominant strains are covered in the now, 30 valent M protein-based vaccine. Illustrating the need for improving vaccine design to extend vaccine coverage across all serotypes of GAS.

8.2.2 M-Related Proteins

The M-related proteins (Mrp) have been predicted to be included in the repertoire of surface antigens on heterologous GAS serotypes [35], which may increase the overall vaccine protective coverage. A study completed by Courtney and Li [133] characterized Mrp as part of the M protein family and determined its virulent potential; operating in tandem with the M protein. The M protein family includes the *Emm*, *Mrp* and *Enn* genes, however their expression varies amongst *emm* types. According to the evidence provided by Courtney, Hasty [134], the co-expression of the M protein with Mrp results in increased virulence, in which Mrp binds to fibrinogen in the host plasma enabling growth of the bacteria. Dale, Niedermeyer [126] elaborated on the characterization of Mrp and determined slight differences occurring in the N terminal region of the protein, comprises 4 Mrp families, MrpI, MrpII, MrpIII and MrpIV. Mice models were used to test the protective efficacy of the rabbit Mrp-antisera following purposeful GAS infection of heterologous serotypes. They also provided evidence that Mrp antibodies increases with age, due to the occurrence of repeated exposure to the surface-bound antigen.

8.2.3 Streptococcus pyogenes Adhesion and Division Protein (SpyAD)

A study completed by Gallotta, Gancitano [127] went on searching for putative vaccine antigens and to so, made use of mass spectrometry to identify peptides, protein arrays to determine immunogenic antigen-antibody complexes and lastly, flow cytometry for the quantification of antibodies produced. They identified a completely novel protein, referred to as the Streptococcus pyogenes Adhesion and Division protein or SpyAD. The protein plays a role in the cell division whilst also promoting GAS adhesion to human host cells. Gallotta, Gancitano [127] performed an in vitro assay and determined that SpyAD binds to collagen VI and keratin I, providing evidence that it may be involved in the initial weak attachment to host cellular exposed proteins. According to De Bentzmann, Plotkowski [135] whom states that the epithelium composition may alter once the host is infected and an inflammatory response is triggered. This may affect the rupturing of basal cells, exposing host receptors for GAS adhesion. Evidence provided by Bensi, Mora [136] illustrated the protective efficacy of SpyAD in a mouse model, producing

antibodies that reduced the colonization of GAS. It has also been predicted that the anti-SpyAD antibodies may disrupt the bacteria from effectively dividing and spreading.

8.3 Cell surface and secreted antigens

8.3.1 C5a Peptidase

Decades ago, a study completed Wexler, Chenoweth [137] discovered a surface antigen on GAS, highly specific for the host chemotaxin C5a, referred to as C5a peptidase or SCPA. According to Chen and Cleary [138] the *scpA* gene is conserved across serotypes of GAS, as they identified the SCPA protein expressed in >160 isolates, which represented 40 serotypes of GAS. The endopeptidase is responsible for cleaving C5a, a protein derived from downstream processes of the complement system for the initiation of an inflammatory response. C5a is a 74-mer chemotaxin peptide that relies on a chemical gradient for the stimulation of phagocytic polymorphonuclear leukocytes (PMNs) and recruiting them to the site of infection; acting as the first line of defence shown by the host for the prevention of bacterial colonization [139]. The elimination of the C5a peptide, provides the bacteria with time to fully establish itself within the host, avoiding phagocytosis [19]. A study completed by O'Connor, Darip [140] measured the protective antibody production directed at SCPA in adults and concluded that over 70% of patients contained a considerable amount of IgA. Whereas only 10% of children were observed to contain anti-SCPA antibodies [141].

8.4 Secreted antigens

8.4.1 Streptococcal Esterase (SSE)

GAS produces virulent enzymes that plays a role in the pathogenesis of the bacteria. A subset of these enzymes are hydrolases, that has the capacity for the degradation of nucleic acids and other host proteins [20]. Hayano and Tanaka [142] discovered an esterase in GAS supernatant containing two antigenic variants, however they did not fully characterize the secreted protein. In a more recent study completed by Liu, Zhu [129], identified the secreted esterase from a GAS isolate and referred to the protein as SSE. They went on to perform protective immunity assays by immunizing mice with the *emm*1-SSE peptide and resulted in the protection of mice subjected to a GAS skin infection. In a study completed by Zhu, Liu [143], went on to describe the function of SSE and its role in GAS virulence. They discovered that the secreted antigen is also in fact regulated by the CovR/S system. They also managed to knockout the *Spy1718* gene in a GAS isolate and subjected mice to subcutaneous GAS infections. The mice containing the knockout GAS isolate, avoided tissue invasion and GAS dissemination into deeper tissues.

8.4.2 Streptolysin O (SLO)

The formidable cytolytic toxins, streptolysin O (SLO) and streptolysin S (SLS) fall within the arsenal of virulent factors contributing to GAS pathogenesis. Müller-Alouf, Geoffroy [144,] Nizet, Beall [145] discovered the conservation and expression by almost all serotypes of GAS respectively. SLO is a toxin that attaches to cholesterol located in the cell membranes of eukaryotic cells, resulting in pore-formation, leading to cell lysis [146]. SLS on the other hand, is not so effective as it loses some of its toxicity upon separation with a carrier molecule, thus resulting in a nonimmunogenic peptide within the host [147]. The streptolysin toxins forms part of the cholesterol-dependent cytolysins also referred to as CDCs [148]; where SLO has been identified to disrupt GAS from being internalized into lysosomes, in which they could be destroyed, ultimately extending GAS survival [149]. A study completed by Johnson, Kurlan [130], took a longitudinal retrospective study design approach to determine the antibody responses from individuals presenting their primary clinician with sore throat. ELISA assays were performed in which recombinant peptides of both SLO and DNase B were prepared. They provided evidence that both peptides evoke the production of their respective anti-SLO and anti-DNase B antibodies capable of inducing bactericidal activity towards GAS. They also identified differences in the production of antibodies based on the presence of GAS; whether it being in its active state (true symptomatic infection) or latent state (carrier infection).

8.4.3 DNase B

It has been hypothesized that secreted DNase B has the ability to avoid the neutrophil extracellular traps (NETs) mediated-killing, contributing to the virulence of invasive GAS isolates. NETs are made up of a combination of neutrophil granules, DNA and chromatin which operates solitarily by capturing surrounding bacteria, independent from phagocytotic killing. DNases were first reported by McCarty [150] and well characterized thereafter by Wannamaker [151]. DNases are considered to be putative vaccine candidates for a number of reasons, 1) almost all GAS isolates are able to construct at least one DNase [151], 2) DNase activity is upregulated upon contact with host tissues and cells [152] and 3) McCarty [150] determined that DNases evoked anti-DNase antibodies from patients having GAS infections, indicating immunogenic potential of the peptide. The latter, has been used as a guideline for determining a true GAS infection, due to the increase in anti-DNase antibodies within the patient post infection [153]. As for their role played in virulence of GAS, they specifically target DNA molecules, crucial to the integrity of any functional cell. This is congruent with other gram positive bacterial studies such as the one completed by Chesbro and Walker [154], indicating the virulence potential of various nucleases. It has been reported that there are four different types of DNases, named DNase A through to D, having differences in their optimum pH, thermal stability, susceptibility and their substrate specificity, whether they are able to cleave DNA only or DNA and RNA [155]. However, it was observed that when measuring the various levels of DNases from >100 heterologous strains, DNase B was found to be produced by all. As reported above by Johnson, Kurlan [130], an increase in anti-DNase B antibodies were evident as

a true infection of GAS within patient sera presenting with sore throat. When GAS was in the latent stage of infection, antibodies towards DNase B plateaued with slight decreases in titres.

Rationale for this research

Introduction

Strategies to reduce the risk of GAS infections are currently focused on the use of antibiotics, specifically penicillin. Penicillin prophylaxis is used to combat GAS pharyngitis, preventing the risk of ARF [156]. Secondary prophylaxis is comprised of monthly benzathine penicillin G injections to reduce the risk of RHD. These techniques suffer greatly in resource-poor communities due to the overall efficacy of clinical infrastructures [121]. Thus, many believe that the long-term prevention mechanism for infections caused by GAS and its sequelae would be the development of a safe, effective, and affordable vaccine.

A meta-analysis completed by Robertson, Volmink [120], described the effectiveness of penicillin in preventing ARF. The study included 10 trials that compared the effectiveness of antibiotics and no antibiotics. Trials that described the effectiveness of penicillin showed a protective effect of 80% for reducing the incidence of ARF. Penicillin was also reported to have minimal side effects however, there were some concerns regarding the effects of administering intramuscular injections in developing countries. These concerns entailed the reuse of needles due to negligence, perceived higher risk of anaphylaxis and the actual pain of the injection itself [157].

Furthermore, failure of penicillin treatment is also due to adherence and limited supply. A study completed by Pelajo, Lopez-Benitez [158] aimed to quantify the adherence to secondary prophylaxis in a Brazilian cohort of children and adolescents with rheumatic fever (RF). They have shown that in patients with recurrent RF episodes, only 14.5% of patients reported adherence to secondary prophylaxis. They also reported a loss to follow-up of 33.5% of patients. Failure to adhere to penicillin is repeated across developing countries [159-161]. A study conducted in 2017 reported on the shortages of benzathine penicillin G, which were afforded due to shortfalls in supply, demand and procurement in the countries assessed [162]. Therefore, strong efforts have been made for an affordable and safe GAS vaccine that would reduce the reliance on penicillin whilst also preventing GAS post-sequelae diseases.

Current vaccine efforts

There is currently no licensed GAS vaccine to date. However, there is a 30-valent M peptide GAS formulation that successfully completed the phase I trial [163]. There are generally two approaches for GAS vaccine development, the first involves the type-specific M peptide approach as seen with the 30-valent M peptide vaccine [35] and the second approach involves identifying conserved GAS antigens, present on all GAS serotypes which produces bactericidal antibodies [164]. Previous research provides

evidence that the M protein elicits the greatest level of protective antibodies in both animal studies [12, 165] and recently, humans [163].

The progress of M peptide-based vaccines are hindered due to the amount of unprotected *emm* types, as there are >230 heterologous *emm* types. The vaccine is also affected by *emm* type prevalence variations amongst different geographical regions, with heterologous *emm* types occurring in populations at high-risk for ARF and RHD [36, 80, 166-168]. Furthermore, the dearth of data on *emm* type distribution, especially in high-risk, low-resource settings, adds to the slow progress in vaccine development. Additionally, M peptide-based vaccines were previously at risk due to the M protein molecular mimicry, with epitopes similar to human tissues, which may have resulted in the autoimmune disease the vaccine was intended to prevent [61]. However, the epitope mimicry problem was resolved by reducing the M peptide epitope to the smallest region (25-50 amino acids) that still provided a bactericidal effect and less likely to cause cross-reactivity with human tissues [169]. An alternative approach to extending coverage to all serotypes of GAS, may be the supplementation of GAS-specific antigens that could provide the necessary bactericidal antibodies; the efficacy of these GAS antigens could only be assessed by evaluating the human immune response.

Human immune response

Previous studies regarding the immune response to GAS antigens involved the detection of protective antibodies evoked by the M protein of various serotypes of GAS. It has also been shown that the presence of specific antibodies for a particular *emm* type would not provide protective efficacy against a heterologous *emm* type [12]. Further studies aimed to measure the bactericidal activity of antibodies evoked by common antigens of GAS; SLO, DNase B or C5a peptidase following a GAS infection [164, 170, 171]. Most efforts relied on an insensitive bactericidal assay to quantitatively inform on the abundance of antibodies.

Due to the complexity of GAS infections, variation of predominant *emm* types and the sheer number of heterologous *emm* types, more interest has been shown in identifying common GAS antigens that will provide additional bactericidal protection. Research completed on the included GAS antigens have identified key virulent determinants and protective antibodies in human sera through genomics, proteomics and reverse vaccinology approaches. *Details on the common GAS antigens were provided in Chapter 1 (Background).*

Efforts to date

The latest prospective study that assessed the human immune response to GAS antigens was completed with participants at low-risk for ARF/RHD (Hysmith, 2017). One-hundred and ninety-five serum samples were collected longitudinally during a 24-month period from 41 children (6-14 years of age) within the USA. The antigen specificity and kinetics of antibody responses were determined by ELISA assays against 18 serotype-specific M peptides and 13 common GAS antigens. The study retrieved a total of 12 heterologous *emm* types depicting 51 new GAS acquisitions. Immune responses to the common GAS antigens were highly variable, with a mean of 3.5 common antigens. Antibody responses to homologous M peptides were observed in 63% (n=32 episodes) of new GAS acquisitions. Asymptomatic cases produced responses with a mean of 3.7 for both, common antigens and homologous M peptides. Whereas symptomatic acquisitions were associated with antibody responses to a mean of 3.0 antigens. Following antibiotic treatment in 31% of new GAS acquisitions, antibody responses were to a mean of 3.1 antigens. If no antibiotics were prescribed (69%), immune responses were increased to a mean of 3.6 antigens. Seven participants were infected with more than one *emm* type of GAS. There were no new GAS acquisitions in participants that had pre-existing antibodies against the homologous M peptide. Sixty-five percent of new GAS acquisitions were asymptomatic yet showed a significant immunological response to the common antigens; streptolysin O and/or DNase B. The clinical protocol, microbiological and immunological assessments in the pilot study will act as a model for the proposed thesis, providing comparable results to determine the efficacy and reliability of the overall analysis.

Motivation for undertaking this research

The *emm* types selected for the 30-valent M protein-based vaccine was based on epidemiological data acquired from the Western world; USA, Canada and Western Europe [35]. Even though the vaccine offers the potential for cross-protection to *emm* types not initially included in the vaccine, there are still significant gaps in *emm* type coverage. This study therefore aims to provide data on the prevalent *emm* types recovered in areas where ARF/RHD is rampant, providing insight that could assist in mitigating the gaps within the current recombinant 30-valent vaccine formulation.

There are limited data regarding the human immune response to GAS antigens. Therefore, this thesis aims to identify the potential efficacy of common GAS antigens in evoking antibodies that can be supplemented into the vaccine with the goal of extending coverage. This arm of the study will also provide insight into the underlying effects occurring within the immune system with regards to the state of GAS infection, be it symptomatic or asymptomatic (GAS carrier state). In terms of the latter, could confirm or refute previous theories of an ongoing inflammatory response with continuous harmful effects irrespective of symptoms.

Aims

To determine the available epidemiological data of GAS reported in Africa

To determine the prevalent *emm* types circulating in a Cape Town cohort

To determine the protective coverage of the 30-valent based vaccine in Africa

To determine the association of GAS antigens with ARF for their usefulness in early diagnosis

To determine the host immune response to GAS antigens following the acquisition of GAS.

Objectives of the research

Study I (Chapter 2) is a systematic review completed on the epidemiology of GAS strains isolated across Africa, a region where the burden of RHD is highest. The review assists in reporting the predominant *emm* types in Africa. The *emm* type prevalence is then tested to identify which *emm* types would be covered in the 30-valent vaccine and thus, report the protective coverage the vaccine would provide. This review will provide further epidemiology of the region and whether the vaccine will provide effective coverage to all strains of GAS regardless of *emm* type. Study I was successfully accepted into the Msphere Journal and reported as a complete paper.

Study II (Chapter 3) is designed to prospectively identify the molecular epidemiology of new GAS acquisitions within a Cape Town cohort. Serial throat cultures are obtained at 2-monthly intervals and at interim times if the participant reported symptoms of pharyngitis. Subsequent *emm* typing of GAS will differentiate the acquisition of a new *emm* type versus persistent carriage of the same type over time. We will determine the frequency of acquisitions, the percentage of total participants that acquire new *emm* types, the *emm* type responsible, and the duration of persistence of GAS in the pharynx. All of this information will provide a comprehensive description of the epidemiology of GAS in a population at high-risk for ARF/RHD.

Study III (Chapter 4) is a systematic review aimed at synthesizing existing data in identifying how the immune response to GAS antigens may be correlated to the early diagnosis of ARF. The Jones Criteria states that the diagnosis of a bona fide case of ARF requires evidence of a preceding GAS infection. A positive throat culture for GAS and antibodies against SLO and DNase B could be used to clearly describe that the individual had a preceding GAS infection prior to ARF symptoms. Therefore, we sought to provide, from published studies, an evidence-based synthesis of the correlation of streptococcal serology to establish the usefulness of immunological data in aiding the diagnosis of ARF. Study III was successfully accepted into the Frontiers Journal and reported as a complete paper.

Study IV (Chapter 5) is designed to assess the serological immune status to GAS antigens of all enrolled participants within a Cape Town cohort over a 24-month observation period. Serum samples will be collected prospectively every 2 months and at interim times if participants experience symptoms of sore throat. Antibody levels from sequential serum samples will be assayed by ELISA against an extensive panel of GAS antigens and will be used in several analyses to answer key questions related to the host-pathogen interactions following natural acquisition of GAS.

Appendices [must be attached at the end of document]: The following three documents have been appended to the end of the thesis:

Appendix A: Ethical approval HREC 461/2018

Appendix B: Supplementary material for Chapter 2: Study I

Appendix C: Supplementary material for Chapter 3: Study II

Appendix D: Supplementary material for Chapter 4: Study III

Appendix D: Supplementary material for Chapter 5: Study IV

Author contributions to included manuscripts

Chapter 2: M.E.E., T.S., and A.M. wrote the protocol and designed the study. T.S. implemented and managed the study. T.S. and K.E. reviewed articles and recorded the data. B.M. was the third arbitrator. T.S. and M.E.E. conducted statistical analyses. T.S., J.B.D., and M.E.E. interpreted the data and wrote the manuscript. All authors read and approved the final manuscript.

Chapter 4: T.S., B.M., and M.E.E were jointly responsible for the conceptualization of the study. T.S., K.R., B.M., and K.E. contributed to the search strategy and performed data extraction. T.S. and M.E.E. designed and executed the analyses, wrote the first draft, and revised drafts of the manuscript. J.B.D. and L.J.Z. contributed to the interpretation of the findings. All authors have read and approved the final manuscript.

Chapter 5: M.E.E. and J.B.D. wrote the protocol and designed the study. M.E.E. implemented and managed the study together with K.E. and R.D. T.S. managed the laboratory at UCT, assisted by K.R. and B.M. assisted by T.P. at UTHSC. C.S. and L.Z. contributed to the clinical aspects of this work. T.S., B.M., M.E.E. and J.B.D. conducted statistical analyses. T.S., B.M., M.E.E. and J.B.D. interpreted the data and wrote the manuscript. All authors read and approved the final manuscript.

References

1. Carapetis JR. Rheumatic heart disease in developing countries. J New England Journal of Medicine 2007;**357**(5):439-441.
2. Tart AH, Walker MJ, Musser JM. New understanding of the group A Streptococcus pathogenesis cycle. Trends Microbiol 2007;**15**(7):318-25.
3. Fischetti VA. M protein and other surface proteins on streptococci. 2016.
4. Sanderson-Smith M, De Oliveira DM, Guglielmini J, McMillan DJ, Vu T, Holien JK, Henningham A, Steer AC, Bessen DE, Dale JB. A systematic and functional classification of Streptococcus pyogenes that serves as a new tool for molecular typing and vaccine development. The Journal of infectious diseases 2014;**210**(8):1325-1338.
5. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine 2009;**151**(4):264-269.
6. Callegari E, A, Chaussee M, S. The Streptococcal Proteome--Streptococcus pyogenes: Basic Biology to Clinical Manifestations. 2016.
7. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, Sriprakash KS, Sanderson-Smith ML, Nizet V. Disease manifestations and pathogenic mechanisms of Group A Streptococcus. Clin Microbiol Rev 2014;**27**(2):264-301.
8. Billroth T. *Lectures on surgical pathology and therapeutics: a handbook for students and practitioners*: New Sydenham Society; 1878.
9. Alouf JE, Horaud T. Streptococcal research at Pasteur Institute from Louis Pasteur's time to date. In. *Streptococci and the Host*: Springer; 1997, 7-14.
10. Evans AC. Studies on hemolytic streptococci: II. Streptococcus pyogenes. Journal of bacteriology 1936;**31**(6):611.
11. Köhler W. The present state of species within the genera Streptococcus and Enterococcus. International journal of medical microbiology: IJMM 2007;**297**(3):133.
12. Lancefield RC. Current knowledge of type-specific M antigens of group A streptococci. The Journal of Immunology 1962;**89**(3):307-313.
13. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. The Lancet infectious diseases 2009;**9**(10):611-616.
14. Simpson WA, Beachey EH. Adherence of group A streptococci to fibronectin on oral epithelial cells. Infection and immunity 1983;**39**(1):275-279.
15. Courtney HS, Hasty DL, Dale JB, Poirier TP. A 28-kilodalton fibronectin-binding protein of group A streptococci. Current microbiology 1992;**25**(5):245-250.
16. Terao Y, Kawabata S, Kunitomo E, Murakami J, Nakagawa I, Hamada S. Fba, a novel fibronectin-binding protein from Streptococcus pyogenes, promotes bacterial entry into epithelial cells, and the fba gene is positively transcribed under the Mga regulator. Molecular microbiology 2001;**42**(1):75-86.
17. Wessels MR, Bronze MS. Critical role of the group A streptococcal capsule in pharyngeal colonization and infection in mice. Proceedings of the National Academy of Sciences 1994;**91**(25):12238-12242.
18. Herman A, Kappler JW, Marrack P, Pullen AM. Superantigens: mechanism of T-cell stimulation and role in immune responses. Annual review of immunology 1991;**9**(1):745-772.
19. Ji Y, McLandsborough L, Kondagunta A, Cleary PP. C5a peptidase alters clearance and trafficking of group A streptococci by infected mice. Infection and immunity 1996;**64**(2):503-510.
20. Sumby P, Barbian KD, Gardner DJ, Whitney AR, Welty DM, Long RD, Bailey JR, Parnell MJ, Hoe NP, Adams GG. Extracellular deoxyribonuclease made by group A Streptococcus assists pathogenesis by enhancing evasion of the innate immune response. Proceedings of the National Academy of Sciences 2005;**102**(5):1679-1684.

21. Perez-Casal J, Caparon M, Scott J. Introduction of the emm6 gene into an emm-deleted strain of Streptococcus pyogenes restores its ability to resist phagocytosis. Research in microbiology 1992;**143**(6):549-558.
22. Fischetti VA. Streptococcal M protein: molecular design and biological behavior. Clinical microbiology reviews 1989;**2**(3):285-314.
23. Dale J, Seyer J, Beachey E. Type-specific immunogenicity of a chemically synthesized peptide fragment of type 5 streptococcal M protein. Journal of Experimental Medicine 1983;**158**(5):1727-1732.
24. Johnson DR, Kaplan EL, Bicova R, Havlicek J, Havlickova H, Kritz P, Motlova J, Sramek J, Organization WH. Laboratory diagnosis of group A streptococcal infections. 1996.
25. Mora M, Bensi G, Capo S, Falugi F, Zingaretti C, Manetti AG, Maggi T, Taddei AR, Grandi G, Telford JL. Group A Streptococcus produce pilus-like structures containing protective antigens and Lancefield T antigens. Proceedings of the National Academy of Sciences 2005;**102**(43):15641-15646.
26. Falugi F, Zingaretti C, Pinto V, Mariani M, Amodeo L, Manetti AG, Capo S, Musser JM, Orefici G, Margarit I. Sequence variation in group A Streptococcus pili and association of pilus backbone types with lancefield T serotypes. The Journal of infectious diseases 2008;**198**(12):1834-1841.
27. Beall B, Facklam R, Thompson T. Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. Journal of clinical microbiology 1996;**34**(4):953-958.
28. Neal S, Beall B, Ekelund K, Henriques-Normark B, Jasir A, Johnson D, Kaplan E, Lovgren M, Reinert RR, Efstratiou A. International quality assurance study for characterization of Streptococcus pyogenes. Journal of clinical microbiology 2007;**45**(4):1175-1179.
29. Facklam R, Beall B, Efstratiou A, Fischetti V, Johnson D, Kaplan E, Kriz P, Lovgren M, Martin D, Schwartz B. emm typing and validation of provisional M types for group A streptococci. Emerging infectious diseases 1999;**5**(2):247.
30. CDC. *Protocol for emm typing*. https://www.cdc.gov/streplab/groupa-strep/emm-typing-protocol.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fstreplab%2Fprotocol-emm-type.html (29/04/2020).
31. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones ofStaphylococcus aureus. Journal of clinical microbiology 2000;**38**(3):1008-1015.
32. Enright MC, Spratt BG, Kalia A, Cross JH, Bessen DE. Multilocus sequence typing of Streptococcus pyogenes and the relationships between emm type and clone. Infection and immunity 2001;**69**(4):2416-2427.
33. Efstratiou A, Lamagni T. Epidemiology of Streptococcus pyogenes--Streptococcus pyogenes: Basic Biology to Clinical Manifestations. 2016.
34. Abdissa A. Characterization of Group A Streptococci Isolated from Throat of Healthy School Children in Ethiopia. Addis Ababa University 2005.
35. Dale JB, Penfound TA, Tamboura B, Sow SO, Nataro JP, Tapia M, Kotloff KL. Potential coverage of a multivalent M protein-based group A streptococcal vaccine. Vaccine 2013;**31**(12):1576-81.
36. Muhamed B. The molecular characterisation of Group A Streptococcus among children with phyryngitis in the Vanguard community (Bonteheuwel). University of Cape Town 2012.
37. Seale AC, Davies MR, Anampiu K, Morpeth SC, Nyongesa S, Mwarumba S, Smeesters PR, Efstratiou A, Karugutu R, Mturi N. Invasive group A streptococcus infection among children, rural Kenya. Emerging infectious diseases 2016;**22**(2):224.
38. Chiang-Ni C, Zheng PX, Wang SY, Tsai PJ, Chuang WJ, Lin YS, Liu CC, Wu JJ. Epidemiology Analysis of Streptococcus pyogenes in a Hospital in Southern Taiwan by Use of the Updated emm Cluster Typing System. J Clin Microbiol 2016;**54**(1):157-62.
39. Chiou CS, Liao TL, Wang TH, Chang HL, Liao JC, Li CC. Epidemiology and molecular characterization of Streptococcus pyogenes recovered from scarlet fever patients in central Taiwan from 1996 to 1999. J Clin Microbiol 2004;**42**(9):3998-4006.

40. Jose JJM, Brahmadathan KN, Abraham VJ, Huang CY, Morens D, Hoe NP, Follmann DA, Krause RM. Streptococcal group A, C and G pharyngitis in school children: a prospective cohort study in Southern India. Epidemiol Infect 2018;**146**(7):848-853.
41. Koh E, Kim S. Decline in erythromycin resistance in group A Streptococci from acute pharyngitis due to changes in the emm Genotypes rather than restriction of antibiotic use. Korean J Lab Med 2010;**30**(5):485-90.
42. Ogawa T, Terao Y, Okuni H, Ninomiya K, Sakata H, Ikebe K, Maeda Y, Kawabata S. Biofilm formation or internalization into epithelial cells enable Streptococcus pyogenes to evade antibiotic eradication in patients with pharyngitis. Microb Pathog 2011;**51**(1-2):58-68.
43. Baroux N, D'Ortenzio E, Amedeo N, Baker C, Ali Alsuwayyid B, Dupont-Rouzeyrol M, O'Connor O, Steer A, Smeesters PR. The emm-cluster typing system for Group A Streptococcus identifies epidemiologic similarities across the Pacific region. Clin Infect Dis 2014;**59**(7):e84-92.
44. McDonald MI, Towers RJ, Fagan P, Carapetis JR, Currie BJ. Molecular typing of Streptococcus pyogenes from remote Aboriginal communities where rheumatic fever is common and pyoderma is the predominant streptococcal infection. Epidemiol Infect 2007;**135**(8):1398-405.
45. Richardson LJ, Towers RJ, Cheng AC, Currie BJ, Carapetis JR, Giffard PM, McDonald MI. Diversity of emm sequence types in group A beta-haemolytic streptococci in two remote Northern Territory Indigenous communities: implications for vaccine development. Vaccine 2010;**28**(32):5301-5.
46. Mhlanga N, Sharp G, Nulsen M. Emm type distribution of group A streptococcus isolates from the throat swabs of children living in areas with a high (Northland and Gisborne) or low (Palmerston North) incidence of acute rheumatic fever. The New Zealand medical journal 2017;**130**(1460):13-20.
47. Williamson DA, Morgan J, Hope V, Fraser JD, Moreland NJ, Proft T, Mackereth G, Lennon D, Baker MG, Carter PE. Increasing incidence of invasive group A streptococcus disease in New Zealand, 2002-2012: a national population-based study. J Infect 2015;**70**(2):127-34.
48. Shulman ST, Tanz RR, Dale JB, Beall B, Kabat W, Kabat K, Cederlund E, Patel D, Rippe J, Li Z. Seven-year surveillance of North American pediatric group A streptococcal pharyngitis isolates. Clinical infectious diseases 2009;**49**(1):78-84.
49. Nelson GE, Pondo T, Toews K-A, Farley MM, Lindegren ML, Lynfield R, Aragon D, Zansky SM, Watt JP, Cieslak PR. Epidemiology of invasive group A streptococcal infections in the United States, 2005–2012. Reviews of Infectious Diseases 2016;**63**(4):478-486.
50. Rudolph K, Bruce MG, Bruden D, Zulz T, Reasonover A, Hurlburt D, Hennessy T. Epidemiology of Invasive Group A Streptococcal Disease in Alaska, 2001 to 2013. J Clin Microbiol 2016;**54**(1):134-41.
51. Rubio-López V, Valdezate S, Álvarez D, Villalón P, Medina MJ, Salcedo C, Sáez-Nieto J-A. Molecular epidemiology, antimicrobial susceptibilities and resistance mechanisms of Streptococcus pyogenes isolates resistant to erythromycin and tetracycline in Spain (1994–2006). BMC microbiology 2012;**12**(1):215.
52. Plainvert C, Dinis M, Ravins M, Hanski E, Touak G, Dmytruk N, Fouet A, Poyart C. Molecular epidemiology of sil locus in clinical Streptococcus pyogenes strains. J Clin Microbiol 2014;**52**(6):2003-10.
53. Koutouzi F, Tsakris A, Chatzichristou P, Koutouzis E, Daikos GL, Kirikou E, Petropoulou N, Syriopoulou V, Michos A. Streptococcus pyogenes emm Types and Clusters during a 7-Year Period (2007 to 2013) in Pharyngeal and Nonpharyngeal Pediatric Isolates. J Clin Microbiol 2015;**53**(7):2015-21.
54. Imohl M, Fitzner C, Perniciaro S, van der Linden M. Epidemiology and distribution of 10 superantigens among invasive Streptococcus pyogenes disease in Germany from 2009 to 2014. PLoS One 2017;**12**(7):e0180757.
55. Pastural E, McNeil SA, MacKinnon-Cameron D, Ye L, Langley JM, Stewart R, Martin LH, Hurley GJ, Salehi S, Penfound TA, Halperin S, Dale JB. Safety and immunogenicity of a 30-valent M protein-

based group a streptococcal vaccine in healthy adult volunteers: A randomized, controlled phase I study. Vaccine 2020;**38**(6):1384-1392.
56. Metzgar D, Zampolli A. The M protein of group A Streptococcus is a key virulence factor and a clinically relevant strain identification marker. Virulence 2011;**2**(5):402-412.
57. Turner CE, Abbott J, Lamagni T, Holden MT, David S, Jones MD, Game L, Efstratiou A, Sriskandan S. Emergence of a New Highly Successful Acapsular Group A Streptococcus Clade of Genotype emm89 in the United Kingdom. MBio 2015;**6**(4):e00622.
58. Shulman ST, Stollerman G, Beall B, Dale JB, Tanz RR. Temporal changes in streptococcal M protein types and the near-disappearance of acute rheumatic fever in the United States. Clinical infectious diseases 2006;**42**(4):441-447.
59. Anthony BF, Kaplan EL, Wannamaker LW, Chapman SS. The dynamics of streptococcal infections in a defined population of children: serotypes associated with skin and respiratory infections. American Journal of Epidemiology 1976;**104**(6):652-666.
60. Bessen DE, Sotir CM, Readdy TL, Hollingshead SK. Genetic correlates of throat and skin isolates of group A streptococci. Journal of Infectious Diseases 1996;**173**(4):896-900.
61. Cunningham MW. Pathogenesis of group A streptococcal infections. Clinical microbiology reviews 2000;**13**(3):470-511.
62. Bessen DE, Carapetis JR, Beall B, Katz R, Hibble M, Currie BJ, Collingridge T, Izzo MW, Scaramuzzino DA, Sriprakash KS. Contrasting molecular epidemiology of group A streptococci causing tropical and nontropical infections of the skin and throat. The Journal of infectious diseases 2000;**182**(4):1109-1116.
63. Carapetis JR, Currie BJ, Kaplan EL. Epidemiology and prevention of group A streptococcal infections: acute respiratory tract infections, skin infections, and their sequelae at the close of the twentieth century. Clinical Infectious Diseases 1999;**28**(2):205-210.
64. Lin JN, Chang LL, Lai CH, Lin HH, Chen YH. Emergence of Streptococcus pyogenes emm102 causing toxic shock syndrome in Southern Taiwan during 2005-2012. PLoS One 2013;**8**(12):e81700.
65. Lacy MD, Horn K. Nosocomial transmission of invasive group a streptococcus from patient to health care worker. Clinical infectious diseases 2009;**49**(3):354-357.
66. McGregor KF, Spratt BG, Kalia A, Bennett A, Bilek N, Beall B, Bessen DE. Multilocus sequence typing of Streptococcus pyogenes representing most known emm types and distinctions among subpopulation genetic structures. Journal of bacteriology 2004;**186**(13):4285-4294.
67. Bowen A, Harris T, Holt D, Giffard P, Carapetis J, Campbell P, McVernon J, Tong S. Whole genome sequencing reveals extensive community-level transmission of group A Streptococcus in remote communities. Epidemiology & Infection 2016;**144**(9):1991-1998.
68. Dale JB, Fischetti VA, Carapetis J, Steer A, Sow SO, Kumar R. Group A Streptococcal vaccines: Paving a path for accellerated development. Vaccine 2013;**31**:216-22.
69. Hu MC, Stroop SD, Reddish MA, Beall B, Dale JB. Immunogenicity of a 26-Valent Group A Streptococcal Vaccine. Infection and Immunity 2002;**70**(4):2171-2177.
70. Dale JB, Penfound TA, Chiang EY, Walton WJ. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine 2011;**29**:8175-8.
71. Jackson SJ, Steer AC, Campbell H. Systematic Review: Estimation of global burden of non-suppurative sequelae of upper respiratory tract infection: rheumatic fever and post-streptococcal glomerulonephritis. Tropical Medicine 2011;**16**(1):2-11.
72. Wessels MR. Streptococcal pharyngitis. New England Journal of Medicine 2011;**364**(7):648-655.
73. Bennett JE, Dolin R, Blaser MJ. *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases E-Book*: Elsevier Health Sciences; 2014.
74. Armstrong GL, Pinner RW. Outpatient visits for infectious diseases in the United States, 1980 through 1996. Archives of internal medicine 1999;**159**(21):2531-2536.

75. Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. The Lancet infectious diseases 2005;**5**(11):685-694.
76. Bisno AL. Acute pharyngitis. New England Journal of Medicine 2001;**344**(3):205-211.
77. Bisno AL. Acute pharyngitis: etiology and diagnosis. Pediatrics 1996;**97**(6):949-954.
78. Engel ME, Muhamed B, Whitelaw AC, Musvosvi M, Mayosi BM, Dale JB. Group A streptococcal emm type prevalence among symptomatic children in Cape Town and potential vaccine coverage. The Pediatric infectious disease journal 2014;**33**(2):208.
79. Danchin MH, Rogers S, Kelpie L, Selvaraj G, Curtis N, Carlin JB, Nolan TM, Carapetis JR. Burden of acute sore throat and group A streptococcal pharyngitis in school-aged children and their families in Australia. Pediatrics 2007;**120**(5):950-957.
80. Tapia MD, Sow SO, Tamboura B, Keita MM, Berthe A, Samake M, Nataro JP, Onwuchekwa UO, Penfound TA, Blackwelder W, Dale JB, Kotloff KL. Streptococcal pharyngitis in schoolchildren in Bamako, Mali. Pediatr Infect Dis J 2015;**34**(5):463-8.
81. Ellis NM, Li Y, Hildebrand W, Fischetti VA, Cunningham MW. T cell mimicry and epitope specificity of cross-reactive T cell clones from rheumatic heart disease. The Journal of Immunology 2005;**175**(8):5448-5456.
82. Faé KC, da Silva DD, Oshiro SE, Tanaka AC, Pomerantzeff PM, Douay C, Charron D, Toubert A, Cunningham MW, Kalil J. Mimicry in recognition of cardiac myosin peptides by heart-intralesional T cell clones from rheumatic heart disease. The Journal of Immunology 2006;**176**(9):5662-5670.
83. Gewitz MH, Baltimore RS, Tani LY, Sable CA, Shulman ST, Carapetis J, Remenyi B, Taubert KA, Bolger AF, Beerman L. Revision of the Jones Criteria for the diagnosis of acute rheumatic fever in the era of Doppler echocardiography: a scientific statement from the American Heart Association. Circulation 2015;**131**(20):1806-1818.
84. Bhardwaj R, Sood A. Clinical profile of acute rheumatic fever patients in a tertiary care institute in present era. Journal of The Association of Physicians of India 2015;**63**.
85. Organization WH. *Rheumatic fever and rheumatic heart disease: report of a WHO Expert Consultation, Geneva, 29 October-1 November, 2001*: World Health Organization; 2004.
86. Zühlke LJ, Beaton A, Engel ME, Hugo-Hamman CT, Karthikeyan G, Katzenellenbogen JM, Ntusi N, Ralph AP, Saxena A, Smeesters PR. Group A streptococcus, acute rheumatic fever and rheumatic heart disease: epidemiology and clinical considerations. Current treatment options in cardiovascular medicine 2017;**19**(2):15.
87. Baumgartner H, Hung J, Bermejo J, Chambers JB, Evangelista A, Griffin BP, Iung B, Otto CM, Pellikka PA, Quiñones M. Echocardiographic assessment of valve stenosis: EAE/ASE recommendations for clinical practice. Journal of the American Society of Echocardiography 2009;**22**(1):1-23.
88. Zuhlke LJ, Engel ME, Watkins D, Mayosi BM. Incidence, prevalence and outcome of rheumatic heart disease in South Africa: a systematic review of contemporary studies. Int J Cardiol 2015;**199**:375-83.
89. Lawrence JG, Carapetis JR, Griffiths K, Edwards K, Condon JR. Acute rheumatic fever and rheumatic heart disease: incidence and progression in the Northern Territory of Australia, 1997 to 2010. Circulation 2013;**128**(5):492-501.
90. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton AZ, Benjamin EJ, Benziger CP. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. Journal of the American College of Cardiology 2020;**76**(25):2982-3021.
91. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, Charlson F, Davis A, Degenhardt L, Dicker D. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet 2015;**386**(9995):743-800.

92. Steer AC, Kado J, Wilson N, Tuiketei T, Batzloff M, Waqatakirewa L, Mulholland EK, Carapetis JR. High prevalence of rheumatic heart disease by clinical and echocardiographic screening among children in Fiji. J Heart Valve Dis 2009;**18**(3):327-35.
93. Pastore S, De Cunto A, Benettoni A, Berton E, Taddio A, Lepore L. The resurgence of rheumatic fever in a developed country area: the role of echocardiography. Rheumatology 2010;**50**(2):396-400.
94. Nkgudi B, Robertson KA, Volmink J, Mayosi BM. Notification of rheumatic fever in South Africa-evidence for underreporting by health care professionals and administrators. South African medical journal 2006;**96**(3):206-208.
95. Singh PI, Carapetis JR, Buadromo EM, Samberkar PN, Steer AC. The high burden of rheumatic heart disease found on autopsy in Fiji. Cardiology in the Young 2008;**18**(1):62-69.
96. Okello E, Wanzhu Z, Musoke C, Kakande B, Mondo CK, Freers J, Twalib A, Lwabi P, Wilson NB, Odoi-Adome R. Cardiovascular complications in newly diagnosed rheumatic heart disease patients at Mulago Hospital, Uganda. Cardiovascular journal of Africa 2013;**24**(3):82.
97. Jack SJ, Williamson DA, Galloway Y, Pierse N, Zhang J, Oliver J, Milne RJ, Mackereth G, Jackson CM, Steer AC. Primary prevention of rheumatic fever in the 21st century: evaluation of a national programme. International Journal of Epidemiology 2018;**47**(5):1585-1593.
98. Bergmark R, Bergmark B, Blander J, Fataki M, Janabi M. Burden of disease and barriers to the diagnosis and treatment of group a beta-hemolytic streptococcal pharyngitis for the prevention of rheumatic heart disease in Dar Es Salaam, Tanzania. The Pediatric infectious disease journal 2010;**29**(12):1135-1137.
99. Ahmed M, Vellani CW, Awiti AO. Medical education: meeting the challenge of implementing primary health care in sub-saharan Africa. Infectious Disease Clinics 2011;**25**(2):411-420.
100. Dajani AS, Ayoub E, Bierman FZ, Bisno AL, Denny FW, Durack D, Ferrieri P, Freed M, Gerber M, Kaplan E. Guidelines for the diagnosis of rheumatic fever: Jones criteria, updated 1992: special writing group of the committee on rheumatic fever, endocarditis, and Kawasaki disease of the council on cardiovascular disease in the young, American Heart Association. Circulation 1993;**87**(1).
101. Ferrieri P. Proceedings of the Jones Criteria workshop. Circulation 2002;**106**(19):2521-2523.
102. DeMuri GP, Wald ER. The Group A Streptococcal Carrier State Reviewed: Still an Enigma. J Pediatric Infect Dis Soc 2014;**3**(4):336-42.
103. Gerber MA, Randolph MF, Mayo DR. The group A streptococcal carrier state: A reexamination. American Journal of Diseases of Children 1988;**142**(5):562-565.
104. Tanz RR, Shulman ST, Barthel MJ, Willert C, Yogev R. Penicillin plus rifampin eradicates pharyngeal carriage of group A streptococci. The Journal of pediatrics 1985;**106**(6):876-880.
105. Pichichero ME, Casey JR. Defining and dealing with carriers of group A streptococci. Contemporary Pediatrics 2003;**20**(1):46-53.
106. Pichichero ME, Marsocci SM, Murphy AML, Hoeger W, Green JL, Sorrento A. Incidence of streptococcal carriers in private pediatric practice. Archives of pediatrics adolescent medicine 1999;**153**(6):624-628.
107. Martin JM, Green M, Barbadora KA, Wald ER. Group A streptococci among school-aged children: clinical characteristics and the carrier state. Pediatrics 2004;**114**(5):1212-9.
108. Kaplan EL, Johnson DR. Unexplained reduced microbiological efficacy of intramuscular benzathine penicillin G and of oral penicillin V in eradication of group A streptococci from children with acute pharyngitis. Pediatrics 2001;**108**(5):1180-1186.
109. Casey JR, Kahn R, Gmoser D, Atlas E, Urbani K, Luber S, Pellman H, Pichichero ME. Frequency of symptomatic relapses of group A β-hemolytic streptococcal tonsillopharyngitis in children from 4 pediatric practices following penicillin, amoxicillin, and cephalosporin antibiotic treatment. Clinical pediatrics 2008;**47**(6):549-554.
110. Bisno AL, Gerber MA, Gwaltney Jr JM, Kaplan EL, Schwartz RHJCID. Practice guidelines for the diagnosis and management of group A streptococcal pharyngitis. 2002:113-125.

111. Shulman ST, Bisno AL, Clegg W, Gerber M, Kaplan E, Lee G, Martin J, Van Beneden C. Clinical Practice Guideline for the Diagnosis and Management of Group A Streptococcal Pharyngitis: 2012 Update by the Infectious Diseases Society of America. Clin Infect Dis 2012;**55**(10):86-102.
112. Dajani A, Taubert K, Ferrieri P, Peter G, Shulman S, Pediatrics AHAJ. Treatment of acute streptococcal pharyngitis and prevention of rheumatic fever: a statement for health professionals. 1995;**96**(4):758-764.
113. Marijon E, Ou P, Celermajer DS, Ferreira B, Mocumbi AO, Jani D, Paquet C, Jacob S, Sidi D, Jouven X. Prevalence of rheumatic heart disease detected by echocardiographic screening. New England Journal of Medicine 2007;**357**(5):470-476.
114. Manyemba J, Mayosi BMJSAMJ. Intramuscular penicillin is more effective than oral penicillin in secondary prevention of rheumatic fever-a systematic review. 2003;**93**(3):212-218.
115. Cilliers A. Treating acute rheumatic fever. In: British Medical Journal Publishing Group; 2003.
116. Essop MR, Nkomo VT. Rheumatic and nonrheumatic valvular heart disease: epidemiology, management, and prevention in Africa. Circulation 2005;**112**(23):3584-3591.
117. Skoularigis J, Sinovich V, Joubert G, Sareli P. Evaluation of the long-term results of mitral valve repair in 254 young patients with rheumatic mitral regurgitation. Circulation 1994;**90**(5 Pt 2):II167-74.
118. Maguire GP, Carapetis JR, Walsh WF, Brown AJMJA. The future of acute rheumatic fever and rheumatic heart disease in Australia. 2012;**197**(3):133-4.
119. Wannamaker LW. The epidemology of streptococcal Infection. Streptococcal infection 1954.
120. Robertson KA, Volmink JA, Mayosi BMJBcd. Antibiotics for the primary prevention of acute rheumatic fever: a meta-analysis. 2005;**5**(1):11.
121. Remenyi B, Carapetis J, Wyber R, Taubert K, Mayosi BMJNRC. Position statement of the World Heart Federation on the prevention and control of rheumatic heart disease. 2013;**10**(5):284.
122. Courtney HS, Dale JB, Hasty D. Differential effects of the streptococcal fibronectin-binding protein, FBP54, on adhesion of group A streptococci to human buccal cells and HEp-2 tissue culture cells. Infection and immunity 1996;**64**(7):2415-2419.
123. Butler JE. Enzyme-linked immunosorbent assay. Journal of immunoassay 2000;**21**(2-3):165-209.
124. Voller A, Bartlett A, Bidwell D. Enzyme immunoassays with special reference to ELISA techniques. Journal of Clinical Pathology 1978;**31**(6):507-520.
125. Crowther JR. *The ELISA guidebook*: Springer Science & Business Media; 2000.
126. Dale JB, Niedermeyer SE, Agbaosi T, Hysmith ND, Penfound TA, Hohn CM, Pullen M, Bright MI, Murrell DS, Shenep LE, Courtney HS. Protective immunogenicity of group A streptococcal M-related proteins. Clin Vaccine Immunol 2015;**22**(3):344-50.
127. Gallotta M, Gancitano G, Pietrocola G, Mora M, Pezzicoli A, Tuscano G, Chiarot E, Nardi-Dei V, Taddei AR, Rindi S, Speziale P, Soriani M, Grandi G, Margarit I, Bensi G. SpyAD, a moonlighting protein of group A Streptococcus contributing to bacterial division and host cell adhesion. Infect Immun 2014;**82**(7):2890-901.
128. Cleary PP, Prahbu U, Dale J, Wexler D, Handley J. Streptococcal C5a peptidase is a highly specific endopeptidase. Infection and immunity 1992;**60**(12):5219-5223.
129. Liu M, Zhu H, Zhang J, Lei B. Active and passive immunizations with the streptococcal esterase Sse protect mice against subcutaneous infection with group A streptococci. Infect Immun 2007;**75**(7):3651-7.
130. Johnson DR, Kurlan R, Leckman J, Kaplan EL. The human immune response to streptococcal extracellular antigens: clinical, diagnostic, and potential pathogenetic implications. Clin Infect Dis 2010;**50**(4):481-90.
131. Dale J, Simmons M, Chiang EC, Chiang EY. Recombinant, octavalent group A streptococcal M protein vaccine. Vaccine 1996;**14**(10):944-948.
132. Dale J, Chiang EY, Liu S, Courtney HS, Hasty DL. New protective antigen of group A streptococci. The Journal of clinical investigation 1999;**103**(9):1261-1268.

133. Courtney HS, Li Y. Non-immune binding of human IgG to M-related proteins confers resistance to phagocytosis of group A streptococci in blood. PLoS One 2013;**8**(10):e78719.
134. Courtney HS, Hasty DL, Dale JB. Anti-phagocytic mechanisms of Streptococcus pyogenes: binding of fibrinogen to M-related protein. Molecular microbiology 2006;**59**(3):936-947.
135. De Bentzmann S, Plotkowski C, Puchelle E. Receptors in the Pseudomonas aeruginosa adherence to injured and repairing airway epithelium. American journal of respiratory and critical care medicine 1996;**154**(4):S155.
136. Bensi G, Mora M, Tuscano G, Biagini M, Chiarot E, Bombaci M, Capo S, Falugi F, Manetti AG, Donato P. Multi high-throughput approach for highly selective identification of vaccine candidates: the Group A Streptococcus case. Molecular & Cellular Proteomics 2012:mcp. M111. 015693.
137. Wexler DE, Chenoweth DE, Cleary PP. Mechanism of action of the group A streptococcal C5a inactivator. Proceedings of the National Academy of Sciences 1985;**82**(23):8144-8148.
138. Chen CC, Cleary P. Complete nucleotide sequence of the streptococcal C5a peptidase gene of Streptococcus pyogenes. Journal of Biological Chemistry 1990;**265**(6):3161-3167.
139. Fernandez HN, Hugli T. Primary structural analysis of the polypeptide portion of human C5a anaphylatoxin. Polypeptide sequence determination and assignment of the oligosaccharide attachment site in C5a. Journal of Biological Chemistry 1978;**253**(19):6955-6964.
140. O'Connor SP, Darip D, Fraley K, Nelson CM, Kaplan EL, Cleary PP. The human antibody response to streptococcal C5a peptidase. Journal of Infectious Diseases 1991;**163**(1):109-116.
141. Park HS, Cleary PP. Active and passive intranasal immunizations with streptococcal surface protein C5a peptidase prevent infection of murine nasal mucosa-associated lymphoid tissue, a functional homologue of human tonsils. Infect Immun 2005;**73**(12):7878-86.
142. Hayano S, Tanaka A. Extracellular esterases of group A streptococci. Infection and immunity 1973;**7**(4):561-566.
143. Zhu H, Liu M, Sumby P, Lei B. The secreted esterase of group a streptococcus is important for invasive skin infection and dissemination in mice. Infection and immunity 2009;**77**(12):5225-5232.
144. Müller-Alouf H, Geoffroy C, Geslin P, Bouvet A, Felten A, Günther E, Ozegowski J-H, Reichardt W, Alouf JE. Serotype, biotype, pyrogenic exotoxin, streptolysin O and exoenzyme patterns of invasive Streptococcus pyogenes isolates from patients with toxin shock syndrome, bacteremia and other severe infections. In. *Streptococci and the Host*: Springer; 1997, 241-243.
145. Nizet V, Beall B, Bast DJ, Datta V, Kilburn L, Low DE, De Azavedo JC. Genetic locus for streptolysin S production by group A streptococcus. Infection and immunity 2000;**68**(7):4245-4254.
146. Palmer M, Vulicevic I, Saweljew P, Valeva A, Kehoe M, Bhakdi S. Streptolysin O: a proposed model of allosteric interaction between a pore-forming protein and its target lipid bilayer. Biochemistry 1998;**37**(8):2378-2383.
147. Ginsburg I. Is streptolysin S of group A streptococci a virulence factor? Apmis 1999;**107**(7-12):1051-1059.
148. Tweten RK. Cholesterol-dependent cytolysins, a family of versatile pore-forming toxins. Infection and immunity 2005;**73**(10):6199-6209.
149. O'Seaghdha M, Wessels MR. Streptolysin O and its co-toxin NAD-glycohydrolase protect group A Streptococcus from Xenophagic killing. PLoS Pathog 2013;**9**(6):e1003394.
150. McCarty M. The inhibition of streptococcal desoxyribonuclease by rabbit and human antisera. Journal of Experimental Medicine 1949;**90**(6):543-553.
151. Wannamaker LW. The differentiation of three distinct desoxyribonucleases of group A streptococci. Journal of Experimental Medicine 1958;**107**(6):797-812.
152. Broudy TB, Pancholi V, Fischetti VA. The in vitro interaction of Streptococcus pyogenes with human pharyngeal cells induces a phage-encoded extracellular DNase. Infection and immunity 2002;**70**(6):2805-2811.
153. Ayoub EM, Wannamaker LW. Evaluation of the streptococcal desoxyribo-nuclease b and diphosphopyridine nucleotidase antibody tests in acute rheumatic fever and acute glomerulonephritis. Pediatrics 1962;**29**(4):527-538.

154. Chesbro W, Walker R. Detection of staphylococcal nuclease elaborated during lethal infections in mice. Infection and immunity 1972;**6**(6):1028.
155. Gray E. Nucleases of group A streptococci. Streptococci and streptococcal diseases. Academic Press, New York, NY 1972:143-155.
156. Karthikeyan G, Mayosi BM. Is primary prevention of rheumatic fever the missing link in the control of rheumatic heart disease in Africa? Circulation 2009;**120**(8):709-713.
157. WHO. *Rheumatic fever and rheumatic heart disease: report of a WHO study group [meeting held in Geneva from 30 March to 4 April 1987]*: World Health Organization; 1988.
158. Pelajo CF, Lopez-Benitez JM, Torres JM, de Oliveira SK. Adherence to secondary prophylaxis and disease recurrence in 536 Brazilian children with rheumatic fever. Pediatric Rheumatology 2010;**8**(1):22.
159. Stewart T, McDonald R, Currie B. Acute rheumatic fever: adherence to secondary prophylaxis and follow up of Indigenous patients in the Katherine region of the Northern Territory. Australian Journal of Rural Health 2007;**15**(4):234-240.
160. Sayed GA. Compliance for secondary prophylaxis for controlling rheumatic fever and rheumatic heart disease in Egyptian children. 2011.
161. Harrington Z, Thomas DP, Currie BJ, Bulkanhawuy J. Challenging perceptions of non-compliance with rheumatic fever prophylaxis in a remote Aboriginal community. Medical journal of Australia 2006;**184**(10):514-517.
162. Nurse-Findlay S, Taylor MM, Savage M, Mello MB, Saliyou S, Lavayen M, Seghers F, Campbell ML, Birgirimana F, Ouedraogo L. Shortages of benzathine penicillin for prevention of mother-to-child transmission of syphilis: An evaluation from multi-country surveys and stakeholder interviews. PLoS medicine 2017;**14**(12):e1002473.
163. Pastural É, McNeil SA, MacKinnon-Cameron D, Ye L, Langley JM, Stewart R, Martin LH, Hurley GJ, Salehi S, Penfound TA. Safety and immunogenicity of a 30-valent M protein-based group a streptococcal vaccine in healthy adult volunteers: A randomized, controlled phase I study. Vaccine 2020;**38**(6):1384-1392.
164. Hysmith ND, Kaplan EL, Cleary PP, Johnson DR, Penfound TA, Dale JB. Prospective Longitudinal Analysis of Immune Responses in Pediatric Subjects After Pharyngeal Acquisition of Group A Streptococci. J Pediatric Infect Dis Soc 2017;**6**(2):187-196.
165. Penfound TA, Chiang EY, Ahmed EA, Dale JB. Protective efficacy of group A streptococcal vaccines containing type-specific and conserved M protein epitopes. Vaccine 2010;**28**(31):5017-22.
166. Barth D, Naicker P, Engel K, Muhamed B, Basera W, Mayosi B, Dale J, Engel M. Molecular Epidemiology of Noninvasive and Invasive Group A Streptococcal Infections in Cape Town. Msphere 2019;**4**(5).
167. Arya DK, Sharma A, Mehta G, Dua M, Johri AK. Molecular epidemiology and virulence characteristics of prevalent group A streptococci recovered from patients in northern India. J Infect Dev Ctries 2014;**8**(3):271-81.
168. Arêas GP, Schuab RB, Neves FP, Barros RR. Antimicrobial susceptibility patterns, emm type distribution and genetic diversity of Streptococcus pyogenes recovered in Brazil. Memórias do Instituto Oswaldo Cruz 2014;**109**(7):935-939.
169. Dale JB. Group A streptococcal vaccines. Infectious disease clinics of North America 1999;**13**(1):227-243.
170. Martins TB, Hoffman JL, Augustine NH, Phansalkar AR, Fischetti VA, Zabriskie JB, Cleary PP, Musser JM, Veasy LG, Hill HR. Comprehensive analysis of antibody responses to streptococcal and tissue antigens in patients with acute rheumatic fever. International immunology 2008;**20**(3):445-452.
171. Shet A, Kaplan EL, Johnson DR, Cleary PP. Immune response to group A streptococcal C5a peptidase in children: implications for vaccine development. The Journal of infectious diseases 2003;**188**(6):809-817.

172. Nordet P, Lopez R, Duenas A & Sarmiento, L. Prevention and control of rheumatic fever and rheumatic heart disease: the Cuban experience (1986-1996-2002): cardiovascular topic. Cardiovascular journal of Africa 2008;19(3), pp.135-140.

173. Bach JF, Chalons S, Mosser A, Forier E, Elana G, Jouanelle J, Kayemba S, Delbois D, Sainte-Aimé C & Berchel C. 10-year educational programme aimed at rheumatic fever in two French Caribbean islands. The Lancet 1996;347(9002), pp.644-648.

Chapter 2

Title: Systematic review and meta-analysis of the prevalence of Group A Streptococcal *emm* clusters in Africa to inform vaccine development.

Taariq Salie[a], Kelin Engel[a], Annesinah Moloi[a,b], Babu Muhamed[a,c], James B. Dale[d], Mark E. Engel[a].

Affiliations

[a]AFROStrep Research Group, Department of Medicine, Faculty of Health Sciences, University of Cape Town & Groote Schuur Hospital, Cape Town, South Africa

[b]Health Systems Research Unit, South African Medical Research Council, Cape Town, South Africa

[c]Children's National Health System, Division of Cardiology, Washington, DC, USA

[d]Division of Infectious Diseases, Department of Medicine, University of Tennessee Health Science Center, Memphis, Tennessee, USA.

Abstract

Background: An *emm* cluster-based system was proposed as a standard typing scheme to facilitate and enhance future studies of Group A Streptococcus (GAS) epidemiological surveillance, M protein function and vaccine development strategies. We provide an evidence-based distribution of GAS *emm* clusters in Africa and assess the potential coverage of the new 30-valent vaccine in terms of an *emm* cluster-based approach.

Method: Two reviewers independently assessed studies retrieved from a comprehensive search and extracted relevant data. Meta-analyses were performed (random effects model) to aggregate *emm* cluster prevalence estimates.

Results: Eight studies (n=1,595 isolates) revealed the predominant *emm* clusters as E6 (18%, 95% confidence interval (CI), 12.6; 24.0%), followed by E3 (14%, 95%CI, 11.2; 17.4%) and E4 (13%, 95%CI, 9.5; 16.0%). There is negligible variation in *emm* clusters as regards regions, age and socio-economic status across the continent. Considering an *emm* cluster-based vaccine strategy, which assumes cross-protection within clusters, the 30-valent vaccine currently in clinical development, would provide hypothetical coverage to 80.3% of isolates in Africa.

Conclusion: This systematic review indicates the most predominant GAS *emm* cluster in Africa is E6 followed by E3, E4 and D4. The current 30-valent vaccine would provide considerable coverage across the diversity of *emm* cluster types in Africa. Future efforts could be directed toward estimating the overall potential coverage of the new 30-valent vaccine based on cross-opsonization studies with representative panels of GAS isolates from populations at highest risk for GAS diseases.

Keywords: Group A Streptococcus, *Streptococcus pyogenes*, GAS, M Protein, *emm* clustering system, epidemiology, surveillance, vaccine, Systematic review

Importance

Low vaccine coverage is of grave public health concern, particularly in developing countries where epidemiological data are often absent. To inform vaccine development for GAS, we report on the epidemiology of the M Protein *emm* clusters from GAS infections in Africa, where GAS-related illnesses and their sequalae including rheumatic fever and rheumatic heart disease, are of a high burden. This first report of *emm* clusters across the continent indicate a high probably of coverage by the M Protein-based vaccine currently undergoing testing, were an *emm*-cluster based approach to be used.

Introduction

Group A Streptococcus (GAS) causes a range of human infections including pharyngitis and impetigo, which can lead to non-suppurative (immune-mediated) sequelae such as acute rheumatic fever (ARF) and rheumatic heart disease (RHD) if not properly managed [1]. Additionally, GAS has the ability to cause invasive infection such as sepsis, necrotizing fasciitis, pneumonia, and streptococcal toxic shock syndrome (STSS) in children and adults [2] with a high fatality rate; furthermore, it is a leading cause of maternal death in some regions [3]. Non-invasive GAS infections mostly affect young children and women living in developing countries [4], while severe invasive infections are both more common in adults (and increases with age) and more common in men than women [5]. The estimated symptomatic GAS pharyngitis annual incidence rate is 0.4 cases per person-year, with over 423 million cases, in children residing in developing countries [1].

The dire complications and huge economic burden of GAS infections and their sequelae, with an estimated global annual incidence of 616 million pharyngitis and 1.78 million severe cases [6], support the urgent need for an effective vaccine that would provide broad coverage of circulating GAS strains [7]. One of the GAS vaccine strategies targets the M-protein on the bacterial surface, which has thermal stability, anti-phagocytic properties and the capacity to evoke antibodies with the greatest bactericidal activity [8]. The hypervariable N-terminal region of the M-protein displays extensive nucleotide differences, thus giving rise to various M-protein amino acid sequences which imparts serological specificity [9]. The 5' *emm* sequence encoding the mature protein is the basis for categorizing different GAS strains through molecular typing methods, which aid in defining the epidemiology of GAS infections.

A 30-valent N-terminal M protein-based vaccine [10] is undergoing clinical trials [11]. The vaccine composition was based on extensive GAS surveillance data from developed regions such as USA and Europe, those isolates that are involved in invasive disease, those associated with superficial infections and those causing autoimmune diseases [12, 13]. However, given the >200 GAS *emm* types characterized to date [14], the absence of highly prevalent GAS subtypes in the current vaccine formulation may diminish the coverage of at-risk populations outside of western countries.

An *emm* clustering system was introduced by Sanderson-Smith and colleagues that phylogenetically analyzed the whole M protein sequences, organizing *emm* types into clusters that have the same or similar sequences and host protein binding properties [15]. This proposed classification allows for the previously identified GAS *emm* types to be categorized into 48 discrete *emm* clusters [15] where more than one *emm* type may be contained within a cluster (Table 1). The *emm* cluster system compliments the *emm* typing system, which may serve to enhance studies relating to M protein function, streptococcal virulence, epidemiological surveillance, and vaccine development [15]. *emm* clusters E1-E6 were placed

into clade X, binding to immunoglobulin and C4BP. While A-C1 through A-C5 and D1-D5 were grouped into clade Y, with a host protein tropism towards plasminogen and fibrinogen.

To date, significant *emm* cluster data have been produced through *emm* typing of GAS, with recent studies reporting on *emm* cluster epidemiology. Shulman documented the most prevalent *emm* clusters in the USA as E4 (27.16%), A-C3 (17.78%) and A-C4 (17.56%) amongst 7,040 isolates [16]. The prevalence of *emm* clusters in three Pacific countries, viz. Australia, Fiji and New Caledonia illustrated that 70%-84% of clusters from isolates were shared, as opposed to comparison of *emm* types having only 14%-30% commonality between countries [17]. In a third study by Chang-Ni in Taiwan, an analysis of both invasive and non-invasive strains revealed that cluster E6 was associated with both types of infections, while clusters D4, E2 and E3 were responsible for causing invasive isolates in their population[18]. Recently, Frost demonstrated that M type–specific and cross-reactive immune responses frequently align with *emm* clusters, raising new opportunities to design multivalent vaccines with broad coverage [19].

A thorough review of *emm* cluster data from Africa has not yet been undertaken. A study that aggregates the African data on clusters is essential to contribute to the growing literature in efforts to develop a GAS vaccine on a global scale, particularly in low-income countries where the burden of disease is greatest. Therefore, this review sought to provide an evidence-based distribution of GAS *emm* clusters in Africa.

Methods

This study employed rigorous methods drawn from the scientific techniques and guidelines offered by the Cochrane Collaboration [20] and by reviews published previously [21, 22]. The review protocol has been registered in the PROSPERO International Prospective Register of Systematic Reviews CRD42017062485.

Review Question

This review asks the following question: What is the prevalence of GAS *emm* clusters in Africa in the current available literature? Is there variation in *emm* cluster prevalence based on geography, age, clinical manifestation or socio-economic status? We further sought to explore the potential coverage of the current 30-valent vaccine using a cluster-based approach.

Table 1 – The *emm* clusters and their corresponding *emm* types. Adopted from Sanderson-Smith, De Oliveira [15]

***emm* types**	***emm* cluster**	**Clade**
4, 60, 78, 165 (st11014), 176 (st213)	E1	Clade
13, 27, 50 (50/62), 66, 68, 76, 90, 92, 96, 104, 106, 110, 117, 166 (st1207), 168 (st1389)	E2	X
9, 15, 25, 44 (44/61), 49, 58, 79, 82, 87, 103, 107, 113, 118, 144 (stknb1), 180 (st2460), 183 (st2904), 209 (st6735), 219 (st9505), 231 (stNS292)	E3	
2, 8, 22, 28, 73, 77, 84, 88, 89, 102, 109, 112, 114, 124, 169 (st1731), 175 (st212), 232 (stNS554)	E4	
34, 51, 134 (st2105), 137 (st465), 170 (st1815), 174 (st211), 205 (st5282)	E5	
11, 42, 48, 59, 63, 65 (65/69), 67, 75, 81, 85, 94, 99, 139 (st7323), 158 (stxh1), 172 (st2037), 177 (st2147), 182 (st2861UK), 191 (st369)	E6	
164 (st106), 185 (st2917), 211 (st7406), 236 (sts104)	Single type *emm* clusters	
36, 54, 207 (st6030)	D1	Clade
32, 71, 100, 115, 213 (st7700)	D2	Y
123, 217 (st809)	D3	
33, 41, 43, 52, 53, 56, 56.2 (st3850), 64, 70, 72, 80, 83, 86, 91, 93, 98, 101, 108, 116, 119, 120, 121, 178 (st22), 186 (st2940), 192 (st3757), 194 (st38), 208 (st62), 223 (stD432), 224 (stD631), 225 (stD633), 230 (stNS1033), 242 (st2926)	D4	
97, 157 (stn165), 184 (st2911)	D5	
46, 142 (st818)	A-C1	
30, 197 (st4119)	A-C2	
1, 163 (st412), 227 (stil103), 238, 239	A-C3	
12, 39, 193 (st3765), 228 (stil62), 229 (stmd216)	A-C4	
3, 31, 133 (st1692)	A-C5	
5, 6, 14, 17, 18, 19, 23, 24, 26, 29, 37, 38 (38/40), 47, 57, 74, 105, 122, 140 (st7395), 179 (st221), 218 (st854), 233 (stNS90), 234 (stpa57)	Single type *emm* clusters	
55, 95, 111, 215 (st804), 221 (stCK249), 222 (stCK401)	Single type *emm* clusters	Outliers

Reprinted from Sanderson-Smith, De Oliveira [15] with permission from the publisher

Search Strategy

A comprehensive strategy was developed to search electronic databases to maximize sensitivity (Table S1 – Appendix B). The search strategies incorporated both free term text that are controlled to suit specific databases individually and Medical Subject Headings (MeSH) adapted to suit each individual database. A combination of terms relating to "*emm* typing", "*emm* clusters", "*emm*/M protein" and "streptococcal diseases" focusing on the African continent by applying the African search filter previously used by Pienaar and colleagues [23]. The following databases were searched as of 29 April 2020; PubMed, Scopus and Google Scholar for grey literature. The search was not restricted to any publication dates or language (however, abstracts must have been clearly written in English for the

study to be considered). Published and unpublished data, such as grey literature including theses and conference proceedings, were also considered for inclusion.

Inclusion criteria

All studies that described the prevalence of *emm* clusters or *emm* types within a given population were included in the review. Participants were restricted to the African continent but were not discriminated by clinical manifestation of GAS or site of GAS isolation. All laboratory-confirmed GAS isolates were molecularly characterized by the *emm* typing method as developed by Beall [24] and in alignment with the Centers for Disease Control and Prevention [25], to classify GAS according to sequence analysis of the 5' hypervariable region of the M protein gene.

Two reviewers applied the search strategy to the relevant databases independently in which the titles and abstracts were evaluated to exclude studies that did not describe the prevalence of GAS. Thereafter, full texts of the included titles and abstracts were retrieved and further evaluated against the inclusion criterion (Table S2 – Appendix B). A comparison was made between individual lists, if the reviewer's lists were not concurrent, discrepancies were discussed, and an arbitrator (third reviewer) was contacted to resolve any disagreements.

Exclusion criteria

Case reports, narrative reviews, opinion pieces and publications lacking prevalence primary data, or referenced methodology is not according to Beall [24], were excluded from the review. Duplicated studies of the same datasets and participants were removed, and the final most recent publication of the data was considered for inclusion.

Data extraction and management

Two reviewers extracted data using a standardized data extraction form and any contradictions were solved through discussion or that of a third reviewer. Search results from the databases listed above, published and unpublished studies were managed with Endnote X9 referencing software. Briefly, data extraction consisted of recording the study demographics (amount of study participants, the geographical region, age group of enrolled participants, the clinical manifestation of disease and socio-economic status) along with the relevant *emm* type/cluster distributions within the population. Socio-economic status for the study settings was determined at a country level, according to The World Bank [26].

Quality assessment

The risk of bias assessment established by Hoy [27] and modified by Werfalli [22], was adapted in questions specific for use in this review (Table S3 – Appendix B). Using a quantitative scoring system, studies were characterized being of a low, moderate or high risk of bias. A study with low risk of bias is of high-quality and a low-quality study is associated with a higher risk of bias. Assessing the risk of bias informs the evaluation of heterogeneity in the pooled analyses.

Analysis

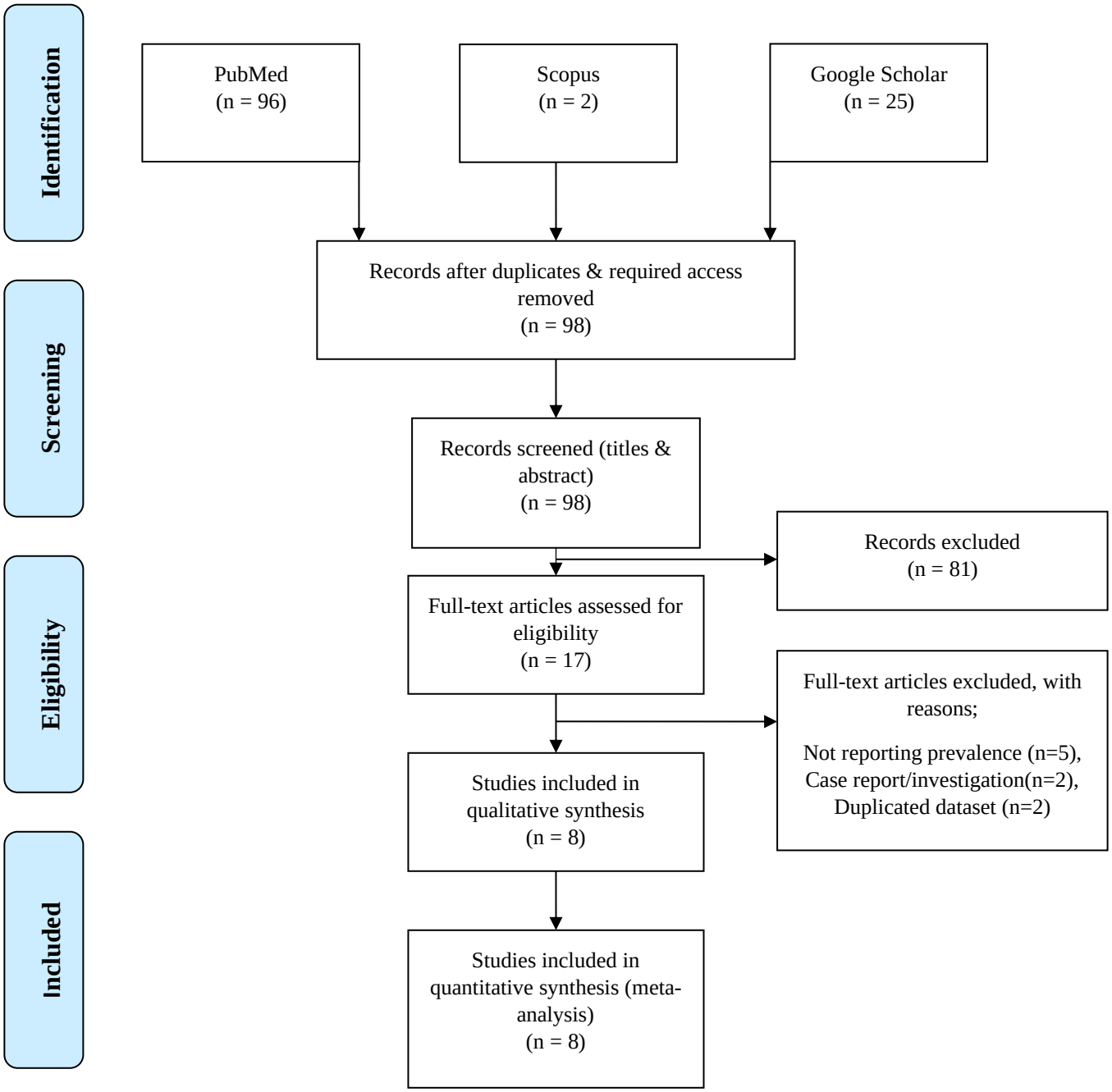
Data synthesis included three steps: (1) characterizing the study demographics (2) documenting *emm* types for *emm* cluster calculations, and (3) assessing potential vaccine coverage. In each study, the prevalence of *emm* types was recalculated by analyzing figures and tables to confirm the authors results and findings and to document the numerators and denominators. In older studies, *emm* typing information needed to be updated using the CDC database [28]. Where *emm* cluster information was not reported, the CDC classification system was used to augment missing data (https://www.cdc.gov/groupastrep/lab.html), as well as the original cluster descriptions [15].

To calculate potential coverage, three tiers were assessed: 1) M peptides in the vaccine, 2) *emm* types that have been shown to be cross-opsonized, and 3) *emm* types that just happen to be in a cluster that are represented by one or more vaccine *emm* types. Quantitative data analysis was completed using Stata version 14.1 (StataCorp, College Station, TX, USA). We applied the Freeman-Tukey double arcsine transformation option using the *metaprop* routine to describe the combined prevalence estimates of all included studies with the standard error across the unadjusted estimates [29]. *emm* cluster distribution was correlated against different variables (resource setting, clinical manifestation and age group) in each of the studies. Lastly, we determined the theoretical protective coverage by *emm* cluster cross-opsonization for *emm* types included in the M protein-based vaccine [12].

Results

The literature search for articles was reported according to the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) Statement [30]. Figure 1 details the search results with the retrieval of 121 articles for consideration from the respective electronic databases. After title screening and the removal of duplicates, we excluded 23 articles. We reviewed the remaining abstracts and excluded a further 81 articles, leaving 17 articles requiring full-text evaluation. Finally, eight articles met the inclusion criteria and were included in the review. A list of the excluded studies with reasons are detailed in Table S4 (Appendix B).

188



189 **Figure 1 – Schematic PRISMA flow diagram of the literature search.** Figure is modelled after
190 Moher, Liberati [30].

Characteristics of included studies

The included articles were published between 2004 and 2019 with sample sizes ranging from 43 and 396 total isolates. Of these, two articles had cross-sectional study designs, while the remaining studies took a prospective passive surveillance approach. The ages of participants included in the studies were also recorded; six articles studied isolates obtained from children (range 0-18 years old) and two, studied patients of all ages. Studies were conducted in local and university hospitals, clinics, outpatient departments and schools situated in the study areas (Table 2).

The country of each article was recorded, with 2 articles obtained from Ethiopia [31, 32], South Africa [33, 34], Tunisia [35, 36] and one article from Kenya [37] and Mali [38]. All the studies included in this review made use of the gold-standard, *emm* typing molecular procedure proposed by Beall [24] and the CDC [25].

Prevalence of GAS *emm* clusters

Five countries within Africa contributed *emm* cluster data to this review (Figure 2). The final dataset included 1,532 isolates representing 126 heterologous *emm* types. Of these, 1,291 isolates, comprising 96 *emm* types, constituted 16 *emm* clusters. Of those remaining, 186 isolates contained 18 single-isolate *emm* clusters [15 *emm* types (143 isolates) representing 15 *emm* clusters belonging to clade Y, while *emm*55, *emm*95 and *emm*111 constituted outliers (43 isolates)] (Table 3). The remaining 12 *emm* types (55 isolates) are amongst those as yet not classified. The predominant clusters were E6 with 294 isolates (18.4%), followed by E3 (n=243, 15.2%) and E4 (n=225, 14.1%). The *emm* clusters with the least number of isolates are D1 and E5, respectively having a single isolate. *emm* cluster A-C1 was not represented. Sixty-three isolates were reported as 'untypable' by authors, thus not assigned an *emm* type, or an 'old' *emm* type that does not correspond with the CDC classification.

There were four regions represented across Africa. Variation of clusters across the regions was not remarkable. Interestingly, single-isolate cluster M55 was specific to Mali in West Africa, containing 16 isolates. The highest single-isolate cluster, M18 (n=41 isolates), was not represented in South Africa. Where age of participants in studies was provided, there was no difference amongst children (<18 years of age) and adults in terms of cluster prevalence. By clinical manifestation, isolates from invasive disease numbered 516 (32.4%) (Figure S1 – Appendix B) with the most prevalent clusters being E3 (n=91 isolates) followed by E6 (n=82) and D4 (n=77). No variation in *emm* clusters by socio-economic status was apparent (data not shown).

Table 2 – Characteristics of included studies

Study ID	Country	Region	Design	Setting (Local, social context)	Socio-economic setting	Population description	Inclusion criteria	Age(s)
Abdissa 2005	Ethiopia	Africa	Cross-sectional	Public primary schools situated in Addis-Ababa, Gondar & Dire-Dawa	Low income	Healthy children attending public primary schools	Healthy school children aged 4-16 in area	4-16 years
Barth 2019	South Africa	Africa	Prospective passive surveillance	Groote Schuur Hospital in Cape Town	Upper middle income	Healthy children attending public primary schools	Patients with confirmed Strep A infection	ALL
Engel 2014	South Africa	Africa	Cross-sectional	Langa Clinic, Vanguard Community Health Centre, UCT, AFROStrep Registry	Upper middle income	Children located in the Vanguard area	Healthy children aged 3-15 years, with sore throat	3-15 years
Hraoui 2011	Tunisia	Africa	Passive surveillance	Microbiology Lab of Charles Nicolle University Hospital in Tunis	Lower middle income	Patients attending the local hospital	Patients with confirmed Strep A infection	ALL
Mzoughi 2004	Tunisia	Africa	Prospective surveillance	Farhat Hached Hospital & Centre PMI, Sousse	Lower middle income	Paediatric outpatients attending clinic	Children aged 2-8 years, with pharyngitis	2-8 years
Seale 2016	Kenya	Africa	Prospective surveillance	Kenya Medical Research Institute, Kilifi County Hospital	Lower middle income	Children admitted for medical care at the hospital	Children located in the area with confirmed iStrep A	0-12 years
Tapia 2016	Mali	Africa	Prospective surveillance	4 Public schools in Djikoroni-para & Sebenikoro	Low income	Children attending 1 of the 4 schools	Children aged 5-16 years with sore throat	5-16 years
Tewodros 2005	Ethiopia	Africa	Prospective surveillance	Black Lion Hospital & 3 elementary schools in Addis Ababa	Low income	Pediatric patients attending the hospital & schools within area	Healthy children at schools & those with confirmed ARF, RHD, APSGN & impetigo	<18 years

iStrep A, invasive Group A Streptococcus; ARF, acute rheumatic fever; APSGN, acute poststreptococcal glomerulonephritis

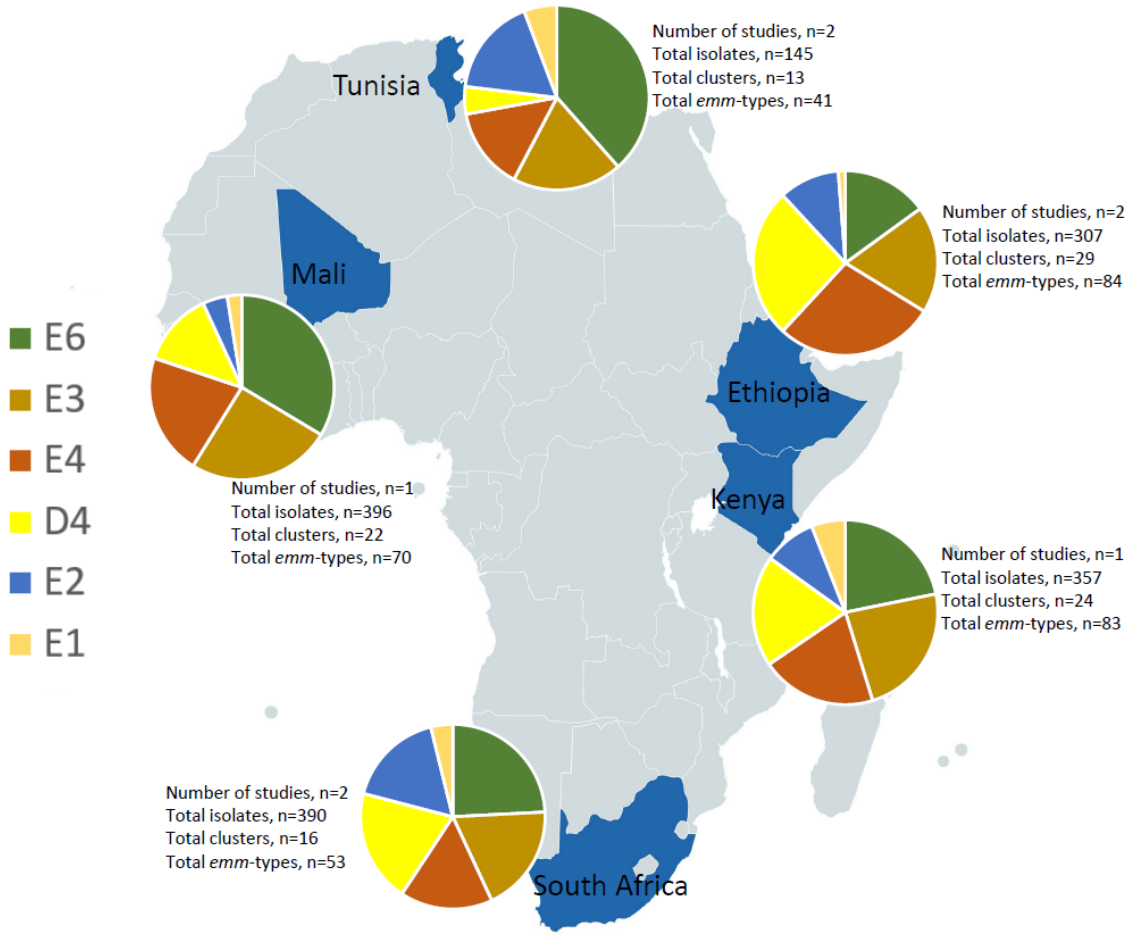


Figure 2 – The five countries included in the review, representing the most abundant *emm* clusters

Table 3 – The *emm* cluster distribution, representing the five countries included into the review and their respective isolate count

***emm* cluster**		**No. of isolates**					**Total no. (%)**
		Ethiopia	**Kenya**	**Mali**	**South Africa**	**Tunisia**	
Clade Y	A-C1	0	0	0	0	0	0 (0.0)
	A-C2	4	1	0	0	0	5 (0.3)
	A-C3	12	3	1	11	9	36 (2.3)
	A-C4	17	3	0	13	7	40 (2.6)
	A-C5	12	0	1	2	4	19 (1.2)
	D1	1	0	0	0	0	1 (0.1)
	D2	10	4	6	1	0	21 (1.4)
	D3	1	3	7	0	0	11 (0.7)
	D4	42	49	36	67	5	199 (13.0)
	D5	3	3	6	9	0	21 (1.4)
Clade X	E1	2	15	7	13	6	47 (3.1)
	E2	17	23	12	58	18	128 (8.4)
	E3	30	59	70	64	20	243 (15.9)
	E4	45	51	59	55	15	225 (14.7)
	E5	1	0	0	0	0	1 (0.1)
	E6	24	55	93	82	40	294 (19.2)
Single type cluster Clade Y	M5	7	0	0	0	0	7 (0.5)
	M6	4	0	0	5	4	13 (0.8)
	M14	1	0	0	0	1	2 (0.1)
	M18	8	12	17	0	4	41 (2.7)
	M19	1	2	7	1	0	11 (0.7)
	M26	0	1	0	0	1	2 (0.1)
	M29	4	0	0	0	0	4 (0.3)
	M38	2	1	0	0	0	3 (0.2)
	M57	1	1	0	0	0	2 (0.1)
	M74	12	4	8	1	0	25 (1.6)
	M105	2	0	1	0	0	3 (0.2)
	M122	0	2	8	0	0	10 (0.7)
	M179	1	10	1	0	0	12 (0.8)
	M218	2	3	2	0	0	7 (0.5)
	M234	0	0	0	1	0	1 (0.1)
Outliers	M55	1	6	16	0	0	23 (1.5)
	M95	2	2	7	2	0	13 (0.8)
	M111	0	4	3	0	0	7 (0.5)
No *emm* cluster[a]		20	11	21	3	0	55 (3.6)
Total[b]		307	357	396	390	145	1532

[a]Refers to those *emm* types that has not been assigned to a particular clade by Sanderson-Smith et al. (2014)
[b]Sixty-three isolates were 'untypable' by the author and was not assigned an *emm* type, or an 'old' *emm* type that does not correspond with the CDC classification.

227

Overall Prevalence of GAS *emm* clusters represented by the *emm* types included in the 30-valent vaccine

Cluster E6 was the most represented *emm* cluster (17.97% (95% confidence interval (CI) 12.6% to 24.0%) amongst African isolates (Figure 3A). This was followed by E3 [14.17% (95% CI 11.2; 17.4)], E4 [12.6% of isolates (95% CI 9.5; 16.0)], D4 [10.88% (95% CI 6.9; 15.5%)] and E2 [9.12% (95% CI 4.6; 14.9)] of isolates (Figure 3B-D). Clusters A-C3, A-C4, A-C5 and E1 each have an effect size of ~2%. Isolates from invasive disease were abundant in clusters D4, E2, E3 and E4 while only E6 had a preponderance of strains from non-invasive disease (Table 4).

M Protein Vaccine Coverage

Just over eighty percent (80.3%) of African GAS isolates are classified in clusters included in the 30-valent vaccine (Figure 4). However, based on *emm* types within the vaccine, together with *emm* types known to be cross-opsonized, the number of African GAS isolates that potentially could be covered by the 30-valent vaccine amounts to 892 of the 1532 isolates corresponding to 58.22% (comprising 599 vaccine type *emm* types and 293 non-vaccine *emm* types) [12]. For the *emm* types representing the remaining 640 isolates (41.78%), there is either no information yet available about possible cross-protection, or the *emm* types would not be expected to cross-react with the 30-valent vaccine antisera because they are in single-*emm* clusters or in clusters not represented by the vaccine. Interestingly, isolates classified as *emm*30 (AC-2), *emm*36 (D1), *emm*51 (E5) and *emm*97 (D5), despite not being in a cluster represented in the vaccine, are nevertheless afforded cross-protection.

With regards to invasive *emm* types in Africa, the overall potential coverage of the vaccine (based on published results of cross-opsonization) was 54.1% for clusters included in the meta-analyses (Table 4). More specifically, coverage for clusters E6, E4 and E2 ranges from 69-74% of invasive isolates; only ~50% of strains would be protected in E3 and coverage for the remaining clusters were below 47% except A-C5 (100% coverage) as there were only two invasive strains reported. Interestingly, the 30-valent vaccine would potentially only provide 12% coverage to invasive isolates belonging to the fourth highest cluster, D4 (n=28 *emm* types).

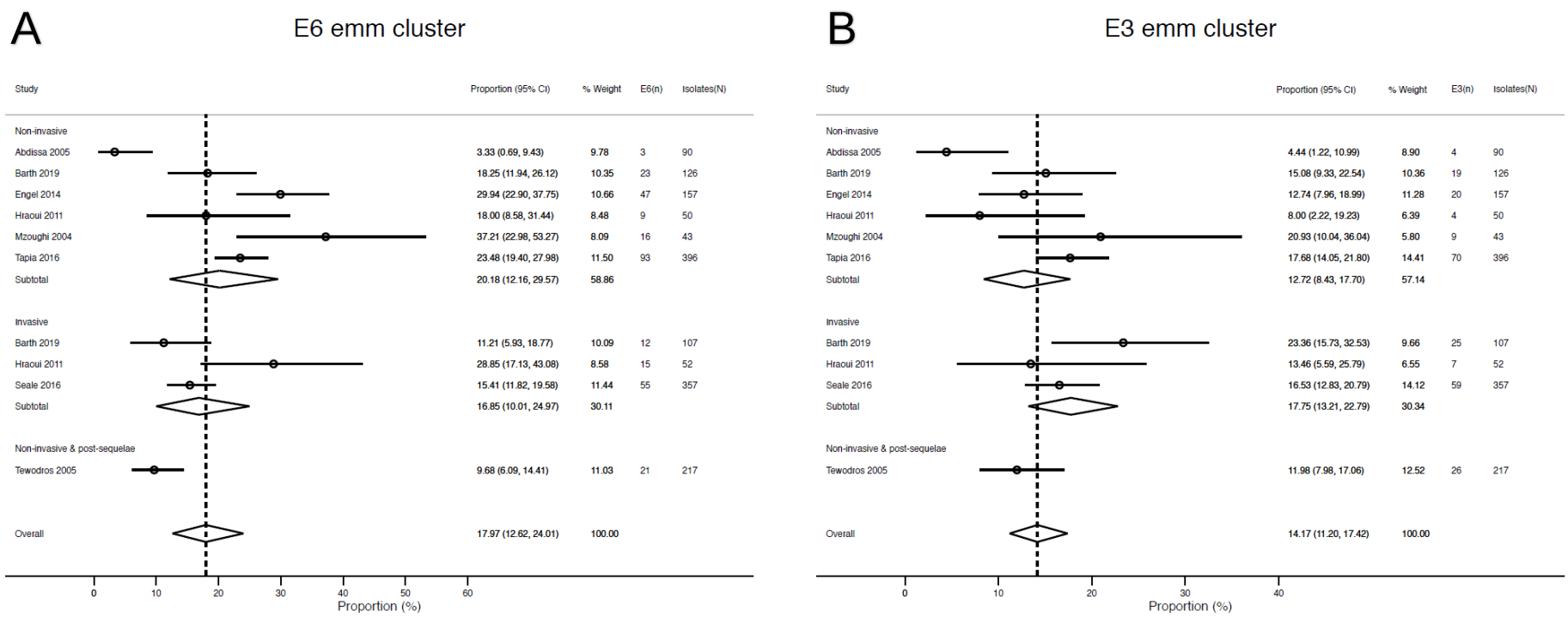


Figure 3 – Forest plots showing the combined prevalence estimates of the four most abundant emm clusters of all included studies.

(A) E6 emm cluster; (B) E3 emm cluster;

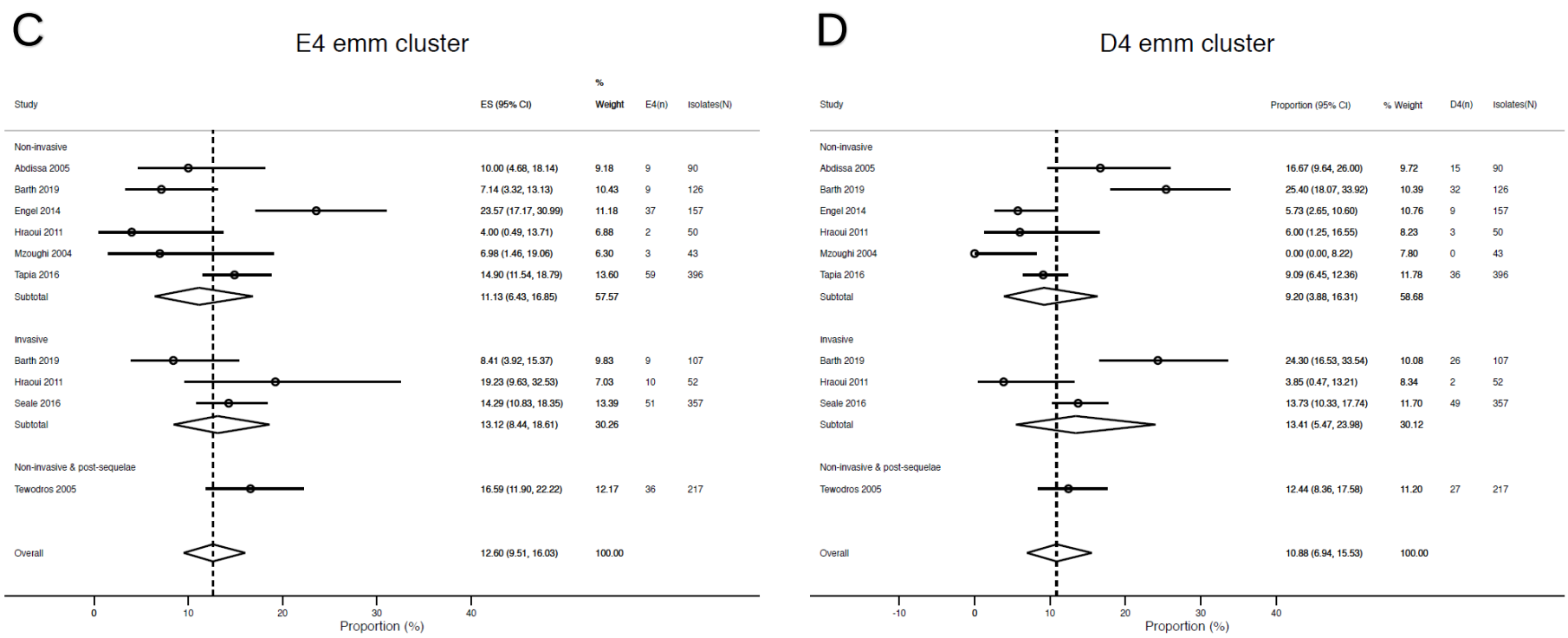


Figure 3 – Forest plots showing the combined prevalence estimates of the four most abundant emm clusters of all included studies (cont').

(C) E4 emm cluster; (D) D4 emm cluster. Studies are represented by author, year of publication, effect size (ES) (proportion [95% CI]). n, number of isolates in cluster; N, total number of isolates in study

Table 4 – Summary of the meta-analyses completed[a]

Cluster	*emm* types in vaccine	*emm* types in African isolates[b]	Combined prevalence (95% CI)	Number of *emm* type isolates			
				NINV (%)	INV (%)	N/D	Total
E6	11, 75, 81	Prot: *11*, ***48***, ***59***, ***65***, *75, 81*, ***85***, ***94*** *(n=235 [79,9%]);* N/P: *42, 63, 67, 99, 158, 182;* NINV: *n=160/191 (83.8%);* INV: *n=59/82 (72.0%)*	18.0% (12.62; 24.01)	191 (65.0)	82 (27.9)	21 (7.1)	294
E3	44, 49, 58, 87,118	Prot: ***9***, ***15***, *44, 49, 58,* ***79***, *87, 118 (n=130 [53.5%]);* N/P: *25, 61, 82, 103, 113, 180, 183, 209;* NINV: *n=73/126 (57.9%);* INV: *n=47/91 (51.6%)*	14.2% (11.20; 17.42)	126 (51.9)	91 (37.4)	26 (10.7)	243
E4	2, 22, 28, 73, 77, 89, 114	Prot: *2,* ***8***, *22, 28, 73, 77, 89,* ***109***, *114 (n=202 [89.8%]);* N/P: *84, 112, 124, 175, 225;* NINV: *n=115/119 (96.6%);* INV: *n=52/70 (74.3%)*	12.6% (9.51; 16.03)	119 (52.9)	70 (31.1)	36 (16.0)	225
D4	83	Prot: ***33***, *83 (n=15 [7.5%]);* N/P: *41, 43, 53, 56, 64, 70, 80, 86, 93, 98, 116, 119, 121, 186, 192, 208, 223, 224, 230;* NINV: *n=5/95 (5.3%);* INV: *n=9/77 (11.7%)*	10.9% (6.94; 15.53)	95 (47.7)	77 (38.7)	27 (13.6)	199
E2	92	Prot: ***66***, ***68***, ***76***, *92,* ***102*** *(n=92 [71.9%]);* N/P: *50, 90, 104, 106, 110, 168;* NINV: *n=49/63 (77.8%);* INV: *n=38/55 (69.1%)*	9.1% (4.61; 14.86)	63 (49.2)	55 (43.0)	10 (7.8)	128
E1	4, 78	Prot: *4, 78 (n=31 [66.0%]);* N/P: *60, 165;* NINV: *n=22/28 (78.6%);* INV: *n=8/17 (47.1%)*	1.9% (0.62; 3.81)	28 (59.6)	17 (36.2)	2 (4.3)	47
A-C4	12	Prot: *12 (n=35 [87.5%]);* N/P: *39, 193, 229;* NINV: *n=17/18 (94.4%);* INV: *n=3/7 (42.9%)*	2.1% (0.37; 4.81)	18 (45.0)	7 (17.5)	15 (38.0)	40
A-C3	1	Prot: *1 (n=32 [88.9%]);* N/P: *238;* NINV: *n=21/22 (95.5%);* INV: *n=1/4 (25.0%)*	2.0% (0.46; 4.31)	22 (61.1)	4 (11.1)	10 (27.8)	36
A-C5	3	Prot: *3 (n=19 [100%]);* N/P: NA; NINV: *n=17/17 (100%);* INV: *n=2/2 100%)*	0.9% (0.01; 2.72)	17 (89.5)	2 (10.5)	0 (0.0)	19

[a]Three hundred and one isolates (19.6%) comprising 39 *emm* types are not included in any of the *emm* clusters contained in the 30-valent vaccine; these include 60 isolates representing seven *emm* clusters; 186 isolates representing single-isolate clusters & 55 isolates were not classified as according to Sanderson-Smith.
[b]Bold *emm* types represent cross-opsonized non-vaccine types. The study completed by Tewodros & Kronvall (2005) did not clearly differentiate its *emm* types according to clinical manifestation. Combined prevalence calculated with M-H meta-analysis procedures. NINV, non-invasive; INV, invasive; Prot, protected; N/P, not protected; N/D, not differentiated; NA, none.

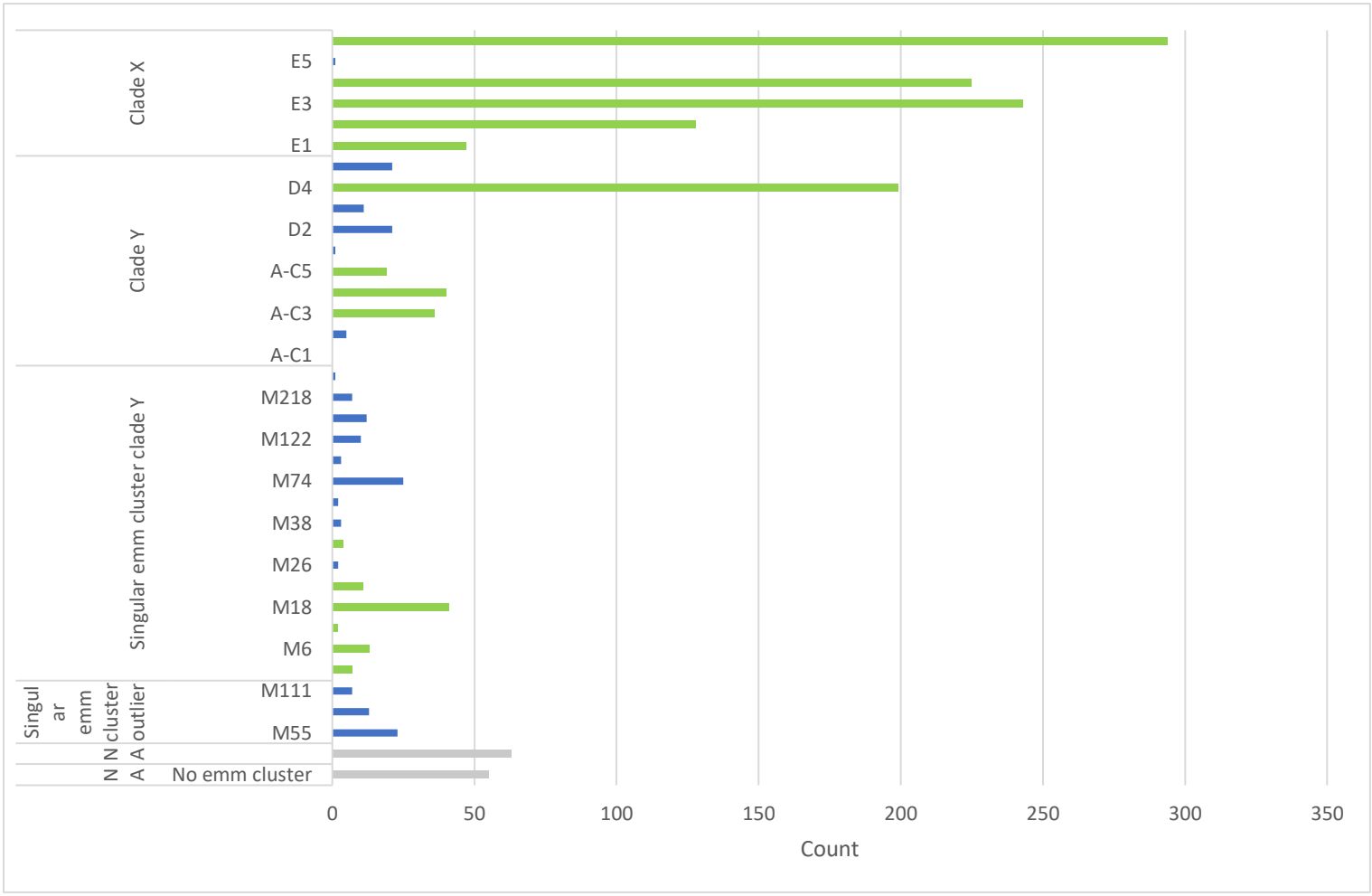


Figure 4 – Count of isolates within emm clusters from African studies included in this systematic review. Green bars indicate emm clusters represented in the 30-valent vaccine. Blue bars represent emm clusters not included in the vaccine. Gray bars represent isolates unassigned to a cluster or were "untypeable" according to the authors' report. Numbers represent the count of isolates across all studies

Assessment of risk of bias of included studies

The results from the assessment are portrayed in Table 5, with two studies having a low risk of bias [34, 38] and the remaining studies being of moderate bias. All the articles narrowed down their target population by focusing on a specific age group, clinical manifestation, or geographical area. The data in all included studies were collected directly from the study participants as opposed to by proxy, confirming the reliability of sample collection and patient demographics. The included studies clearly described the phenotypes of patients, providing an acceptable case definition or diagnostic algorithm. Studies focusing on invasive GAS infections isolated GAS from normally sterile sites such as blood, cerebrospinal fluid, joints, bones or synovium amongst others. Non-invasive GAS was isolated from skin or throat via swabs of the infected area.

Table 5 – Summary of risk of bias assessment

Study ID	Risk of bias:[a]							Quality score	Risk of bias
	1	2	3	4	5	6	7		
Abdissa 2005	0	1	1	1	1	1	0	5	Moderate
Barth 2019	0	1	1	1	1	1	1	6	Low
Engel 2014	0	1	1	1	1	1	0	5	Moderate
Hraoui 2011	0	1	1	1	1	1	0	5	Moderate
Mzoughi 2004	0	1	1	1	NCS[b]	1	0	4	Moderate
Seale 2016	0	1	1	1	1	1	0	5	Moderate
Tapia 2016	0	1	1	1	1	1	1	6	Low
Tewodros 2005	0	1	1	1	1	1	0	5	Moderate

[a]Risk of bias categories: 1 & 2, Representativeness of population; 3 & 6, Data collection; 4, Case definitions; 5, Study instrument reliability; 7, Limitations
[b]NCS, Not Clearly Stated.

Discussion

This systematic review provides evidence for the distribution of *emm* clusters of GAS in Africa, specifically focusing on the epidemiological differences within Africa and added value of the *emm* clustering system in assisting with vaccine development. Using prevalence data obtained from eight studies representing five countries within Africa, this report identified the predominant *emm* clusters in Africa, namely E6 followed by E3, E4 and D4. We further report that the *emm* clusters contained in the current 30-valent vaccine could provide just over 80% coverage across the diversity of *emm* cluster types in Africa.

Comparing results to other *emm* clustering epidemiology studies, it is clear that the dominant *emm* clusters vary between regions. Only cluster E3 in the present study is common with the Pacific region [17]. When comparing the current data to the USA, E4, the highest cluster in USA is the third highest cluster, whereas A-C3 and A-C4 together only amount to ~2% of the total strains isolated in Africa [16]. This study emphasizes that *emm* clusters E6, E3 and D4, prevalent in the African populations where the burden of GAS infections is highest [39], should take prominence alongside clusters E4, A-C3 and A-C4.

We note that there are a number of *emm* clusters containing a single *emm* type as they do not share similar binding properties or sequences. Also, there are many *emm* types that have as yet not been categorized into a particular cluster, as this may be due to their recent emergence post the proposed cluster system. This should be the focus of future studies in which more associations with human host protein binding could be tested to determine any other similarities between single-*emm* clusters.

Steer et al, reported that the African and Asian regions had the greatest diversity of *emm* types [14]. This could be due to a variety of factors causing site-tissue tropism and disease manifestation, promoting the dominance of heterologous *emm* types in different regions [40]. Our review provides no evidence for marked variation across the continent amongst most of the more prominent *emm* clusters. When considering the ages of participants infected with GAS, there appears to be no differences compared to that of the overall estimates. There is an increased risk for the transmission of GAS in poorer countries due to household crowding and the lack of income for proper healthcare [41]. Evaluating socio-economic status amongst our studies revealed little to no differences in *emm* cluster data.

By clinical manifestation in terms of the invasive nature of the infection, amongst non-invasive infections, cluster E6 was the most abundant cluster. This is in accordance with previous

studies conducted in similarly impoverished areas in India [42, 43] and Brazil [44], that identified *emm* types belonging to cluster E6 (*emm*75, *emm*81) as the predominant isolates. However, in invasive disease, the predominant *emm* cluster is E3, followed by E6 and D4, which are comparable to the *emm* cluster data shown in Southern Taiwan [18]. A study conducted on invasive isolates in the USA suggests that cluster E3 and E6 is the third and fifth highest cluster, respectively [45].

In terms of the current 30-valent vaccine [12], with the assumption that the *emm* type prevalence data from the eight included studies could be generalised for the entire continent, vaccine coverage would be 55.92% of strains isolated in Africa. Frost had shown cross-reactive protection of a single *emm* type with the remaining *emm* types within the same cluster, specifically that of E4 [19]. Thus, hypothetically assuming that if a single *emm* type in the 30-valent vaccine would provide cross-protection to the remaining isolates within the cluster, a *emm* cluster-based vaccine would then extend coverage to ~80% protection against GAS (Figure 4). Of interest cluster D4, which comprises 28 heterologous *emm* types, and ranked high in this analysis, has only a single representation (*emm*83) included in the vaccine. If cross-protection were to occur within clusters, more *emm* types belonging to cluster D4 ought to gain a particular importance for inclusion into new vaccines, especially since D4 (10.9% of isolates) is the fourth highest abundant cluster within Africa. It is also important to note that coverage extended to invasive isolates was sub-optimal (n=219, 54.1%, inclusive of cross-protection afforded to non-vaccine types).

One of the main strengths of this review is attributed to the use of multiple databases searched, using an African search filter and a robust approach to the meta-analysis of the data. We systematically and purposefully assessed all the data available with no language exclusions, or restrictions to a clinical manifestation of disease, using the most recently published standard quality assessment tools for prevalence studies. We also assessed the risk of bias present in the individual articles, showing that the quality was reasonably high, thus allowing for comparisons across the studies. The main limitations of the review are due to the lack of epidemiological data obtained from low to middle income countries in Africa, especially given their relatively high burden of GAS infections. The inclusion of more articles reporting on the prevalence of GAS may further assist in distinguishing differences amongst the geographical location, age and socio-economic categories. A further limitation to the results of our systematic review is the significant heterogeneity in the prevalence estimates produced in the meta-analysis, however, this is expected when pooling prevalence studies. We made use of the

Freeman-Tukey double arc-sine transformation to stabilize the variance of primary studies before pooling, thus limiting the impact of studies with either small or large prevalence on the overall pooled estimates, as well as across major subgroups [29].

Conclusion

In conclusion, this systematic review provides the latest evidence for the distribution of *emm* clusters of GAS in Africa. We show that there is negligible variation in *emm* clusters as regards regions, age and socio-economic status across the continent. We further report that the current 30-valent vaccine will provide considerable coverage across the diversity of *emm* cluster types in Africa, thus providing direction for future work to include coverage of clusters D4, E2-E4 and E6, given that they comprise 83% of the total isolates obtained in Africa.

References

1. Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. The Lancet infectious diseases 2005;**5**(11):685-694.
2. Bronze MS, Dale JB. The reemergence of serious group A streptococcal infections and acute rheumatic fever. The American journal of the medical sciences 1996;**311**(1):41-54.
3. van Dillen J, Zwart J, Schutte J, van Roosmalen J. Maternal sepsis: epidemiology, etiology and outcome. Current opinion in infectious diseases 2010;**23**(3):249-254.
4. Madden S, Kelly L. Update on acute rheumatic fever: It still exists in remote communities. Canadian Family Physician 2009;**55**(5):475-478.
5. Sanyahumbi S, Colquhoun S, Wyber R, Carapetis J. Global disease burden of Group A Streptococcus. Streptococcus pyogenes: Basic biology to clinical manifestations 2016:661-704.
6. Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. Lancet Infect Dis 2005;**5**(11):685-94.
7. Steer AC, Dale JB, Carapetis JR. Progress toward a global group a streptococcal vaccine. J The Pediatric infectious disease journal 2013;**32**(2):180-182.
8. Tandon R. Preventing rheumatic fever: M-protein based vaccine. Indian heart journal 2014;**66**(1):64-67.
9. Fischetti VA. Streptococcal M protein: molecular design and biological behavior. Clinical microbiology reviews 1989;**2**(3):285-314.
10. Dale JB, Penfound TA, Chiang EY, Walton WJ. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine 2011;**29**(46):8175-8.
11. Pastural E, McNeil SA, MacKinnon-Cameron D, Ye L, Langley JM, Stewart R, Martin LH, Hurley GJ, Salehi S, Penfound TA, Halperin S, Dale JB. Safety and immunogenicity of a 30-valent M protein-based group a streptococcal vaccine in healthy adult volunteers: A randomized, controlled phase I study. Vaccine 2020;**38**(6):1384-1392.
12. Dale JB, Penfound TA, Chiang EY, Walton WJ. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine 2011;**29**:8175-8.
13. Hu MC, Stroop SD, Reddish MA, Beall B, Dale JB. Immunogenicity of a 26-Valent Group A Streptococcal Vaccine. Infection and Immunity 2002;**70**(4):2171-2177.
14. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. The Lancet infectious diseases 2009;**9**(10):611-616.
15. Sanderson-Smith M, De Oliveira DM, Guglielmini J, McMillan DJ, Vu T, Holien JK, Henningham A, Steer AC, Bessen DE, Dale JB. A systematic and functional classification of Streptococcus pyogenes that serves as a new tool for molecular typing and vaccine development. The Journal of infectious diseases 2014;**210**(8):1325-1338.
16. Shulman ST, Tanz RR, Dale JB, Steer AC, Smeesters PR. Added value of the emm-cluster typing system to analyze group A Streptococcus epidemiology in high-income settings. Clinical Infectious Diseases 2014;**59**(11):1651-1652.
17. Baroux N, D'Ortenzio E, Amedeo N, Baker C, Ali Alsuwayyid B, Dupont-Rouzeyrol M, O'Connor O, Steer A, Smeesters PR. The emm-cluster typing system for Group A Streptococcus identifies epidemiologic similarities across the Pacific region. Clin Infect Dis 2014;**59**(7):e84-92.
18. Chiang-Ni C, Zheng PX, Wang SY, Tsai PJ, Chuang WJ, Lin YS, Liu CC, Wu JJ. Epidemiology Analysis of Streptococcus pyogenes in a Hospital in Southern Taiwan by Use of the Updated emm Cluster Typing System. J Clin Microbiol 2016;**54**(1):157-62.
19. Frost HR, Laho D, Sanderson-Smith ML, Licciardi P, Donath S, Curtis N, Kado J, Dale JB, Steer AC, Smeesters PR. Immune cross-Opsonization within emm clusters following group A Streptococcus skin infection: broadening the scope of type-specific immunity. Clinical infectious diseases 2017;**65**(9):1523-1531.

20. Deeks J, Higgins J, Altman D, Green S. Cochrane handbook for systematic reviews of interventions version 5.1. 0 (updated March 2011). The Cochrane Collaboration 2011.
21. Barth DD, Mayosi BM, Jabar A, Engel ME. Prevalence of group A streptococcal disease in North and Sub-Saharan Africa: a systematic review protocol. BMJ open 2015;**5**(8):e008646.
22. Werfalli M, Musekiwa A, Engel ME, Ross I, Kengne AP, Levitt NS. The prevalence of type 2 diabetes mellitus among older people in Africa: a systematic review study protocol. BMJ open 2014;**4**(6):e004747.
23. Pienaar E, Grobler L, Busgeeth K, Eisinga A, Siegfried NJHI, Journal L. Developing a geographic search filter to identify randomised controlled trials in Africa: finding the optimal balance between sensitivity and precision. 2011;**28**(3):210-215.
24. Beall B, Facklam R, Thompson T. Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. Journal of clinical microbiology 1996;**34**(4):953-958.
25. CDC. *Protocol for emm typing*. https://www.cdc.gov/streplab/groupa-strep/emm-typing-protocol.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fstreplab%2Fprotocol-emm-type.html (29/04/2020).
26. TWB. *The World Bank: World Bank Country and Lending Groups*. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (29/04/2020).
27. Hoy D, Brooks P, Woolf A, Blyth F, March L, Bain C, Baker P, Smith E, Buchbinder R. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. Journal of clinical epidemiology 2012;**65**(9):934-939.
28. CDC. *emm- Designations and Clusters*. https://www.cdc.gov/streplab/groupa-strep/resources.html (29/04/2020).
29. Nyaga VN, Arbyn M, Aerts MJAoPH. Metaprop: a Stata command to perform meta-analysis of binomial data. 2014;**72**(1):39.
30. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine 2009;**151**(4):264-269.
31. Abdissa A. Characterization of Group A Streptococci Isolated from Throat of Healthy School Children in Ethiopia. Addis Ababa University 2005.
32. Tewodros W, Kronvall G. M protein gene (emm type) analysis of group A beta-hemolytic streptococci from Ethiopia reveals unique patterns. Journal of clinical microbiology 2005;**43**(9):4369-4376.
33. Engel ME, Muhamed B, Whitelaw AC, Musvosvi M, Mayosi BM, Dale JB. Group A streptococcal emm type prevalence among symptomatic children in Cape Town and potential vaccine coverage. The Pediatric infectious disease journal 2014;**33**(2):208.
34. Barth D, Naicker P, Engel K, Muhamed B, Basera W, Mayosi B, Dale J, Engel M. Molecular Epidemiology of Noninvasive and Invasive Group A Streptococcal Infections in Cape Town. Msphere 2019;**4**(5).
35. Hraoui M, Boubaker IB-B, Doloy A, Samir E, Redjeb SB, Bouvet A. Epidemiological markers of Streptococcus pyogenes strains in Tunisia. Clinical microbiology and infection 2011;**17**(1):63-68.
36. Mzoughi R, Bouallegue O, Selmi H, Ben Said H, Essoussi A, Jeddi M. Group A streptococci in children with acute pharyngitis in Sousse, Tunisia. 2004.
37. Seale AC, Davies MR, Anampiu K, Morpeth SC, Nyongesa S, Mwarumba S, Smeesters PR, Efstratiou A, Karugutu R, Mturi N. Invasive group A streptococcus infection among children, rural Kenya. Emerging infectious diseases 2016;**22**(2):224.
38. Tapia MD, Sow SO, Tamboura B, Keita MM, Berthe A, Samake M, Nataro JP, Onwuchekwa UO, Penfound TA, Blackwelder W, Dale JB, Kotloff KL. Streptococcal pharyngitis in schoolchildren in Bamako, Mali. Pediatr Infect Dis J 2015;**34**(5):463-8.
39. Zühlke LJ, Beaton A, Engel ME, Hugo-Hamman CT, Karthikeyan G, Katzenellenbogen JM, Ntusi N, Ralph AP, Saxena A, Smeesters PR. Group A streptococcus, acute rheumatic fever and

rheumatic heart disease: epidemiology and clinical considerations. Current treatment options in cardiovascular medicine 2017;**19**(2):15.
40. Bessen DE, Carapetis JR, Beall B, Katz R, Hibble M, Currie BJ, Collingridge T, Izzo MW, Scaramuzzino DA, Sriprakash KS. Contrasting molecular epidemiology of group A streptococci causing tropical and nontropical infections of the skin and throat. The Journal of infectious diseases 2000;**182**(4):1109-1116.
41. Bowen A, Harris T, Holt D, Giffard P, Carapetis J, Campbell P, McVernon J, Tong S. Whole genome sequencing reveals extensive community-level transmission of group A Streptococcus in remote communities. Epidemiology & Infection 2016;**144**(9):1991-1998.
42. Sagar V, Kumar R, Ganguly N, Chakraborti A. Comparative analysis of emmtype pattern of Group A Streptococcus throat and skin isolates from India and their association with closely related SIC, a streptococcal virulence factor. BMC microbiology 2008;**8(1)**(150).
43. Dhanda V, Kumar R, Thakur J, Chakraborti A. emm type distribution pattern of group A streptococcus in North India: Need for a new preventive approach. The Indian journal of medical research 2010;**132**(6):741.
44. Arêas GP, Schuab RB, Neves FP, Barros RR. Antimicrobial susceptibility patterns, emm type distribution and genetic diversity of Streptococcus pyogenes recovered in Brazil. Memórias do Instituto Oswaldo Cruz 2014;**109**(7):935-939.
45. Smeesters PR, Laho D, Beall B, Steer AC, Van Beneden CA. Seasonal, Geographic, and Temporal Trends of emm Clusters Associated With Invasive Group A Streptococcal Infections in US Multistate Surveillance. Clin Infect Dis 2017;**64**(5):694-695.

Funding

This work is based on the research supported in part by the National Research Foundation of South Africa (grant number 116287). Any opinion, finding, conclusion, or recommendation expressed in this material is that of the authors, and the NRF does not accept any liability in this regard. M.E.E., T.S., K.E., B.M., and J.B.D. are supported by a grant (NW17SFRN33630027) from the American Heart Association. J. B. Dale received funds from the U.S. Public Health Service (NIH grants R01 AI-132117 and R01 AI-010085).

Chapter 3

Title: The longitudinal assessment of Group A Streptococcal *emm* types in Cape Town, South Africa.

Introduction

Group A Streptococcus (GAS) is a strict human pathogen that plays a role in the development of a variety of human infections and diseases such as pharyngitis (sore throat), scarlet fever, impetigo (skin infection), to more severe invasive diseases such as necrotizing fasciitis, Streptococcal toxic shock syndrome (STSS) and post-streptococcal autoimmune sequelae such as post-streptococcal glomerulonephritis (PGSN), acute rheumatic fever (ARF) and rheumatic heart disease (RHD) [1]. A Previous study in South Africa has reported children as the major reservoir for GAS pharyngitis with approximately ~20% incidence rate, with a peak age between 5 to 15 years [2]. GAS has an arsenal of virulence features capable of infecting the host and tissues, resulting in disease. The M protein is considered a significant feature due to its anti-phagocytotic capabilities, host mimicry and confers strain specificity [3].

The M protein, encoded by the *emm* gene, is an elongated α-helical coiled-coil structure that is anchored to the cell wall of the bacteria. The N-terminal region (or A repeat region) of the M protein which comprises around 50 or so amino acids differentiating to a total of over 230 *emm* types [4, 5]. M protein serotyping of GAS (referred to as *emm* typing of GAS) is considered by many as the gold standard typing scheme [5, 6] and thus, used in the current study. The characterization of prevalent *emm* types of GAS in a specific region forms the foundation of tracking the epidemiology, biology and genetic structure of the organism and their respective diseases [7]. A previous study highlighted the usefulness of longitudinal assessment of a specific population for accurate categorization of isolates from GAS infections, as well as or the variations that may occur in the population at both the group and individual level [8], Furthermore, longitudinal studies allow for the evaluation of the relationship between immune responses and development of infection, thus potentially providing further insight to the immunity of the human host, in order to understand the effects that may occur after a new acquisition of GAS.

Efforts are currently underway to develop a 30-multivalent M protein-based vaccine [9, 10]. This 30-valent vaccine has recently reported results obtained from phase I clinical trials which was well-tolerated by twenty-three Strep A-vaccinated participants with no clinical evidence of autoimmunity. Furthermore, it elicited significant increases in mean antibody levels to 24 of the 30 M peptide antigen components determined by ELISA assays (Pastural, 2019). Given the abundance of *emm* types against vaccine formulation, documenting the distribution of GAS *emm* types provides data which is likely to inform vaccine development in assessment of potential coverage. Our earlier systematic review demonstrated

the inadequate potential of the putative vaccine against isolates obtained from African participants, where protection was only afforded to 58.22% of isolates obtained from five countries across Africa; a region where the burden of GAS diseases are significant [11].

This report therefore aims to provide comprehensive, longitudinal data of children over a two-year period, describing the epidemiological features that exist within the population using molecular typing methods that will aid in vaccine development and essentially determine the efficacy in terms of putative coverage within the region.

Methods

Setting and participants

The longitudinal study design was applied to characterize the epidemiological features and serum immune responses of children (aged 5-17) coming from two juxtaposed peri-urban townships, namely Langa (pop=52,401) and Bonteheuwel (pop=45,967), collectively called Vanguard in Cape Town, South Africa. The communities are of lower socioeconomic status and comprise various racial and class groupings. A previous study in South Africa reported children as a major reservoir for GAS pharyngitis with an annual incidence of 0.837 cases/100 person-years (unpublished data) and prevalence of approximately 21% in children 5 to 15 years of age [2] in children presenting with sore throat at local clinics.

Recruitment and enrolment

The children were recruited in the study from a local clinic within each township, assessed from among individuals in the waiting area of the clinic site that had an initial sore throat. After informed consent had been taken from the parent/ legal guardian and assent from the child >7 years, a questionnaire was completed by the study nurse to capture the demographics. Physical examinations were conducted on the pharynx of each participant following the collection of samples, symptoms were recorded on a CRF (case report form) and documented in the AFRO*Strep* RedCap database. Treatment decisions were made by the clinician according to standard practice without input from the research team. Recruitment was conducted from April 2018 to March 2020; unfortunately, recruitment was halted due to the COVID pandemic.

Processing of samples

Throat culture specimens and blood samples were obtained from the enrolled participants, A follow-up throat swab and blood sample was requested at 2-monthly intervals. If a child experienced a sore throat episode during the follow-up period, samples were obtained.

Throat swabs were sent to the NHLS laboratories for standard serogrouping to determine the presence of β-haemolytic Streptococci. Positive-Streptococci pure cultures were sent to the AFRO*Strep* laboratory for characterizing the *emm* types. Cultures were stored at -80°C until *emm* typing procedure.

Definitions

Children were grouped into categories based on the longitudinal data obtained over the enrolled period. Carriers of GAS were those participants that had >1 GAS culture, with no corresponding symptoms. Recurrent episodes of GAS were the participants that had >1 GAS culture, regardless of *emm* type and symptoms. A participant with a single GAS-positive culture with subsequent negative cultures was classified into the single episode category. Participants having had no infections throughout the study were regarded as having no infections.

A GAS acquisition was described as the presence of GAS from the initial enrolment visit with a positive throat-culture. A new GAS acquisition was defined as a GAS-positive culture isolated from a participant following the enrolment visit, as it was not possible to determine the acquisition date of GAS, as described by Johnson, Kurlan [12]. If GAS was present in two sequential visits, a new GAS acquisition was considered if the *emm* types differed.

Treatment

The study nurse did not intervene or influence the administration of antibiotics. However, laboratory results stating whether GAS was present at the initial enrolment or either of the follow up visits were added into the participants' clinic folder. The clinic folders were reviewed to obtain the total number of participants that received antibiotics throughout the entire study, which may have an altered effect on the immune response. The antibiotics administered by the clinician were either 250mg or 500mg of penicillin V or amoxicillin depending on the severity of case symptoms.

Molecular investigations for characterization of GAS

The *emm* typing technique involves a number of steps. Briefly, the procedure begins with the culture and identification of GAS from the cryovial specimens received from the microbiology lab (NLHS). DNA is extracted from the isolates using the Wizard Genomic DNA Purification Kit (*Qiagen*) and using the NanoDrop 2000 (*ThermoFisher Scientific*) to quantify the DNA and determine the quality thereof.

The extracted DNA was subjected to PCR amplification of the *emm* gene, according to the CDC protocol [5, 13]. This molecular method selectively amplifies a region on the N-terminal of the M gene. The primers used in procedure anneal to a conserved sequence internal to the M gene. The primers used

were synthesized at the Department of Molecular and Cell Biology, University of Cape Town, South Africa (Table 1).

Table 1 – Primer sets used in *emm* typing procedure

Primer set	Forward/Reverse strand	Reference
Primer set I	5'-TAT TSG CTT AGA AAA TTA A-3' / 5'-GCA AGT TCT TCA GCT TGT TT-3'	Beall, Facklam [5,] CDC [13]
Primer set II	5'-TAT T(C/G) GCT TAG AAA ATT AA-3' / 5'-GCA AGT TCT TCA GCT TGT TT-3'	Beall, Facklam [5,] CDC [13]
Primer set III	5'- TAT TSG CTT AGA AAA TTA A-3' / 5'-TTCT TCA AGC TCTT TGTT-3'	Frost, Laho [14]

The PCR products were separated and visualized using agarose gel electrophoresis, purified using the MinElute PCR purification kit (Qiagen, Valencia, CA, USA) which incorporates spin-column technology using silica-gel membranes with selective binding properties to purify DNA directly from PCR products.

The PCR products were sent to the University of Stellenbosch, Central DNA Sequencing Facility for sequencing. The DNA sequences generated were submitted to the CDC *Streptococcal pyogenes* database [15] for comparison with existing sequences; thereafter the submitted sequence was assigned an *emm* type; e.g., *emm*2 and when necessary a subtype, *emm*2.1.

Result

Overall representation of participants

A total of 263 children were screened, of which 256 were successfully enrolled into the study over a 23-month period, with 91 withdrawals or lost to follow-up (LTFU). The study originally attempted a 2-monthly period between visits however, this proved difficult due to unavailability of parents bringing their children to the study site for sample collection, resulting in adherence issues. The average period between visits were ~89 days (min = 4, max = 617, CI 95%, 5.619). The average number of visits per patient was 3.4, when excluding those withdrawals or LTFU after the initial enrolment visit, there was an average of 4.5 visits per patient, with follow-ups ranging from 2 to 11 visits among participants over the period.

Overall representation of β-haemolytic streptococci (βHS)

The overall representation of samples are portrayed in Table 2, clearly showing an increase in children enrolled during the winter parts of the year, with more infections occurring. School holidays over the months of June also increased the chances of successful visits. Prevalence of ***β***HS and GAS among those seeking treatment for pharyngitis at the time of enrolment was 70/256 (27.3%) and 54/256 (21.1%), respectively. Overall, 124 ***β***HS isolates were recovered from enrolment and follow-up visits (GAS = 83, GCS = 25, GGS = 16). Sixty-eight of the GAS cultures were successfully emm typed, while the remaining isolates failed to yield sequence data (n=9) or a pure GAS culture from the NHLS was never received (n=2) or produced no growth on blood agar plates during culture (n=4). The sixty-eight successfully sequenced GAS were isolated from 60 participants, of which 46 were obtained upon entry to the study and 22 were new GAS acquisitions during the study.

Table 2 – Representation of samples collected across all participants

Year	No. of children enrolled	No. of throat cultures	No. of ***β***HS -positive throat cultures (%)			
			GAS	GCS	GGS	Total ***β***HS
April 2018 – Sep 2018 (Period I)	136	211	23 (10.9)	11 (5.2)	4 (1.9)	38 (18.1)
Oct 2018 – Mar 2019 (Period II)	32	152	15 (9.9)	3 (2.0)	4 (2.6)	22 (14.5)
April 2019 – Sep 2019 (Period III)	63	258	31 (12.0)	6 (2.3)	4 (1.6)	41 (15.9)
Oct 2019 – Mar 2020 (Period IV)	25	237	14 (5.9)	5 (2.1)	4 (1.7)	23 (9.7)
Total	256	858	83 (9.7)	25 (2.9)	16 (1.9)	124 (14.5)

Participants according to classification groups

The characterization of participants into categories were based on the infection outcome (enrolment infection, single episode infection, recurrent infections, carrier infections or no infections) throughout their time within the study (Table 3). Ninety-five participants had no infection, the most abundant category, which is surprising as they reside in a reservoir for GAS. This was followed by the enrolment infection category (n=39 ***β***HS, n=34 GAS), single episode infections (n=21 ***β***HS, n=11 GAS) while recurrent and carrier infections were the least. If the participant had an enrolment infection and later went on to obtain another infection, that participant was categorized as a recurrent infection (n=11 ***β***HS, n=7 GAS).

Interestingly, a participant in the recurrent infection category, went from being a carrier, showing no symptoms at visit 2 (*emm*82), to having GAS isolated at visit 4 (*emm*82) with symptoms, indicating the versatility of the bacteria and how it would affect the host.

There were three instances whereby a participant could have been categorized into more than one group or not categorized at all, these cases were: participant 158, had 2 sequential GAS-positive cultures, differing *emm* types (visit 1, *emm*80 & visit 2, *emm*49) with symptoms on both occasions (enrolment infection and/or recurrent infection group); participant 125 had 2 sequential GAS-positive cultures, initially with symptoms (visit 2, failed reaction) then moved into a carrier state with no symptoms (visit 3, *emm*12) and participant 161 had 2 sequential GAS-positive cultures, differing *emm* types (visit 3, *emm*49 & visit 4, *emm*223) both without symptoms (enrolment infection and/or carrier group). The remaining participants (n=85) only had 1 visit and were not categorized (Table 4).

Table 3 – The number of participants according to classification groups

Classification	Definition	No. of participants (***β***HS / GAS)
Enrolment infection	If positive throat culture obtained at the enrolment visit, and no infection for the remaining visits.	39 / 34
Single episode infection	If a single positive throat culture was preceded and followed by negative throat cultures during a single school year.	21 / 11
Recurrent infections	If >1 GAS culture was isolated, regardless of emm type and symptoms	11/ 7
Carriers	if GAS were recovered from 2 or more sequential surveillance cultures obtained in absence of symptoms.	5 / 3
No infection	Children whose throat cultures were negative for GAS throughout a single school year were designated as having no infections.	95

*85 participants had 1 visit and were excluded from categorization

Table 4. Interesting cases that could have been grouped in more than one category and those participants in the recurrent infection category with corresponding symptoms and antibiotics

PID	Initial infection				Follow-up infection		
	Visit	Symptoms	antibiotics	*emm* type	Visit	Symptoms	*emm* type
3	1	+	Y	*emm*22	3	+	*emm*1
23	1	+	Y	*emm*25	4	-	*emm*95
109	1	+	Y*	*emm*12	7	-	*emm*82
125	2	+	N	Failed rxn	3	-	*emm*12
139	1	+	Y#	Failed rxn	5	-	Failed rxn
142	2	-	N	*emm*82	4	+	*emm*82
158	1	+	Y	*emm*80	2	+	*emm*49
161	3	-	N	*emm*49	4	-	*emm*223
163	5	-	Y$	*emm*75	7	-	*emm*18
194	1	+	N	*emm*184	3	-	*emm89*

-, asymptomatic; +, symptomatic; Failed rxn, failed reaction (untypable), *, received antibiotics after initial infection and before follow-up infection, #, date of antibiotics received unknown, $, antibiotics received before initial infection

Among the 185 participants GAS-negative at enrolment, 9.7% (n=18) had positive GAS culture results at subsequent follow-up visits; of these 5 participants had more than two GAS positive results. Fifty-four participants were GAS-positive at the enrolment visit, of which, 11.1% (n=6) had GAS-positive results at subsequent visits, of which only two had symptoms when GAS was isolated on both occasions. Seventy-two participants provided a total of 83 positive GAS cultures. Of these, 26/83 (31.3%) were obtained post-enrolment, the majority (23/26, 88.5%) were isolated from asymptomatic participants.

Comparison between primer sets

The *emm* typing results were completed using three alternative primer sets (Table 1), to confirm the *emm* type sequence. We can confirm that the primer set developed by Frost, Davies [16] had the best success rate (71.1%) at generating appropriate sequence data in comparison to primer set 1 (53.0%) and 2 (19.3%) respectively.

Characterization of *emm* types and vaccine coverage

The characterization of *emm* types are portrayed in Table 5, with a diversity of 27 heterologous *emm* types obtained throughout the study period. The most abundant *emm* type was *emm*1 (n=10), followed by *emm*12 (n=7) and *emm*49 (n=6). When grouped according to *emm* clusters, cluster E3 (n=19) was the most predominant cluster followed by A-C3 (n=10), E6 and E4 (n=8), and A-C4 and D4 (n=7) isolates respectively (Figure 1). When considering the efficacy of the 30-multivalent vaccine in a high-risk population, 63.2% (n=43) of GAS isolates could theoretically be covered by the vaccine, which

would increase to 72.1% (n=49) considering previously published emm types that were cross-opsonized by 30-valent vaccine antisera [17] (Figure 1).

Table 5 – Characterization of *emm* types of GAS isolates over the study period

emm type	No. of isolates				
	Period I	Period II	Period III	Period IV	Total
1	4		4	2	10
2				1	1
6			1		1
8	2				2
12	1		6		7
18				1	1
22	1		1	1	3
25	1			1	2
44	1	1	1		3
49	2	1	3		6
58			2		2
70	2				2
75	1	1	1	1	4
80		2			2
82		1	2		3
87		1		1	2
89		1		1	2
92		1			1
94	3			1	4
95	1	2	1		4
116	1				1
184			1		1
192	1				1
223			1		1
231		1			1
244	1				1
Total	22	13	24	10	68

Period I: April 2018 – September 2018, Period II: October 2018 – March 2019, Period III: April 2019 – September 2019, Period IV: October 2019 – March 2020.

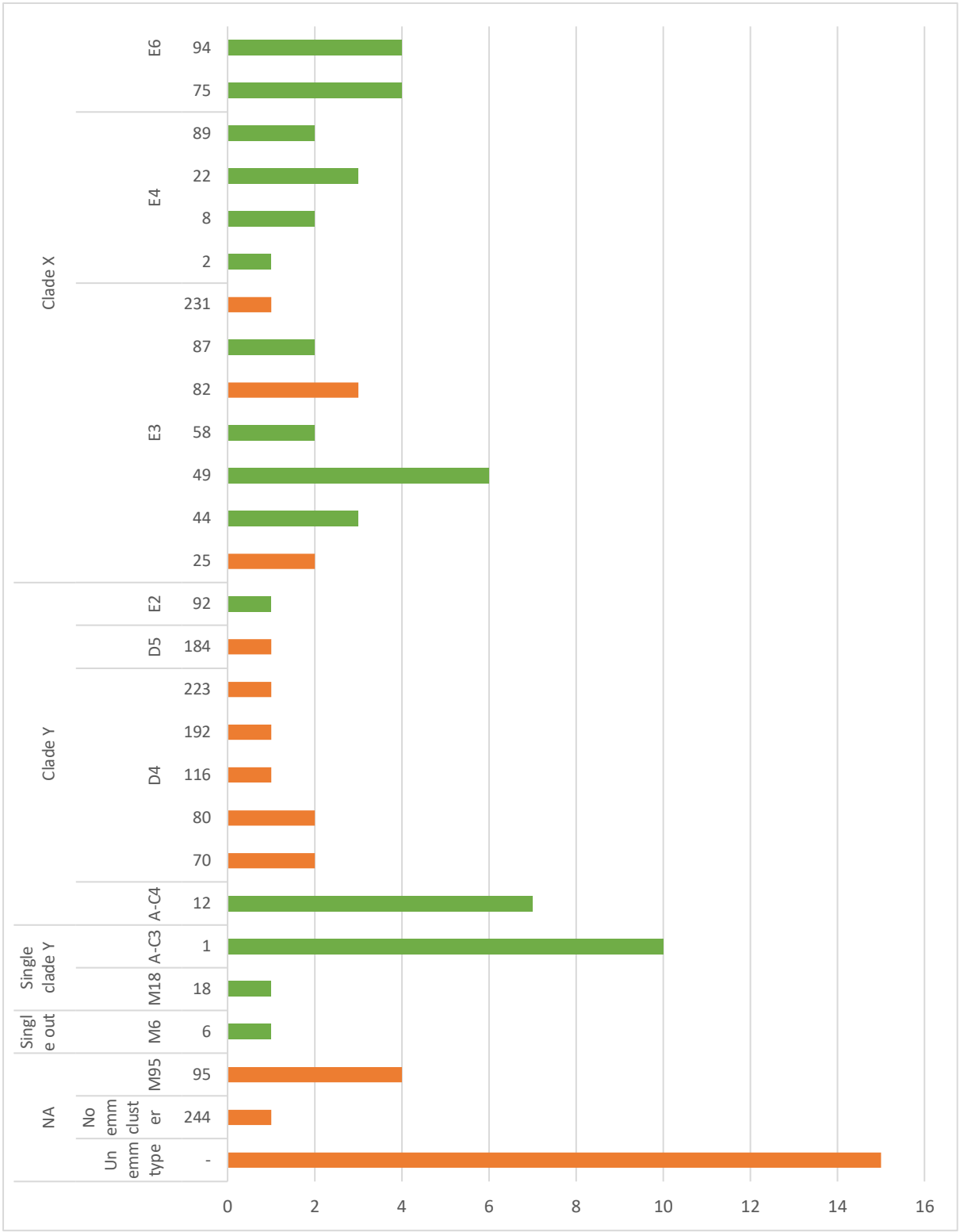


Figure 1 – Graphical representation of *emm* type distribution with corresponding *emm* cluster and clade. Green bars represent the *emm* types included in the 30-valent vaccine (n=43, 63.2% coverage), with cross-protection afforded to emm8 and emm94 (n=49, 72.1% coverage). Un *emm* type, unknown *emm* types that includes: untypable (no *emm* sequence) = 9, failed cultures = 4, missing sample = 2; No *emm* cluster, has not been classified according to Sanderson-Smith (2014)

181

Symptoms associated with GAS acquisitions

Of those GAS acquisitions with symptoms (n=57), the most common symptoms were tonsillar swelling (n=55), followed by hoarseness and anterior cervical node > 1.5cm in diameter (n=14) respectively and having a temperature above 38°C (n=13) (Table S1 – Appendix C), which is coherent with that described by the Centor criteria [18] and the clinical decision rule described by Engel, Cohen [19] in the same population.

Administration of antibiotics and corresponding symptoms of sore throat

Children were administered antibiotics according to the clinician at the clinic (penicillin V or amoxicillin, 250mg/500mg) depending on symptom severity (Table S1 – Appendix C). Sixty-nine participants received antibiotics throughout the study period, for seven of these, the dates were unknown. Of those prescribed antibiotics, 34.8% (n=24) were GAS-positive at the time of administration. Antibiotics were administered on the enrolment visit for 46 of the 69 cases, as most participants entered the study having symptoms of sore throat. Eight participants were given antibiotics despite having no symptoms of sore throat however, seven of these participants received antibiotics in between their scheduled visits and physical examination and documentation of symptoms were not conducted by the study nurse. The most prominent symptoms of participants that received antibiotics regardless of whether GAS was cultured or not, were tonsillar swelling (n=43), followed by hoarseness (n=13) and a cough and tonsillar erythema with (n=11) respectively. Fourteen participants were administered antibiotics, with having only one of the symptoms, 13 of which were tonsillar swelling and one case had hoarseness. Of participants that received antibiotics (n=69), 16 got infected with GAS, 13 infected with other $\boldsymbol{\beta}$HS, 5 participants were infected with both, GAS and other $\boldsymbol{\beta}$HS. For those participants that maintained the 2-monthly visitations, we calculated the average number of days from receiving antibiotics to infection with GAS/$\boldsymbol{\beta}$HS, were 157.1 days (min=51, max=546 days) and an average of 109.3 days (min=63, max=186 days), between the first infection and the second.

Discussion

This study documents the longitudinal characterization of new GAS acquisitions and corresponding *emm* types over a period of 2 years within a high-risk population aged, 5-15 years. We can conclude that there has been a shift in the molecular epidemiology of GAS in the population compared to a previous study. We identified 83 GAS isolates (124 $\boldsymbol{\beta}$HS) throughout the study period, of which, 69 were sequenced. The most prevalent GAS *emm* types were *emm*1 (n=10), followed by *emm*12 (n=7) and *emm*49 (n=6). In terms of *emm* clusters, E3 was the most prevalent cluster, followed by A-C3 and E6. In terms of the 30-valent, StrepAnova vaccine, it would provide protective coverage to 72.1% of

isolates obtained in this region. Additionally, we evaluated the effectiveness of different primer sets used in the *emm* typing procedure and provided information regarding the symptoms of sore throat and administration of antibiotics.

In a study conducted in the same population a decade previously [20], the more prevalent *emm* types included *emm*48 (n=17); *emm*12 (n=13); *emm*4 (n=11); *emm*89 (n=9) and *emm*94 (n=9). In this study, only *emm*12 ranks among the most prevalent *emm*-types isolated. Compared with the distribution of emm types within Africa, we affirm clear differences as compared with those reported in Chapter 2 (Study I) [11], where *emm*18, *emm*65, *emm*75, *emm*76 and *emm*81 are among top *emm* types reported. The changing molecular epidemiology of GAS infections is consistent with previous studies that showed *emm* type replacements withing large [21] and smaller [22] geographic regions.

When considering the 30-valent vaccine coverage in the current population, 72.1% of *emm* types would be afforded protection. However, compared with the overall data from Africa, the multi-valent vaccine only covers 58.22% of isolates including those afforded cross-protection [11]. This study along with the review provide evidence that the 30-multivalent, StrepAnova vaccine, would be hampered due to its inability to protect against the sheer number of *emm* types and the ever-changing prevalence of different *emm* types per region, especially in areas where the burden of GAS infections and GAS sequalae diseases are rampant. If vaccine development strategies had to consider the *emm* clustering framework, as hypothesized in Salie, Engel [11], which states that if an *emm* type is currently within the 30-valent formulation, it would extend coverage to all *emm* types within the *emm* cluster it belongs to. If this were the case, the hypothesized *emm* cluster vaccine formulation would increase coverage to 85.3% in the current study and similarly in the overall African data (80.3%).

All participants that received antibiotics on the enrolment visit showed no sign of symptoms on their next follow-up visit, which holds true to the potential of antibiotics in the eradication of symptoms of sore throat within 1-2 days [23, 24]. Furthermore, based on the clinical decision rules applied by the clinicians, 45 participants received antibiotics on the enrolment visit; however, only 24 were correctly prescribed treatment based on the acquisition of GAS.

In comparison to Hysmith (2017) and Quinn, Vander Zwaag [25], identified asymptomatic rates of 65% and 68% respectively. Another longitudinal study completed by Martin, Green [8] had an asymptomatic rate of 75%. This further indicates that most GAS infections are asymptomatic, like most other pathogenic infections. Furthermore, Martin et al. (2004), states that the definition of a carrier is a participant that had more than 2 sequential positive GAS cultures obtained more than 1 week apart in the absence of symptoms. Using this definition in the current study, we would have 5 carriers of βHS, 3 of which would be considered GAS carriers specifically. If we were to change this definition, based

on the period between visits (2 months) and possibly longer due to the inconsistencies of the participants, the new definition of a potential carrier would be that of a single episode of GAS with no symptoms; resulting in 26 potential GAS carriers. It is also important to keep in mind that the recovery of GAS in the throat of participants does not necessarily mean that GAS is responsible for the symptoms of sore throat. Thus, we require extra information with regards to the immune response of the host, to be able to discriminate and confirm an active GAS infection, a harmless carrier state and/or an underlying asymptomatic GAS infection. It will be interesting to see the immune response in these participants, specifically the GAS carriers to determine the antibody responses to GAS antigens and whether there is an underlying response.

One of the main limitations of the longitudinal study design was regarding the participants and their involvement to maintain and uphold the 2-monthly follow up visitations across the 2 years period This proved to be difficult, as 91 participants withdrew or were LTFU. Furthermore, the current COVID-19 pandemic, caused a stop to all recruitment and follow-up activities thus, we did not meet the overall study objective of obtaining a total of 300 participants and seeing each participant for 12 visits. Apart from this aspect of the study, another limitation was due to the unsuccessful sequence data and failed reactions, reducing the overall prevalence estimates of *emm* types recovered. Six were unviable samples altogether and were not submitted to PCR *emm* typing whereas, nine samples failed to render an *emm* type following sequencing.

Conclusion

This study provides evidence of the longitudinal assessment of participants and their associated GAS infections, determining the prevalence of dominant *emm* types, the overall efficacy of the 30-valent vaccine, the symptoms associated with infection as well as the intake of antibiotics. The data obtained in this study should be viewed in conjunction with the characterization of the host immune response to shed light on confirming an underlying immune response that potentially results in the post-sequelae diseases such as ARF and RHD.

References

1. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, Sriprakash KS, Sanderson-Smith ML, Nizet V. Disease manifestations and pathogenic mechanisms of Group A Streptococcus. Clin Microbiol Rev 2014;**27**(2):264-301.
2. Engel ME, Muhamed B, Whitelaw AC, Musvosvi M, Mayosi BM, Dale JB. Group A streptococcal emm type prevalence among symptomatic children in Cape Town and potential vaccine coverage. The Pediatric infectious disease journal 2014;**33**(2):208.
3. Fischetti VA. Streptococcal M protein: molecular design and biological behavior. Clinical microbiology reviews 1989;**2**(3):285-314.
4. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. The Lancet infectious diseases 2009;**9**(10):611-616.
5. Beall B, Facklam R, Thompson T. Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. Journal of clinical microbiology 1996;**34**(4):953-958.
6. Neal S, Beall B, Ekelund K, Henriques-Normark B, Jasir A, Johnson D, Kaplan E, Lovgren M, Reinert RR, Efstratiou A. International quality assurance study for characterization of Streptococcus pyogenes. Journal of clinical microbiology 2007;**45**(4):1175-1179.
7. Efstratiou A, Lamagni T. Epidemiology of Streptococcus pyogenes--Streptococcus pyogenes: Basic Biology to Clinical Manifestations. 2016.
8. Martin JM, Green M, Barbadora KA, Wald ER. Group A streptococci among school-aged children: clinical characteristics and the carrier state. Pediatrics 2004;**114**(5):1212-9.
9. Pastural E, McNeil SA, MacKinnon-Cameron D, Ye L, Langley JM, Stewart R, Martin LH, Hurley GJ, Salehi S, Penfound TA, Halperin S, Dale JB. Safety and immunogenicity of a 30-valent M protein-based group a streptococcal vaccine in healthy adult volunteers: A randomized, controlled phase I study. Vaccine 2020;**38**(6):1384-1392.
10. Dale JB, Penfound TA, Tamboura B, Sow SO, Nataro JP, Tapia M, Kotloff KL. Potential coverage of a multivalent M protein-based group A streptococcal vaccine. Vaccine 2013;**31**(12):1576-81.
11. Salie T, Engel K, Moloi A, Muhamed B, Dale JB, Engel ME. Systematic Review and Meta-analysis of the Prevalence of Group A Streptococcal emm Clusters in Africa To Inform Vaccine Development. Msphere 2020;**5**(4):e00429-20.
12. Johnson DR, Kurlan R, Leckman J, Kaplan EL. The human immune response to streptococcal extracellular antigens: clinical, diagnostic, and potential pathogenetic implications. Clin Infect Dis 2010;**50**(4):481-90.
13. CDC. *Protocol for emm typing*. https://www.cdc.gov/streplab/groupa-strep/emm-typing-protocol.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fstreplab%2Fprotocol-emm-type.html (29/04/2020).
14. Frost HR, Laho D, Sanderson-Smith ML, Licciardi P, Donath S, Curtis N, Kado J, Dale JB, Steer AC, Smeesters PR. Immune cross-Opsonization within emm clusters following group A Streptococcus skin infection: broadening the scope of type-specific immunity. Clinical infectious diseases 2017;**65**(9):1523-1531.
15. CDC. *emm- Designations and Clusters*. https://www.cdc.gov/streplab/groupa-strep/resources.html (29/04/2020).
16. Frost H, Davies MR, Velusamy S, Delforge V, Erhart A, Darboe S, Steer A, Walker MJ, Beall B, Botteaux A. Updated emm-typing protocol for Streptococcus pyogenes. Clinical Microbiology and Infection 2020;**26**(7):946. e5-946. e8.
17. Dale JB, Penfound TA, Chiang EY, Walton WJ. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine 2011;**29**:8175-8.
18. Centor RM, Witherspoon JM, Dalton HP, Brody CE, Link K. The diagnosis of strep throat in adults in the emergency room. Medical Decision Making 1981;**1**(3):239-246.

19. Engel ME, Cohen K, Gounden R, Kengne AP, Barth DD, Whitelaw AC, Francis V, Badri M, Stewart A, Dale JB. The Cape Town clinical decision rule for streptococcal pharyngitis in children. The Pediatric infectious disease journal 2017;**36**(3):250.
20. Muhamed B. The molecular characterisation of Group A Streptococcus among children with phyryngitis in the Vanguard community (Bonteheuwel). University of Cape Town 2012.
21. Shulman ST, Tanz RR, Dale JB, Beall B, Kabat W, Kabat K, Cederlund E, Patel D, Rippe J, Li Z. Seven-year surveillance of North American pediatric group A streptococcal pharyngitis isolates. Clinical infectious diseases 2009;**49**(1):78-84.
22. Kaplan EL, Wotton JT, Johnson DR. Dynamic epidemiology of group A streptococcal serotypes associated with pharyngitis. The Lancet 2001;**358**(9290):1334-1337.
23. Del Mar C, Glasziou PP, Spinks A. Antibiotics for sore throat. Cochrane Database of Systematic Reviews 2004(2).
24. Zwart S, Sachs AP, Ruijs GJ, Gubbels JW, Hoes AW, de Melker RA. Penicillin for acute sore throat: randomised double blind trial of seven days versus three days treatment or placebo in adults. Bmj 2000;**320**(7228):150-154.
25. Quinn RW, Vander Zwaag R, Lowry PN. Acquisition of group A streptococcal M protein antibodies. Pediatric infectious disease 1985;**4**(4):374-378.

Chapter 4

Title: Utility of Human Immune Responses to GAS Antigens as a Diagnostic Indicator for ARF: Systematic Review

M. Taariq Salie[1], Kimona Rampersadh[1], Babu Muhamed[1,2], Kélin C. Engel[1], Liesl J. Zühlke[3], James B. Dale[4] and Mark E. Engel[1*].

Affiliations

[1] Department of Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

[2] Division of Cardiology, Children's National Health System, Washington, DC, United States

[3] Children's Heart Disease Research Unit, Department of Paediatrics, University of Cape Town, Cape Town, South Africa

[4] Division of Infectious Diseases, University of Tennessee Health Science Center (UTHSC), Memphis, TN, United States.

Abstract

Background: Previous studies have established that streptococcal antibody titer is correlated with a diagnosis of acute rheumatic fever (ARF). However, results vary in the usefulness of GAS antibodies, particularly anti-streptolysin-O (ASO) and anti-DNase B, in confirming a recent GAS infection. Therefore, we sought to provide, from published studies, an evidence-based synthesis of the correlation of streptococcal serology to establish the usefulness of immunological data in aiding the diagnosis of ARF. These findings are anticipated to have implications where echocardiography is not freely available, especially where ARF is rampant.

Methods: We conducted a comprehensive search across a number of databases. Applying a priori criteria, we selected articles reporting on studies, regardless of study design, that evaluate the levels of antibodies against GAS-specific antigens in ARF subjects against control values or a published standard. Data were extracted onto data extraction forms, captured electronically, and analyzed using Stata software. Risk of bias was assessed in included studies using the Newcastle-Ottawa Scale (NOS).

Results and Conclusion: The search strategy yielded 534 studies, from which 24 met the inclusion criteria, reporting on evaluation of titres for SLO (n = 10), DNase B (n = 9), anti-streptokinase (ASK) (n = 3) amongst others. Elevation in titres was determined by comparison with controls and upper limit of normal (ULN) antibody values as determined in healthy individuals. Meta-analysis of case-controlled studies revealed moderate odds ratio (OR) correlations between ARF diagnosis and elevated titres for SLO (OR = 10.57; 95% CI, 3.36–33.29; 10 studies) and DNase B (OR = 6.97; 95% CI, 2.99–16.27; 7 studies). While providing support for incorporating SLO and DNase B in the diagnosis of ARF, we present the following reflections: an elevation in SLO and DNase B levels are not consistently associated with an ARF diagnosis; increasing the number of GAS proteins in the test is warranted to improve sensitivity; paired (acute and convalescent) samples could provide a more accurate indication of a rising titre. Use of community-based controls as a standard is not a reliable marker by which to gauge recent GAS infection.

Keywords: GAS antigens, anti-streptolysin-O, anti-DNase B, ARF diagnosis, systematic review

Introduction

Acute rheumatic fever (ARF), which develops within 2-6 weeks after a preceding non-invasive group A streptococcal (GAS) infection, such as streptococcal pharyngitis or scarlet fever, is not treated correctly, affects ~300,000 - 500,000 people across the globe each year, the majority of whom live in developing countries. ARF symptoms include fever, arthritis, carditis, rash (erythema marginatum), subcutaneous nodules, and/or Sydenham's chorea [2-4]. Since these symptoms are related to other diseases, the Jones criteria has been used since 1944 as a clinical standard in the diagnosis of ARF and rheumatic heart disease (RHD); amongst its criteria, is laboratory evidence of recent streptococcal infection, either through culture or an elevation in serum streptococcal antibodies [3]. An accurate diagnosis of ARF ensures proper treatment and reduces the risk of recurrent disease and the development of rheumatic heart disease [5]. ARF is often underdiagnosed, as much as by 50%, as reported in a Fijian hospital-based study which compared clinical data against primary care records from health care clinics [6]. Okello similarly reports a likelihood of under-representation of the actual number of cases presenting to primary care in Uganda [7], thus highlighting the need to develop simple and practical approaches to diagnosing ARF in primary care in low-resource settings.

The Jones criteria [8] has recently been updated by the American Heart Association (AHA) to suit all types of populations with appropriate recommendations [3, 9]. With recent advancements in medical technologies such as 2D echocardiography and Doppler flow assessments, analysing images of the heart valves has been made easier, with clinicians following the guidelines as described in the Jones criteria [3]. However, these tools may not be readily available in areas where ARF is rampant. The Jones criteria has guidelines as to assess the remaining major clinical manifestations which, however, may be difficult to evaluate as these symptoms could be mistaken for or attributed to other illnesses. Symptoms of arthritis may be found in septic arthritis, juvenile idiopathic arthritis, lyme disease, or sickle cell anemia [10]. Carditis may be found in mitral valve prolapse, fibroelastoma, cardiomyopathy or Kawasaki disease. Symptoms associated with chorea may be found in Wilson disease, tic disorder, encephalitis, or other autoimmune diseases such as systemic lupus erythematosus and systemic vasculitis. Therefore, included in the Jones criteria are other mechanisms, such as evidence of a preceding GAS infection to eliminate any doubt in the diagnosis of ARF [10].

The Jones criteria describes evidence of a preceding GAS infection with any one of three circumstances; a positive throat culture for GAS, a positive GAS carbohydrate (GAC) antigen test, or a rise in GAS antibody titres (anti-streptolysin O (ASO) or anti-DNase B) that requires paired samples (at diagnosis and 3 weeks post suspected diagnosis) [10]. These GAS-specific antigens form part of the arsenal of GAS for survival and infiltration of the bacteria into human tissues thus, these antibodies are present in the sera of GAS-infected individuals [11-13]. Streptolysin O (SLO) is a cytolytic toxin released by GAS for cell lysis [14, 15]. DNase B is an extracellular virulent protein with DNA-degrading activity [16, 17]. These two virulent factors, along with GAC, are mentioned within the Jones criteria; however, there are many other GAS-specific antigens in recent times that have shown potential as putative biomarkers relating to the presence of GAS [18-25].

Few countries have specific guidelines in terms of an ARF diagnosis, specifically including a preceding GAS infection (Table S1 – Appendix D). We sought, through a comprehensive systematic review of published studies, to conduct an evidence-based synthesis of the utility of streptococcal serology in aiding the diagnosis of ARF. Primarily, our review aimed to assess available published literature regarding the association of antibodies against GAS antigens with ARF. We anticipate that our findings could have implications for the design of future diagnostic tests to confirm recent GAS infection in suspected ARF patients.

Methods

Search strategy and selection criteria

According to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [26], we performed a systematic literature review from two peer-reviewed databases (PubMed and Scopus) with predefined search terms (Table S2 – Appendix D). This review asks the following question: What is the utility of GAS serology, specifically SLO and/or DNase B, in providing evidence of a recent GAS infection in diagnosing ARF? In addition, we further sought to explore the potential of other GAS-specific antigens which may provide additional support of a recent GAS infection. The search strategy incorporated both free term text and Medical Subject Headings (MeSH) adapted to suit the particular database. Keywords incorporated a combination of terms relating to group A streptococcus, GAS antigens, SLO, DNase B, serology, immune response and acute rheumatic fever. Results were complemented by hand searching and citation searches in Google Scholar. The search was not restricted to publication dates or language. Additionally, grey literature including theses and conference proceedings, were also considered for inclusion.

Studies were included if immunological assays were used to evaluate the expression of antibodies evoked by GAS-specific antigens in ARF cases and controls within the same population or the use of a control standard based on the titres of healthy individuals (upper limit of normal, ULN) from the same region. Cases needed to be clinically diagnosed as ARF (peer-reviewed guidelines were not a prerequisite), while controls were documented as those having no history of ARF or RHD. In addition, longitudinal studies evaluating immune responses at more than one time point following new GAS acquisitions were also included. Case reports, narrative reviews, opinion pieces and publications lacking expression data, or referenced methodology and/or accepted guidelines, were excluded from the review. Duplicated studies of datasets and participants were removed, with the final, most recent, publication of the data assessed for inclusion.

Two reviewers independently applied the search strategy to the relevant databases. Articles were managed using the Rayyan QCRI web/mobile application [27]. Titles and abstracts were evaluated to exclude studies that did not describe the expression of GAS-specific antibodies. Thereafter, full texts

of the included titles and abstracts were retrieved and further evaluated against the inclusion criterion (Table S3 – Appendix D). Rayyan QCRI has a built-in "blind" filter function which prevented reviewers from observing the other's judgements. Discrepancies were resolved through discussion, involving an arbitrator (third reviewer) where necessary.

Two reviewers extracted data using a standardized data extraction form and any contradictions were solved through discussion with a third reviewer. Search results from the databases listed above, published and unpublished studies were managed with Endnote X9 referencing software. Briefly, data extraction consisted of recording the study demographics (number of study participants, geographical region), diagnostic measures, GAS-specific antigens and relevant antibody titre measurements describing elevation.

The risk of bias assessment tool (Table S4 – Appendix D) established by Wells [28] was adapted in questions specific for use in this review. Using the Newcastle-Ottawa Quality Assessment Scale (NOS) for case-control studies, which were characterized as being of a low or high risk of bias. A study with a low risk of bias is considered to be of high-quality and a low-quality study, with a higher risk of bias. Risk of bias was incorporated into the evaluation of heterogeneity in the pooled analyses.

Data analysis

Odds ratio estimates together with their 95% confidence intervals (CIs) were calculated to represent the association between GAS-specific antigens and ARF. The Mantel-Haenszel method was used to pool together odds data from individual studies [29]. Variability between studies was evaluated both visually by assessing forest plots and formally by the heterogeneity tests using $\chi 2$-based Q and I^2 statistic [30]. As expected, the studies varied in the constitution of participants and in the types of assays conducted; thus, a random-effects model was used for analysis [31].

We conducted statistical data analyses using Stata version 14.1 (StataCorp, College Station, TX, USA) to estimate the combined effect size (odds ratio and 95% CIs) between GAS antigens and ARF and to generate comparative effect forest plots. Studies were analysed in subgroups based on the inclusion of an accepted guideline at the time of diagnosis of ARF cases. Where a meta-analysis was not feasible,

either because data were too heterogeneous or insufficient to allow for meaningful pooling, we compiled a narrative report of the results. Antigens, for which only a single study was available, thus precluding conducting a meta-analysis, were presented as an odds ratio with its 95% CI. We utilized the respective authors' definition of ULN in defining a rising titre.

Results

Our search strategy yielded 534 articles which reduced to 479 after excluding duplicates (Figure 1). An additional five studies were included through citation searching, thus leaving 484 for consideration for this review. Following screening of titles and abstracts, 53 articles were deemed potentially relevant and available for full-text evaluation. Twenty-four studies met with the inclusion criteria, of which 14 were amenable to metanalysis. Table 1 shows the characteristics of the included studies. The included articles were published between 1955 and 2020 with sample sizes (cases and controls) ranging from 43 to 2,118 enrolled participants. Studies were conducted in local and university hospitals, clinics, outpatient departments, and schools situated in the study areas. Studies were conducted in the USA (n=7), Japan (n=5), India (n=4), Egypt (n=3), with one article from each of Pakistan, Trinidad, Madagascar, Ethiopia, UK, Australia, New Zealand. Participants ranged in age from 1 to 89 years. All the articles narrowed down their target population to a specific age group, mainly that of children. Only one article [32] made an effort to obtain participants from any age group so as to reflect the national population. A list of the excluded studies with reasons are detailed in Table S5 (Appendix D).

The overall quality of studies was moderate, with twelve studies deemed as having a low risk of bias (ie, a high NOS score; Table S4 – Appendix D). The included studies clearly described the phenotypes of patients, providing an acceptable case definition and guideline or diagnostic algorithm. Studies diagnosed ARF according to the Jones criteria, with controls examined as having no prior history of ARF or RHD from the same population and age-matched to the cases. Seven of the studies classed as of high risk of bias (NOS < 5) were completed before 1990 with authors failing to clearly define the

controls or cases with appropriate diagnostic guidelines [33-39]. One study, Zainab, Saleem [40] recruited no controls but instead used an ULN cut-off published previously from within the same region.

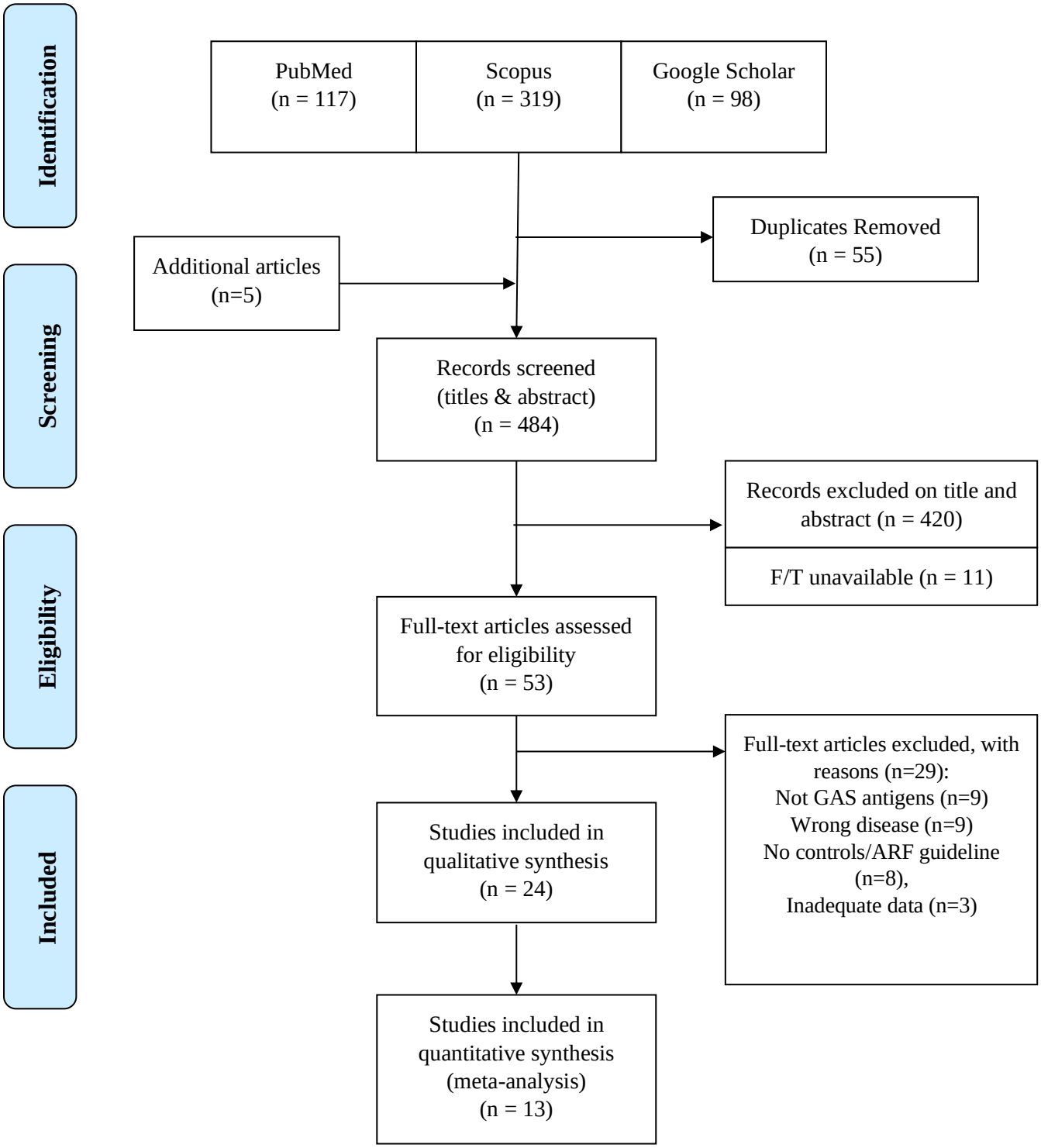


Figure 1 – Schematic PRISMA flow diagram of the literature search. Figure is modelled after Moher, Liberati [1]

Table 1 – Characteristics of included studies

Study ID	Country	Setting	Diagnostic guideline	Antigen / Detection method	Participants	Age
[a]Ayoub, Nelson [41]	Grenada / USA	ND	Jones criteria	SLO, DNase B & GAC - assays described in previous publication (not available)	Grenada: RF (n=32), controls (n=30), Florida: RF (n=32), controls (n=32)	5-32yrs
[b]Das, Dileepan [42]	USA / India	ND	NCS	DNase B – ELISA vs DNA methyl green micromethod	ARF (20), controls (n=20)	ND
[a]Fujikawa and Ohkuni [34]	Japan	ND	RF & RHD Guideline of Japanese Circulation Society	SLO – streptozyme test, DNase B – hemoprobe B test, GAC – ASP kit, SE – enzyme antibody-antigen reaction	ARF (n=8), controls (n=354)	6-15yrs
[a]Fujikawa and Okuni [35]	Japan	ND	F & RHD Guideline of Japanese Circulation Society	SLO, DNase B & SK – multiple enzyme test (streptozyme test)	ARF (n=21), controls (n=178)	6-15yrs
[a]Fujikawa, Kawakita [33]	Japan	ND	F & RHD Guideline of Japanese Circulation Society	SLO – described previously, DNase B – hemoprobe B test	RF (n=46), controls (n=278)	3 age groups: <6yrs, 6-16yrs & >16yrs
[b]Gomaa, Ali [43]	Egypt	Outpatient RF Clinic	Jones criteria	SLO – turbidimetric immunoassay & ELISA	ARF (n=80), controls (n=80)	ARF – 14.5yrs (mean), control – 15.2yrs (mean)
[a]Halbert, Swick [37]	USA	ND	NCS	SLO – agar precipitin technique	RF (n=33), non-RF (n=35)	ND
[a]Hanson-Manful, Whitcombe [9]	New Zealand	Hospitals	Jones criteria	SLO – turbidimetric technique using SLO kit, DNase B – enzyme inhibition assay, SpnA – bead-based immunoassay	ARF (n=16), controls (n=36)	ARF – 10.6yrs (mean), Controls – 6yrs (mean)
[a]Hokonohara, Yoshinaga [36]	Japan	ND	NCS	SLO – described by other author DNase B – hemoprobe B test, GAC – hemmaglutination method	RF (n=28), controls (n=NCS)	5-16yrs

[d]Hysmith, Kaplan [44]	USA	University associated clinics	-	SLO, DNase B, SCPA, Mrp, J14, SpyCEP, SSE, SOF, SpyAD & FBP54 – ELISA	PIDs (n=41)	6-15yrs
[d]Johnson, Kurlan [13]	USA	University associated clinics	-	SLO & DNase B – ELISA	PIDs (n=160)	6-15yrs
[a]Julie, Arivelo [32]	Madagascar	Hospital	NCS	SLO – latex agglutination technique	ARF (n=1690), control (n=428)	1-89yrs
[d]Kaplan, Top Jr [45]	USA	NCS	-	SLO, DNase B & NADase – assays described in previous publication (not available)	PIDs (n=49)	3-6yrs
[a]Kawakita, Takeuchi [39]	Japan	Elementary school	NCS	SLO – spectrophotometric method, DNase B – micro method, NADase – reduction by alcohol dehydrogenase	ARF (n=3), controls (n=361)	6-11yrs
[a]Kotby, Habeeb [46]	Egypt	Hospital	Jones criteria	SLO – rapid latex agglutination	ARF (n=60), controls (n=200)	3 age groups: <6yrs, 6-10yrs & >10yrs
[b]Read, Reid [47]	Trinidad	Hospital	Jones criteria	SLO – antibody titre kit	RF (n=44), controls (n=34)	ND
[b]Read, Fischetti [48]	USA	Hospital	Rheumatic Fever Service of The Rockefeller University Hospital	SLO – in vitro cellular migration of white blood cells	RF (n=NCS), controls (n=NCS)	ND
[c]Sagar, Bergmann [49]	India	ND	Jones criteria	SCI, SCPA & PSA – ELISA	RF (n=24), controls (n=25)	ND
[a]Saini, Kumar [50]	India	Hospital	Jones criteria	SLO – NCS	ARF (n=26), controls (n=84)	5-15yrs
[d]Shet, Kaplan [51]	USA	NCS	-	SLO, DNase B & SCPA – ELISA	PIDs (n=202)	2-12yrs

[b]Tewodros, Norgren [52]	Ethiopia	ND	NCS	SK – ELISA	ARF (n=11), controls (n=10)	3-12yrs
[b]Thakur and Prakash [53]	India	ND	NCS	GAC – ELISA	ARF (n=50), controls (n=50)	ND
[a]Widdowson, Maxted [38]	UK	Outpatient clinic	NCS	SLO – spectrophotometric method, DNase B – micro method	RF group (n=6), controls (n=44)	16-18yrs
[a]Zainab, Saleem [40]	Pakistan	Hospital	Jones criteria	SLO – kit human tex ASOT	ARF (n=50) (Historic control values)	5-15yrs

ND, no data; NCS, not clearly stated; RF, rheumatic fever; PIDs, participants
SLO, streptolysin O; GAC, group A carbohydrate; SE, streptococcal esterase; SK, streptokinase; SpnA, GAS nuclease A; SCPA, C5a peptidase; Mrp, M-related peptides; J14, C-repeat M peptide; SpyCEP, serine protease; SSE, serine esterase; SOF, serum opacity factor; SpyAD, GAS adhesion and division protein; FBP54, fibronectin-binding protein; SCI, collagen-like surface protein; PSA, putative surface antigen
[a], meta-analysis; [b], mean titre data; [c], single article-antigen; [d], longitudinal data

174

Association of SLO Antibody with ARF

Ten studies (11 populations; controls, n=1972; cases, n=1947) providing data on Anti-SLO titres were amenable to meta-analysis (Figure 2).

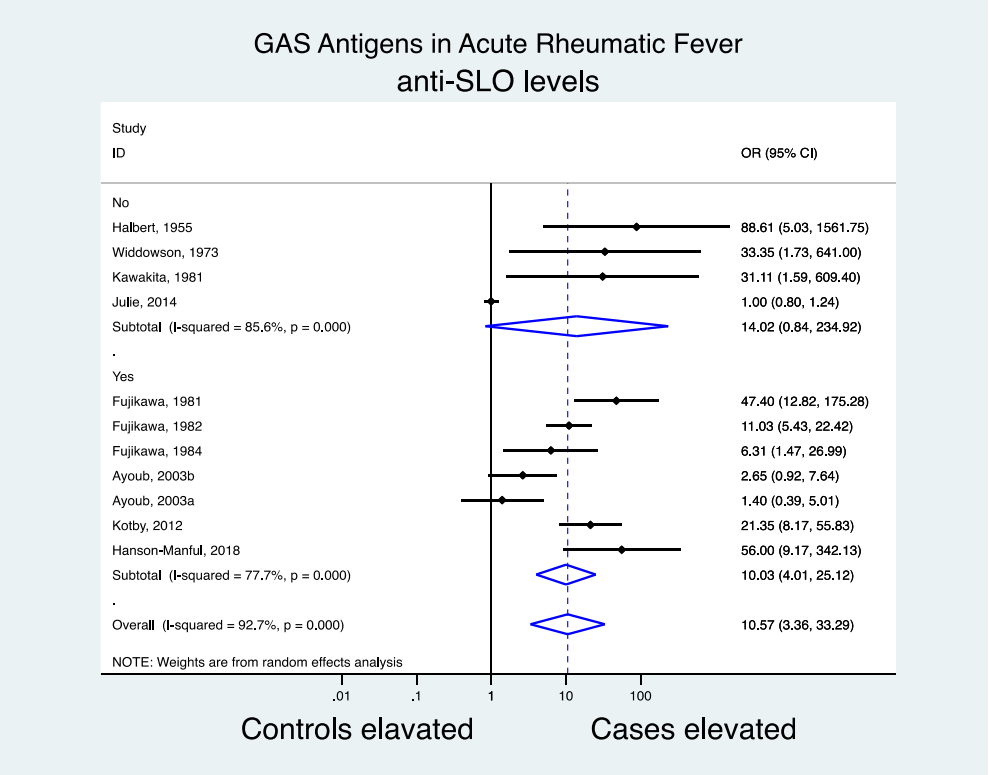


Figure 2 – Forest plot evaluating the odds of association between Anti-SLO titres and ARF. Subgrouping based on whether a guideline was used to diagnose ARF

Overall, ARF cases showed a greater association of an SLO antibody rise in comparison to controls (OR, 10.57 (95% CI [3.36,33.29]; (I^2, 92.7%)). In a subgroup analysis according to whether a guideline for the diagnosis of ARF was used or not, the association was not found to be statistically significant among studies not utilizing a guideline. Amongst studies not included in the meta-analysis due to the absence of a control group, each report a rise in anti-SLO titre: Zainab (2020) in 24 of 50 cases (48%) of ARF diagnosed according to the Jones criteria (published standard = 200IU/ml), Saini (2019) in 15 of 26 (58%) cases of ARF as per the Jones criteria guideline, (published standard = 262IU/ml), Hokonohara (1987) in 24 of 67 cases of ARF (36%) against published standard = 240IU/ml.

Association of DNase B Antibody with ARF

Nine studies provided data on Anti-DNase B titres, of which seven (cases, n=164, controls, n=1031) were amenable to meta-analysis (Figure 3). DNase B antibody levels were significantly increased in ARF cases (OR, 6.97 (95% CI [2.99,16.27]; (I^2, 67.4%)) in comparison to controls. This result was consisted across all studies, irrespective of the use of a guideline in diagnosing ARF cases. The two studies not included in the meta-analysis did not have control groups as a comparator. Saini (2019), used a published standard of 134IU/ml and reported an elevation of anti-DNase B titres in 85% (22 of 26) of cases while Hokonohara (1987), had 41 cases of ARF and showed a 60% (n=25) titre elevation.

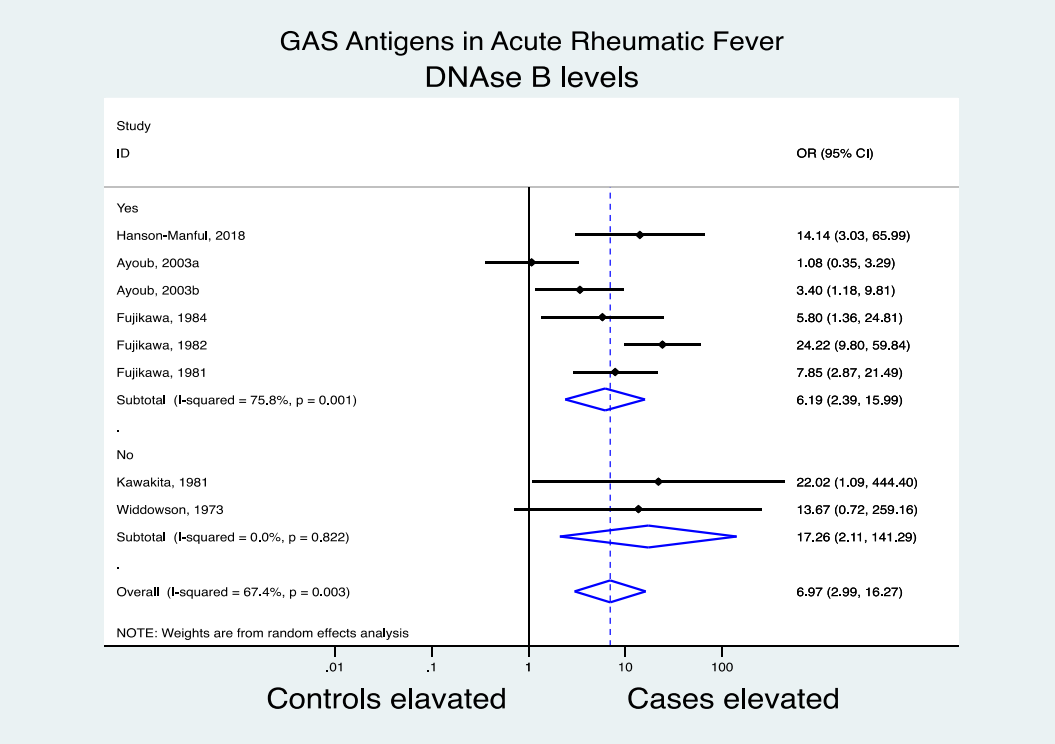


Figure 3 – Forest plot evaluating the odds of association between Anti-DNase B titres and ARF. Subgrouping based on whether a guideline was used to diagnose ARF cases

A sensitivity analysis revealed a reduction in the association of anti-SLO levels with ARF in studies with low-risk scores for bias, (OR, 4.87 (95% CI, 1.07 to 22.17); I^2, 93,1%). Anti-DNase B meta-

analysis comprising studies of a low risk of bias, revealed a non-statistically significant association between antibody titres and ARF.

Association of Other GAS-Specific Antigens With ARF: Streptokinase and GAS Carbohydrate

Three studies provided data on anti-streptokinase (ASK) titres, of which only two (controls, n=532; cases, n=29) were amenable to meta-analysis while two studies (controls, n=416; cases, n=72) provided data on Anti-GAC (AGAC) titres (Figure 4).

ASK antibody levels were significantly increased in ARF patients, (OR, 5.09 (95% CI [2.07,12.50], (I^2, 13%)) while GAC responses showed no significance in elevation between cases and controls (OR, 2.79 (95% CI [0.87,8.99]; (I^2, 67.2%)). Hokonohara (1987), not included in the meta-analysis, showed an elevation of 54% (36 of 67) for ASK and 45% (19 of 41) for AGAC titres in cases of ARF.

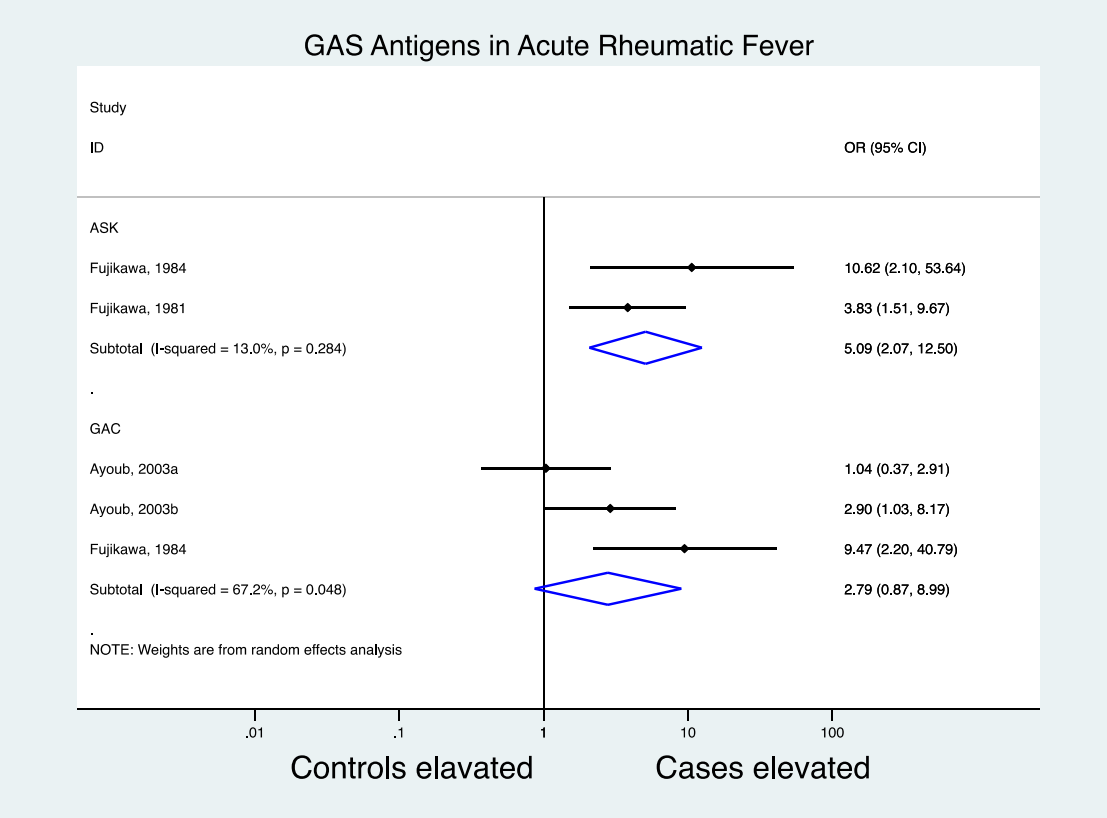


Figure 4 – Forest plot evaluating the odds of association between Anti-DNase B titres and ARF. Subgrouping based on antigen

Narrative review of studies not included in the meta-analysis

Four studies [13, 44, 45, 51] reported on the longitudinal assessment of human immune responses to GAS-specific antigens following a new GAS acquisition. Kaplan (1971), in evaluating the immune response of 49 participants (aged 3-6 years) against SLO, DNase B and NADase antigens, showed that GAS pharyngeal-infected participants had an elevated response to GAS antigens compared with participants without infection: SLO, 57% increase vs 22%, DNase B, 50% vs 11% and NADase, 43% vs 22%. Johnson (2010) reported on the immune response of 160 participants (aged 6-15 years), from whom 3,491 cultures and 1,679 serum samples were obtained. Over the study period they identified 58 new GAS acquisitions in 45 participants, of which 34.5% (n=20) participants showed a significant increase in SLO and DNase B titres. Thirty-six (62.1%) GAS acquisitions were associated with an increase in antibody titre to SLO or DNase B, while 28 showed an increase to SLO and 28 for DNase B. Thus, had only SLO or only DNase B antibody titres been analyzed, eight GAS acquisitions would have been missed. Johnson (2010), provided evidence that of 54 serum samples for ASO and 51 samples for anti-DNase B showing an increase in titre following a new GAS acquisition, ~60% were below the ULN described, resulting in the mischaracterization of preceding GAS infections. Furthermore, in amongst GAS carriers, 239 samples had an increased ASO titre above the ULN and 307 samples for anti-DNase B, while only 9.6% (ASO) and 6.5% (anti-DNase B) of these were associated with a true titre increase following a GAS acquisition. It was also shown that in the absence of a culture-positive for GAS, ASO and anti-DNase B titres were still higher than the ULN for extended periods of time. The study undertaken by Hysmith (2017) presents the most recent and comprehensive investigation of antibody responses following GAS acquisition. From 41 participants (aged 6-15 years) over a period of 2 years, 51 new GAS acquisitions were documented, with 34 showing an increase in antibody titres against SLO and/or DNase B, illustrating an overall sensitivity of 67% in predicting a new GAS acquisition. Including SCPA (C5a peptidase) and one additional GAS-shared antigen to SLO and DNase B, improved the overall sensitivity to 76% and 98% respectively.

Studies that only reported average mean titers were summarized in Table S6 (Appendix D), provided data were available for SK [52], SLO [43, 47, 48], GAC [53], and DNase B [42]. In all the studies, the average mean

titres from cases of ARF were considerably higher in comparison to that of the controls. Studies reporting on less common GAS antigens are summarized in Table S7 (Appendix D), with the use of GAS specific antigens: GAS nuclease A (Spn A) [9], collagen-like surface protein (SCI), putative surface antigen (PSA), SCPA [54], streptococcal esterase (SE) [34], Nicotinamide adenine dinucleotidase (NADase) [39], and superoxide dismutase (SOD) [55], in which only Spn A showed a significant OR (95% CI, 56.00 [9.17; 342.13]).

Discussion

We have presented a comprehensive review of the literature on group A streptococcal antibody responses and their utility in making clinical diagnosis of ARF. Our meta-analysis provides evidence for a significant association between ARF and anti-SLO, anti-DNase B and ASK. Indicating the usefulness of these immunological markers in supporting an ARF diagnosis; providing evidence of a recent GAS infection. This finding is supported by individual studies showing a consistently higher average mean titre in ARF cases over controls. However, there is currently no evidence of an association between GAC antibodies and ARF.

We grouped our findings according to whether peer-reviewed guidelines were used in the clinical diagnosis of ARF. Excluding studies which did not employ a guideline, ie those classified as having a higher risk of bias in terms of case definition, reduced the combined estimate of association for anti-SLO and anti-DNase B and ARF from OR 10.57 (95% CI [3.36, 33.29]) to OR 10.03 (95% CI [4.01, 25.12]) and OR 6.97 (95% CI [2.99, 16.27]) to OR 6.19 (95% CI [2.39, 15.99]) respectively. This may indicate caution in terms of investigating only a single antigen in establishing a recent GAS infection.

Given differences in techniques used to measure antibody titres across the studies, sample size variation, regional differences resulting in the variation of ULN titre levels and published standards used (Table 1), a high degree of heterogeneity was to be expected; hence we employed the random-effect model for

the meta-analysis. Sensitivity analyses of studies with a low risk of bias score revealed a reduction in, although still significant, the association of anti-SLO levels with ARF while anti-DNase B analyses showed no statistically significant association between antibody titres and ARF. Unfortunately, the dearth of studies precluded conducting further meaningful subgroup analyses.

We provide a summary of literature meeting our inclusion criteria, but not amenable to meta-analysis through a narrative review. Additional single antigen studies provide further support for the significant association of SK, SLO, GAC, DNase B, Spn A and SE with ARF but not for SCI, PSA, SCPA and SOD. For completeness, though not encompassing ARF cases, we included four studies reporting on the longitudinal assessment of human immune response to GAS-specific antigens following a new GAS acquisition in acute and convalescent samples. The studies provided meaningful data in terms of the effectiveness of SLO, DNase B, NADase and other antigens in detecting a preceding GAS infection. These findings suggest the need to employ an array of antigens to increase the sensitivity of assays confirming a preceding GAS infection, that could be amended within a ARF diagnostic guideline.

Within the studies reporting study limitations, the use of the ULN to describe elevation in titres as evidence of a recent infection is confounding given its dependence on the controls used within the study. Numerous reports suggest that the ULN of antibody titres towards GAS antigens varies with age and geographical location. It has also been reported that titres are lowered in adults in comparison with children [56, 57]. Johnson (2010) and Hysmith (2017) identified a number of cases where the participant showed an increase in titre following a new GAS acquisition where the peak titres did not exceed the ULN described for the specific population. Given that cases of GAS carriers demonstrated prolonged elevated responses that exceeded the ULN, caution is warranted since using ULN solely to describe elevation may result in false-negative or false-positive GAS-associations. Thus, these studies strongly suggest that evaluating the rise in titre in paired sequential samples as the most effective way in describing a preceding GAS infection.

This systematic review employed rigorous methods as proposed by the Cochrane Collaboration [58] in synthesizing published resources on the utility of GAS serology in confirming a preceding GAS infection. However, the availability of individual patient data would have further enhanced our findings. As is often the case, data were not reported so as to allow inclusion of some studies into the meta-analysis. Also, there remains a lack of studies in this area.

Conclusion

Providing evidence for a preceding GAS infection remains a challenge. Future studies to evaluate serological tests for evidence of a preceding GAS infection should be designed to overcome the major limitations of the existing evidence base. This can be readily accomplished by ensuring a well-defined case definition as in clear symptoms of ARF with paired sequential sampling of the target population. Furthermore, utilizing an array of GAS antigens is more likely to provide greater sensitivity in providing evidence of a recent infection.

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine 2009;**151**(4):264-269.
2. Bhardwaj R, Sood A. Clinical profile of acute rheumatic fever patients in a tertiary care institute in present era. Journal of The Association of Physicians of India 2015;**63**.
3. Gewitz MH, Baltimore RS, Tani LY, Sable CA, Shulman ST, Carapetis J, Remenyi B, Taubert KA, Bolger AF, Beerman L. Revision of the Jones Criteria for the diagnosis of acute rheumatic fever in the era of Doppler echocardiography: a scientific statement from the American Heart Association. Circulation 2015;**131**(20):1806-1818.
4. Organization WH. *Rheumatic fever and rheumatic heart disease: report of a WHO Expert Consultation, Geneva, 29 October-1 November, 2001*: World Health Organization; 2004.
5. Ralph AP, Noonan S, Wade V, Currie BJ. The 2020 Australian guideline for prevention, diagnosis and management of acute rheumatic fever and rheumatic heart disease. Med J Aust 2021;**214**(5):220-227.
6. Parks T, Kado J, Colquhoun S, Carapetis J, Steer A. Underdiagnosis of acute rheumatic fever in primary care settings in a developing country. Trop Med Int Health 2009;**14**(11):1407-13.
7. Okello E, Wanzhu Z, Musoke C, Kakande B, Mondo CK, Freers J, Twalib A, Lwabi P, Wilson NB, Odoi-Adome R. Cardiovascular complications in newly diagnosed rheumatic heart disease patients at Mulago Hospital, Uganda. Cardiovascular journal of Africa 2013;**24**(3):82.
8. Dajani AS, Ayoub E, Bierman FZ, Bisno AL, Denny FW, Durack D, Ferrieri P, Freed M, Gerber M, Kaplan E. Guidelines for the diagnosis of rheumatic fever: Jones criteria, updated 1992: special writing group of the committee on rheumatic fever, endocarditis, and Kawasaki disease of the council on cardiovascular disease in the young, American Heart Association. Circulation 1993;**87**(1).
9. Hanson-Manful P, Whitcombe AL, Young PG, Carr PEA, Bell A, Didsbury A, Mitchell EA, Dunbar PR, Proft T, Moreland NJ. The novel Group A Streptococcus antigen SpnA combined with bead-based immunoassay technology improves streptococcal serology for the diagnosis of acute rheumatic fever. Journal of Infection 2018;**76**(4):361-368.
10. Gerber MA, Baltimore RS, Eaton CB, Gewitz M, Rowley AH, Shulman ST, Taubert KA. Prevention of rheumatic fever and diagnosis and treatment of acute Streptococcal pharyngitis: a scientific statement from the American Heart Association Rheumatic Fever, Endocarditis, and Kawasaki Disease Committee of the Council on Cardiovascular Disease in the Young, the Interdisciplinary Council on Functional Genomics and Translational Biology, and the Interdisciplinary Council on Quality of Care and Outcomes Research: endorsed by the American Academy of Pediatrics. Circulation 2009;**119**(11):1541-1551.
11. Gerber MA, Gray ED, Ferrieri P, Kaplan EL. Enzyme-linked immunosorbent assay of antibodies in human sera to streptococcal DNase B. The Journal of laboratory and clinical medicine 1980;**95**(2):258-265.
12. Podbielski A, Zarges I, Flosdorff A, Weber-Heynemann J. Molecular characterization of a major serotype M49 group A streptococcal DNase gene (sdaD). Infection and immunity 1996;**64**(12):5349-5356.
13. Johnson DR, Kurlan R, Leckman J, Kaplan EL. The human immune response to streptococcal extracellular antigens: clinical, diagnostic, and potential pathogenetic implications. Clin Infect Dis 2010;**50**(4):481-90.
14. Fontaine MC, Lee JJ, Kehoe MA. Combined contributions of streptolysin O and streptolysin S to virulence of serotype M5 Streptococcus pyogenes strain Manfredo. Infection and immunity 2003;**71**(7):3857-3865.
15. O'Seaghdha M, Wessels MR. Streptolysin O and its co-toxin NAD-glycohydrolase protect group A Streptococcus from Xenophagic killing. PLoS Pathog 2013;**9**(6):e1003394.
16. Broudy TB, Pancholi V, Fischetti VA. The in vitro interaction of Streptococcus pyogenes with human pharyngeal cells induces a phage-encoded extracellular DNase. Infection and immunity 2002;**70**(6):2805-2811.

17. Buchanan JT, Simpson AJ, Aziz RK, Liu GY, Kristian SA, Kotb M, Feramisco J, Nizet V. DNase expression allows the pathogen group A Streptococcus to escape killing in neutrophil extracellular traps. Current Biology 2006;**16**(4):396-400.
18. Cleary PP, Prahbu U, Dale J, Wexler D, Handley J. Streptococcal C5a peptidase is a highly specific endopeptidase. Infection and immunity 1992;**60**(12):5219-5223.
19. Courtney HS, Dale JB, Hasty D. Differential effects of the streptococcal fibronectin-binding protein, FBP54, on adhesion of group A streptococci to human buccal cells and HEp-2 tissue culture cells. Infection and immunity 1996;**64**(7):2415-2419.
20. Courtney HS, Dale JB, Hasty DL. Mapping the fibrinogen-binding domain of serum opacity factor of group a streptococci. Curr Microbiol 2002;**44**(4):236-40.
21. Batzloff M, Hayman W, Davies M, Zeng M, Pruksakorn S, Brandt E, Good M. Protection against group A streptococcus by immunization with J8-diphtheria toxoid: contribution of J8-and diphtheria toxoid-specific antibodies to protection. The Journal of infectious diseases 2003;**187**(10):1598-1608.
22. Sabharwal H, Michon F, Nelson D, Dong W, Fuchs K, Manjarrez RC, Sarkar A, Uitz C, Viteri-Jackson A, Suarez RSR. Group A streptococcus (GAS) carbohydrate as an immunogen for protection against GAS infection. The Journal of infectious diseases 2006;**193**(1):129-135.
23. Liu M, Zhu H, Zhang J, Lei B. Active and passive immunizations with the streptococcal esterase Sse protect mice against subcutaneous infection with group A streptococci. Infect Immun 2007;**75**(7):3651-7.
24. Penfound TA, Chiang EY, Ahmed EA, Dale JB. Protective efficacy of group A streptococcal vaccines containing type-specific and conserved M protein epitopes. Vaccine 2010;**28**(31):5017-22.
25. Dale JB, Niedermeyer SE, Agbaosi T, Hysmith ND, Penfound TA, Hohn CM, Pullen M, Bright MI, Murrell DS, Shenep LE, Courtney HS. Protective immunogenicity of group A streptococcal M-related proteins. Clin Vaccine Immunol 2015;**22**(3):344-50.
26. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hrobjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021;**10**(1):89.
27. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Systematic Reviews, 5, 210. In; 2016.
28. Wells GA, Shea B, O'Connell Da, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. In: Oxford; 2000.
29. Robins J, Greenland S, Breslow NE. A general estimator for the variance of the Mantel-Haenszel odds ratio. In: *American journal of epidemiology*. *1986*. Citeseer.
30. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;**21**(11):1539-58.
31. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010;**1**(2):97-111.
32. Julie ZN, Arivelo RZ, Hendriso RD, Ramamonjisoa A, Andry R. Current practice about the evaluation of antibody to streptolysin O (ASO) levels by physicians working in Antananarivo, Madagascar. African health sciences 2014;**14**(2):384-389.
33. Fujikawa S, Kawakita S, Kosakai N, Oda T, Ohkuni M, Shiokawa Y, Watanabe N, Yamada T. Significance of Anti-deoxyribonuclease-B (ADN-B) Determination in Clinical Practice: THE 6th CONFERENCE ON PREVENTION FOR RHEUMATIC FEVER AND RHEUMATIC HEART DISEASE. Japanese Circulation Journal 1982;**46**(11):1180-1183.
34. Fujikawa S, Ohkuni M. Clinical Significance of Anti-streptococcal Esterase (ASE) Determination in Rheumatic Fever and Other Streptococcal Diseases: THE 8TH CONFERENCE ON

PRVENTION FOR RHEUMATIC FEVER AND RHEUMATIC HEART DISEASE. Japanese circulation journal 1984;**48**(12):1330-1333.
35. Fujikawa S, Okuni M. Diagnosis of Streptococcal Infection: Previous or Recent: Proceedings of the 5th Conference on Prevention for Rheumatic Fever and Rheumatic Heart Disease. Japanese circulation journal 1981;**45**(12):1382-1383.
36. Hokonohara M, Yoshinaga M, Baba Y. Study of Antibody Response to 4 Streptococcal Antigens in Rheumatic Fever and Kawasaki Disease with or without Cardiovascular Lesions: THE 11th CONFERENCE ON PREVENTION FOR RHEUMATIC FEVER AND RHEUMATIC HEART DISEASE. Japanese circulation journal 1987;**51**(12):1353-1356.
37. Halbert SP, Swick L, Sonn C. The use of precipitin analysis in agar for the study of human streptococcal infections: II. Ouchterlony and Oakley technics. The Journal of experimental medicine 1955;**101**(5):557-576.
38. Widdowson JP, Maxted W, Newrick C, Parkin D. An outbreak of streptococcal sore throat and rheumatic fever in a Royal Air Force Training camp; significance of serum antibody to M-associated protein. Epidemiology & Infection 1974;**72**(1):1-12.
39. Kawakita S, Takeuchi T, Inoue J, Onishi T, Uemura Y. Infection of Group A Streptococcus and Antibody Response to Extracellular Antigens: Proceedings of the 5th Conference on Prevention for Rheumatic Fever and Rheumatic Heart Disease. japanese Circulation journal 1981;**45**(12):1384-1390.
40. Zainab S, Saleem N, Manzoor A, Khaliq S, Wasim A, Khaliq S, Shehzad F. Antistreptolysin O titer and C-reactive protein levels in pediatric patients of acute rheumatic fever. The Professional Medical Journal 2020;**27**(07):1335-1339.
41. Ayoub EM, Nelson B, Shulman ST, Barrett DJ, Campbell JD, Armstrong G, Lovejoy J, Angoff GH, Rockenmacher S. Group A streptococcal antibodies in subjects with or without rheumatic fever in areas with high or low incidences of rheumatic fever. Clinical and diagnostic laboratory immunology 2003;**10**(5):886-890.
42. Das S, Dileepan T, Johnson D, Kaplan E, Cleary PP. Enzyme-linked immunosorbent assay for group A Streptococcal anti-DNase B in human sera, using recombinant proteins-Comparison to the DNA methyl green micromethod. Journal of immunological methods 2017;**451**:111-117.
43. Gomaa MH, Ali SS, Fattouh AM, Hamza HS, Badr MM. MBL2 gene polymorphism rs1800450 and rheumatic fever with and without rheumatic heart disease: an Egyptian pilot study. Pediatric Rheumatology 2018;**16**(1):24.
44. Hysmith ND, Kaplan EL, Cleary PP, Johnson DR, Penfound TA, Dale JB. Prospective Longitudinal Analysis of Immune Responses in Pediatric Subjects After Pharyngeal Acquisition of Group A Streptococci. J Pediatric Infect Dis Soc 2017;**6**(2):187-196.
45. Kaplan EL, Top Jr FH, Dudding BA, Wannamaker LW. Diagnosis of streptococcal pharyngitis: differentiation of active infection from the carrier state in the symptomatic child. Journal of Infectious Diseases 1971;**123**(5):490-501.
46. Kotby AA, Habeeb NM, Ezz El Arab S. Antistreptolysin O titer in health and disease: levels and significance. Pediatric reports 2012;**4**(1):25-29.
47. Read S, Reid H, Fischetti V, Poon-King T, Ramkissoon R, McDowell M, Zabriskie J. Serial studies on the cellular immune response to streptococcal antigens in acute and convalescent rheumatic fever patients in Trinidad. Journal of clinical immunology 1986;**6**(6):433-441.
48. Read SE, Fischetti VA, Utermohlen V, Falk RE, Zabriskie JB. Cellular Reactivity Studies to Streptococcal Antigens MIGRATION INHIBITION STUDIES IN PATIENTS WITH STREPTOCOCCAL INFECTIONS AND REHEUMATIC FEVER. The Journal of clinical investigation 1974;**54**(2):439-450.
49. Sagar V, Bergmann R, Nerlich A, McMillan DJ, Schmitz DPN, Chhatwal GS. Variability in the distribution of genes encoding virulence factors and putative extracellular proteins of Streptococcus pyogenes in India, a region with high streptococcal disease burden, and implication for development of a regional multisubunit vaccine. Clinical and Vaccine Immunology 2012;**19**(11):1818-1825.

461 50. Saini N, Kumar D, Swarnim S, Bhatt D, Kishore S. Comparison of antistreptolysin O and anti-
462 deoxyribonucleic B titers in healthy children to those with acute pharyngitis, acute rheumatic fever,
463 and rheumatic heart disease aged 5–15 years. Annals of pediatric cardiology 2019;**12**(3):195.
464 51. Shet A, Kaplan EL, Johnson DR, Cleary PP. Immune response to group A streptococcal C5a
465 peptidase in children: implications for vaccine development. The Journal of infectious diseases
466 2003;**188**(6):809-817.
467 52. Tewodros W, Norgren M, Kronvall G. Streptokinase activity among group A streptococci in
468 relation to streptokinase genotype, plasminogen binding, and disease manifestations. Microbial
469 pathogenesis 1995;**18**(1):53-65.
470 53. Thakur A, Prakash K. Detection of antibody to C-carbohydrate of group A streptococci with
471 enzyme-treated whole bacterial cells as antigen for ELISA. Journal of medical microbiology
472 1996;**45**(3):214-218.
473 54. Kumar R, Chakraborti A, Aggarwal A, Vohra H, Sagar V, Dhanda V, Sharma Y, Majumdar S,
474 Hoe N, Krause R. Streptococcus pyogenes pharyngitis & impetigo in a rural area of Panchkula district
475 in Haryana, India. The Indian journal of medical research 2012;**135**(1):133.
476 55. McMillan DJ, Davies MR, Good MF, Sriprakash KS. Immune response to superoxide
477 dismutase in group A streptococcal infection. FEMS Immunology & Medical Microbiology
478 2004;**40**(3):249-256.
479 56. Steer AC, Vidmar S, Ritika R, Kado J, Batzloff M, Jenney AW, Carlin JB, Carapetis JR. Normal
480 ranges of streptococcal antibody titers are similar whether streptococci are endemic to the setting
481 or not. Clin Vaccine Immunol 2009;**16**(2):172-5.
482 57. Blyth CC, Robertson PW. Anti-streptococcal antibodies in the diagnosis of acute and post-
483 streptococcal disease: streptokinase versus streptolysin O and deoxyribonuclease B. Pathology
484 2006;**38**(2):152-6.
485 58. Deeks J, Higgins J, Altman D, Green S. Cochrane handbook for systematic reviews of
486 interventions version 5.1. 0 (updated March 2011). The Cochrane Collaboration 2011.

487

488 **Funding**

489 This work was based on the research supported in part by the National Research Foundation of South
490 Africa (grant number 116287). Any opinion, finding, conclusion, or recommendation expressed in this
491 material was that of the authors, and the NRF does not accept any liability in this regard. MS, KR, BM,
492 KE, JD, and ME were supported by grant (NW17SFRN33630027) from the American Heart
493 Association. JD received funds from the U.S. Public Health Service (NIH grants R01 AI-132117 and
494 R01 AI-010085).

495 **Supplementary Material**

496 The Supplementary Material for this article can be found online at:
497 https://www.frontiersin.org/articles/10.3389/fcvm. 2021.691646/full#supplementary-material

Chapter 5

Title: Serum immune responses to group A streptococcal antigens following pharyngeal acquisitions among Children in Cape Town, South Africa

M. Taariq Salie[1], Babu Muhamed[1, 2], Kélin Engel[1], Kimona Rampersadh[1], Rezeen Daniels[1], Thomas A. Penfound[2], Craig A. Sable[3], Liesl J Zühlke[4,5], James B. Dale[2], Mark E. Engel[1]

[1]Department of Medicine (AFROStrep Research Initiative) and Cape Heart Institute, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

[2]Department of Medicine, University of Tennessee Health Science Center, Memphis, TN, USA

[3]Children's National Health System, Washington, DC, USA

[4]South African Medical Research Council, Francie Van Zyl Drive, Tygerberg, Cape Town

[5]Division of Paediatric Cardiology, Department of Paediatrics, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

Abstract

Introduction: There is limited information on the human immune response following infection with group A streptococcus (Strep A). Animal studies have shown that, in addition to the M protein, shared Strep A antigens elicit protective immunity. This study aimed to investigate the kinetics and specificity of antibody responses against a panel of Strep A antigens in a cohort of school-aged children in Cape Town, South Africa.

Methods: Participants provided serial throat cultures and serum samples at two-monthly follow-up visits. Strep A recovered were emm-typed and serum samples were analysed by ELISA to assess immune responses to thirty-five Strep A antigens (10 shared and 25 M peptides).

Results: Serologic evaluations were performed on serial serum samples from 42 selected participants (from 256 enrolled) based on the number of follow-up visits, the frequency of visits, and throat culture results. Among these, there were 44 Strep A acquisitions, 36 of which were successfully emm-typed. Participants were grouped into three clinical event groups based on culture results and immune responses. A preceding infection was most convincingly represented by a Strep A-positive culture with an immune response to at least one shared antigen and M peptide (11 clinical events) or a Strep A-negative culture with antibody responses to shared antigens and M peptides (9 clinical events). More than a third of participants demonstrated no immune response despite a positive culture.

Conclusion: This study provided important information regarding the complexity and variability of human immune responses following pharyngeal acquisition of Strep A, as well as demonstrating the immunogenicity of Strep A antigens currently under consideration as potential vaccine candidates.

Introduction

Group A Streptococcus (GAS) is a strict human pathogen that causes a variety of infections and diseases such as pharyngitis, scarlet fever, and impetigo, as well as more severe diseases such as necrotizing fasciitis, streptococcal toxic shock syndrome (STSS) and the post-streptococcal autoimmune sequelae acute rheumatic fever (ARF), rheumatic heart disease (RHD), and post-streptococcal glomerulonephritis (PSGN) [1]. Despite being one of the most studied bacterial pathogens, there is not currently a licensed GAS vaccine. One impediment to vaccine development is that there is insufficient information regarding human immune responses to GAS antigens following natural infection. Additionally, there is not an established immune correlate of protection.

Most previous studies have focused on immune responses to the M protein because of its central role as a determinant of virulence and because M antibodies are associated with protection against infection in animals [2, 3] and humans [4]. The multiplicity of genomically distinct M types (now >200) has prompted the development of multivalent M protein-based subunit vaccines to achieve sufficient efficacy [5, 6]. Alternative approaches to the development of vaccines with potentially broad coverage involve the identification of protective antigens that are shared by all or most GAS [7]. While most of the shared antigens being considered as vaccine candidates have demonstrated variable levels of protective immunogenicity in animal models, the role of these antigens in human immunity to GAS infections is unknown.

In a previous report we demonstrated highly variable antibody responses to shared antigens and more consistent, yet incomplete responses to M peptides in participants from the United States that were enrolled in a study of pediatric autoimmune neuropsychiatric disorders associated with GAS infections (PANDAS) [8]. Limitations of the previous study included the limited number of infecting M types, the fact that the participants had previously been identified as meeting criteria for a diagnosis of PANDAS, from geographic regions at very low risk for ARF and RHD.

Our recent systematic review, highlighted limitations in using upper limit of normal (ULN) values of immune responses as a baseline for determining a positive reaction to an antigen given the natural variation of titres across population groups[27]. Thus, there is value in evaluating whether antibody change by itself could serve as an utility in determining positive responses to antigens following a GAS infection, regardless of antibody concentration within populations.

The current study was designed to determine the kinetics and antigen specificity of antibody responses to ten shared GAS antigens and twenty-five M peptides in children with new infections from Cape Town, South Africa where ARF and RHD are highly prevalent [9-11]. The longitudinal clinic-based protocol was designed to enrol children who had throat cultures and serum samples obtained at two-month intervals over a 24-month period. The GAS isolates were M typed and ELISAs were performed to assess changes in antibody levels over time. Overall, the results show that the human immune

responses following GAS acquisition are complex and variable. Despite the lack of concrete and universal patterns following the acquisition of GAS, the data here highlight the uniqueness of the bacteria and the complexities it has with the host. However, some significant patterns emerged which may contribute to our understanding of protective immunity and the limitations presented in this study could be taken into consideration for future studies.

Methods

Setting and participants. This longitudinal study was designed to determine the epidemiology and serum immune responses of children (aged 5-17) recruited from two juxtaposed peri-urban townships in Cape Town, South Africa. The communities are of lower socioeconomic status and comprise various racial and class groupings. A previous study in South Africa reported children as a major reservoir for GAS pharyngitis with an annual incidence of 0.837 cases/100 person-years (unpublished data) and prevalence of approximately 21% in children 5 to 15 years of age [12] in children presenting with sore throat at local clinics.

Recruitment and enrolment. The participants were recruited from a local clinic within each township. Individuals in the waiting area of the clinic who presented with a complaint of a sore throat were invited to participate. After informed consent had been obtained from the parent/legal guardian and assent from children >7 years, a questionnaire was completed by the study nurse to capture the demographics. A physical examination was performed, after which a throat culture and blood sample were obtained. Physical findings and symptoms were recorded on a case report form and documented in the AFRO*Strep* RedCap database. Treatment decisions were made by the clinician according to standard practice without input from the research team. The study protocol and procedures were approved by the Ethics Committee of the University of Cape Town and the Institutional Review Board of the University of Tennessee. While this study did not include participants without pharyngitis, an upper limit of normal (ULN) value was determined from sera obtained from "healthy" participants during scheduled follow-up visits, with the following criteria: laboratory-confirmed GAS negative, asymptomatic beyond a minimum of 100 days from previous symptomatic clinic visit.

Processing of samples. Throat culture specimens and blood samples were obtained upon enrolment and subsequent throat swabs and blood samples were requested at 2-monthly intervals for a period of 24 months. If a child experienced a sore throat between scheduled visits, they were instructed to return to the clinic for an additional throat culture and serum sample. Throat swabs were sent to the National Health Laboratory Service (NHLS) laboratories for culture to identify β-haemolytic streptococci and for standard serogrouping. Pure cultures were sent to the AFRO*Strep* laboratory for *emm*-typing. Cultures were stored at -80°C until the *emm* typing procedure.

Definitions. A GAS acquisition was described as a throat culture positive for GAS during the initial enrolment visit. A new GAS acquisition was defined as a GAS-positive culture isolated from a participant following the enrolment visit. If GAS was present in two sequential visits, a new GAS acquisition was defined if the *emm* types were different. The participants were categorised into three main clinical event groups based on the acquisition of GAS, immune responses to shared antigens and M peptides, as well as symptoms at the time of infection. Upper Limit of Normal optical density (OD) values was defined as the 80th percentile of healthy participants.

M typing. The M typing procedure was performed according to previously published protocols [13-15], using three alternative primer sets [13, 16, 17]. PCR primers were synthesized at the Department of Molecular and Cell Biology, University of Cape Town, South Africa. The PCR products were sent to the University of Stellenbosch Central DNA Sequencing Facility for sequencing. The DNA sequences generated were submitted to the CDC *Streptococcus pyogenes* database [18] for comparison with existing sequences. The sequence was assigned an *emm* type; e.g., *emm*2 and when necessary a subtype, *emm*2.1.

GAS antigens. Ten shared GAS antigens and 25 M peptide antigens were used to assess the human immune responses (Table 1). The M peptides were selected based on the *emm* types recovered from throat cultures in this study. With the exception of DNaseB, the shared antigens were selected based on prior studies indicating that they were potential vaccine antigens. Recombinant proteins were cloned, expressed, and purified as previously described [8].

Antibody assays. ELISA was performed as previously described [8]. Briefly, all sera were tested in duplicate against each antigen at a constant concentration of 5µg/ml. The optimal dilution of test sera was selected for each antigen to yield an OD within the straight-line portion of the ELISA curve. This involved pre-experiments against the antigens, testing the dilutions in serial fashion ranging from 1:100 through 1:25600. The sera were diluted 1:200 for assays with the M peptides, SpyAD, SSE, and Mrp's. For SLO, SCPA, and DNaseB, the sera were diluted 1:12800, 1:3200 and 1:3200, respectively [8]. These dilutions were optimal for detecting the longitudinal change in antibody responses. As a negative control, sera from participants in a previous study[8] were pooled to yield an OD below or equal to the background levels without primary antibody. IVIG was used as the positive control. Positive control sera against SSE (1/200 dilution) and SCPA (1/3200 dilution) were included in each assay to assess day-to-day variation of the ELISA.

The mean OD values from repeated assays were calculated and used as the final OD. All assays were repeated and results that differed by more than 10% were repeated. A significant immune response or increase in response was defined by plotting OD values on a standard ELISA curve. As previously determined by Hysmith [8] and Johnson [19], for an OD reading higher than 0.25, a 40% increase in antibody level from the previous visit was considered a significant immune response to that antigen. Pre-existing

high antibody levels were defined as an OD reading above 0.8 and remained elevated for one or more subsequent visits.

Results

Two-hundred and fifty-six children were enrolled in the study. The protocol specified a two-monthly interval of follow-up for 24 months. However, the average interval between visits for all participants was 89 days (range, 4 to 617 days). Ninety-one participants withdrew or were lost to follow-up (LTFU). The average number of visits per patient was 3.4. Excluding withdrawals or LTFU after the initial enrolment visit, there was an average of 4.5 visits per patient, with the number of subsequent visits ranging from 2 to 11 for the study period.

Prevalence of ***β***-haemolytic streptococci (***β***HS) and GAS among those seeking treatment for pharyngitis at time of enrolment into the study was 70/256 (27.3%) and 54/256 (21.1%), respectively. Overall, 124 ***β***HS isolates were recovered from enrolment and follow-up visits (GAS = 83, GCS = 25, GGS = 16). Sixty-eight of the GAS cultures were successfully emm typed, while the remaining isolates failed to yield sequence data (n=9) or a pure GAS culture from the NHLS was never received (n=2) or produced no growth on blood agar plates during culture (n=4). The sixty-eight successfully sequenced GAS were isolated from 60 participants, of which 46 were obtained upon entry to the study and 22 were new GAS acquisitions during the study**.**

emm type prevalence and potential vaccine coverage. Twenty-seven different *emm* types were obtained throughout the study (Supplementary Figure 1). The most prevalent *emm* type was *emm*1 (n=10), followed by *emm*12 (n=7) and *emm*49 (n=6). When grouped according to *emm* clusters, cluster E3 (n=19) was represented most frequently followed by A-C3 (n=10), E6 and E4 (n=8), and A-C4 and D4 with 7 isolates each. When considering the theoretical coverage of the 30-valent vaccine in a high-risk population, 63.2% (n=43) of GAS isolates could theoretically be covered by the vaccine, which would increase to 72.1% (n=49) considering previously published *emm* types that were cross-opsonized by 30-valent vaccine antisera [5].

GAS acquisitions and symptoms. Seventy-two participants provided a total of 83 positive GAS cultures. Of these, 26/83 (31.3%) were obtained post-enrolment, the majority (23/26, 88.5%) were isolated from asymptomatic participants. The most common physical examination findings were tonsillar swelling (n=55), hoarseness (n=14), anterior cervical node > 1.5cm in diameter (n=14), and a temperature above 38°C (n=13). Children were administered antibiotics according to standard of care at the community clinics (penicillin V or amoxicillin, 250mg/500mg). Sixty-nine participants received antibiotics throughout the study (Supplementary Table 1).

Immune responses to GAS antigens. Serologic evaluations were performed on serial serum samples from 42 participants that were selected based on the number of follow-up visits, the frequency of visits, and throat culture results. The average number of follow-up visits within this group was 4.8 (range 2-11). Forty participants had one or more positive throat cultures during the study period; the remaining two participants had negative throat cultures over five visits.

A total of 202 serum samples were assayed for antibody levels against 10 shared GAS antigens and 25 N-terminal M peptides (Table 1). During the initial review of the serologic results, it was apparent that immune responses were highly variable and no clear patterns emerged. To facilitate analysis of the data, throat culture results and immune responses to shared antigens and M peptides were used to group the data into clinical events (Table 2). Events in group 1 were defined by a positive throat culture for GAS and immune responses to at least one shared antigen and/or an M peptide. Group 2 was similar in terms of the acquisition of GAS but no significant antibody response to any of the shared antigens and M peptides. Group 3 events were characterized by a negative culture for GAS and immune responses to at least one shared antigen and/or an M peptide. Among the 40 participants in this cohort, were 65 clinical events defined by these criteria (Table 2).

Antibody levels (Table 3) in serial serum samples from each participant were considered as either new responses, which in combination with throat culture results defined the clinical event group, or as pre-existing elevated antibody levels using the criteria described above. To assess the overall immunogenicity of GAS shared antigens, the total number of events with new antibody responses was evaluated (Table 4). The frequency of antibody responses was highest against Mrp4B, which is the full-length Mrp4 protein containing semi-conserved N-terminal epitopes and highly conserved C-terminal sequences. The frequency of new antibody responses against SpyCEP, SLO, SpyAD, SSE and SCPA were observed in decreasing frequency (Table 4). Although M-related proteins are expressed by the majority of GAS M types, their N-terminal sequences are semi-conserved among a limited number of Mrp types [20]. As expected, antibodies against Mrp2U and Mrp4U, the N-terminal peptides that define the Mrp type, were observed less frequently and no serum samples contained antibodies against Mrp49U, which is a less prevalent Mrp type [21].

Assessment of immune responses within clinical event groups. By definition, event groups 1 had the highest number of new antibody responses to M peptides (Table 4). Within event group 1, the M peptide antibody response was not always consistent with the M type of the throat culture isolate (Table 3). Nine of eleven GAS were successfully M typed and 5 of the M peptide antibody responses matched the M type recovered at the start of the clinical event. The frequency of antibody responses for each of the groups against shared antigens are displayed in Table 4. Event group 3 (25%) had the highest percentage

of potential responses followed by group 1 (23%). Group 2, despite being GAS-positive, demonstrated no antibody responses to any of the antigens.

New antibody responses to N-terminal M peptides were observed in 20 clinical event groups 1 and 3, which was part of the criterion used to define these groups. Within these groups there were antibody responses against 24 different M peptides (Table 3). Three individuals mounted immune responses against more than one M peptide, which may have represented multiple infections during the intervening period between serum samples. Five of the 20 events included in clinical event groups 1 and 3 had pre-existing high levels of antibody against 6 different M peptides. In none of the 11/16 group 1 clinical events was an *emm* type acquired against which the participant had pre-existing antibodies. Similarly, in the 9/21 group 3 events, none of the M peptide-specific immune responses occurred in subjects with pre-existing high levels of antibodies against the same M peptide. Twenty of 40 participants had pre-existing high levels of M peptide antibodies against 13 different M types (Table 3). Four individuals had pre-existing antibodies against the same emm type that was recovered from the pharynx. Three of these were observed in clinical event group 2 and one was a group 1 event.

Based on the observations above, within clinical event groups 1 (11/16) and 3 (9/21) are most consistent with symptomatic or asymptomatic GAS infections with predicted and temporally related antibody responses to shared antigens and M peptides (Figure 1). Event group 2 was defined based on the absence of immune responses to shared antigens or M peptides, despite a positive throat culture for GAS. None of the event groups was defined using the criterion of pre-existing high levels of antibodies against shared antigens. However, tabulating events in each group with and without high levels of antibodies against five shared antigens (SpyCEP, SLO, SpyAD, SCPA, SSE) revealed significant differences between event group 2 and other event groups (Figure 2). Assuming that events in group 2, which were all associated with a positive throat culture but no immune responses, represent colonization without infection, due to their pre-existing elevated levels of antibodies against one or more shared antigens.

Discussion

This longitudinal study provides support for the utility of GAS-shared antigens and M peptides in studying human immune responses in children from an endemic area in Cape Town following GAS infection. We conclude that (1) there has been a shift in the molecular epidemiology of GAS in the population compared to a previous study, (2) individuals demonstrate GAS-specific antibody responses in the absence of symptomatic GAS infection, (3) the GAS antigen-specific serum antibody responses following pharyngeal acquisitions are complex, and (4) compared to subjects with antibody responses consistent with GAS infection, the cohort of participants with positive throat cultures but no antigen-specific antibody responses were more likely to have pre-existing elevated antibody levels against five shared GAS antigens.

Our study reports both a temporal and continental distribution in *emm*-type epidemiology. In our study conducted a decade prevously[22], the more prevalent *emm* types included *emm*48 (n=17); *emm*12 (n=13); *emm*4 (n=11); *emm*89 (n=9) and *emm*94 (n=9). In this study, only *emm*12 ranks among the most prevalent *emm*-types isolated. Compared with distribution of emm types within Africa, we affirm clear differences as compared with those reported in a recent prevalence systematic review [23] where *emm*18, *emm*65, *emm*75, *emm*76 and *emm*81 are among top emm types reported. The changing molecular epidemiology of GAS infections is consistent with previous studies that showed *emm* type replacements withing large [24] and smaller [25] geographic regions.

Evidence for asymptomatic immunologically significant infections

The entire dataset comprises 68 successfully sequenced GAS acquisitions (recovered from 60 participants) of which, 26 were not recovered from the enrolment visit and were defined as new GAS acquisitions. Of these, only three had symptoms prior to the isolation of GAS. For event groups 1 and 4, 16/20 experienced new GAS acquisitions by either culture, immune responses or both and none of those subjects reported symptoms of sore throat. Previous studies have found that between 65%-75% GAS acquisitions were asymptomatic [8, 26, 27]. This along with the immunological evidence provided in the current study indicates that most new GAS acquisitions may be asymptomatic. This observation is particularly relevant in the context of primary prevention of ARF, which relies on antibiotic treatment of antecedent symptomatic GAS pharyngitis. Asymptomatic immunologically significant GAS infections could theoretically trigger ARF but would not be amenable to primary antibiotic prevention. This finding thus lends support to the anecdotal notion that asymptomatic GAS infection, without subsequent treatment, may be driving RHD incidence. In the current study, 26/55 GAS-positive participants (53 = culture, 2 = immune responses) were correctly identified on enrolment as needing treatment by Clinical Decision Rules. However, 19 participants prescribed antibiotics were GAS-negative. Additionally, in event groups 1 and 3, which most likely represented true infections, only 2/20 participants received antibiotic treatment. Thus, the suggestion by Ganesan and Mayosi to treat all sore throats may be warranted[28].

Antigen-specific antibody responses following GAS acquisitions were variable

All of the shared GAS antigens used in this study, with the exception of Mrp49U, were immunogenic in at least one subject. Across all clinical event groups, the frequency of immune responses to the six shared non-Mrp antigens was highest against SpyCEP>SLO>SpyAD>DNaseB>SCPA and SSE. We believe that clinical event groups 1 and 3 demonstrate the most convincing evidence of GAS infection as opposed to colonization. The immunogenicity of shared antigens in the current study differs somewhat from our previous study [8], which showed that increases in antibody levels following new GAS acquisitions were observed most frequently against DNaseB and SpyAD and least often against SpyCEP. In the current study, 60% of the participants in groups 1 and 3 demonstrated antibody

increases against either SLO or DNaseB, which is similar to the 67% observed in the previous study. While there is not a standardized ELISA for new immune responses to SLO and DNaseB, these results confirm previous observations showing the relatively low sensitivity of these antibody responses in detecting prior GAS infection [8, 19, 29]. This study also highlights the usefulness of employing a change in the antibody response from one time point to the next for accurately describing a positive association with a particular antigen. This is as opposed to comparison of a single reading against that of ULN values in a population which may produce false negatives and false positives, given that a rise in titre, although significant, may remain below the ULN and vice versa (Supplementary Figure 2).

A recent review promotes the identification of antigenic targets and the development of assays that act as correlates of protection to inform Strep A vaccine development; an assay that could be used in all settings, especially in remote areas where the burden of disease is highest and overall medical infrastructure is especially desired[30]. A major goal of this study was to provide details of the human immune response to five of the GAS shared antigens (SpyCEP, SLO, SpyAD, SCPA, SSE) that, among others, are currently being evaluated as vaccine candidates [7]. As stated in Table 1, the antigens tested for protection was deliberate because of their role in Strep A infections. These virulence factors are either secreted or attached to the bacterium, with a purpose of aiding adhesion to the pharynx, cleaving defence mechanisms to avoid phagocytosis, and cell division for the growth of the bacteria. If there was in fact immunity to these correlates of protection, it could potentially be an avenue capable of preventing non-invasive colonisation or infection altogether thus, preventing downstream processes that will lead to autoimmune manifestations [30]. We acknowledge, though, the scope of the study was insufficient to identify a specific antigen that may provide protection, the observation that all 5 shared antigens were immunogenic following natural infection is supportive. Protective efficacy of one or more of the shared antigens or M antigens will most likely be determined during vaccine studies in open populations or GAS challenge studies [31]. This work could, nevertheless, act as a framework for future studies, paying close attention to the limitations of the study, to maximise the potential of obtaining data on protectiveness and furthermore, assisting in identifying correlates of infection, that could be utilized as a signature for earlier diagnosis of RHD.

Identifying clinical event groups using throat culture and immune response results permitted a more detailed evaluation of the entire data set. There were 28 group 2 events identified in 27 participants which were defined by a positive throat culture but no new immune response to any of the antigens tested. Participants with GAS positive throat cultures and no apparent immune responses to any antigens (group 2) were significantly more likely to have pre-existing high levels of antibodies against shared antigens. Levels of pre-existing antibodies in group 2 against SSE, SpyAD, SCPA, SLO and SpyCEP were notably higher than those in event groups 1 and 3. One explanation of these results is that pre-existing immunity to one or more of these antigens conferred some resistance to infection but did not prevent acquisition of GAS in the pharynx. Alternatively, it is also possible that these individuals

experienced a recent infection and continued to carry the GAS that was recovered from the throat. In support of this is the fact that 22/28 group 2 events were initiated during the enrolment visit of the study which precluded previous culture or serologic results. However, it is notable that none of the participants demonstrated any increases in antibody levels to any antigens and only 3 participants had elevated levels of antibodies against the M peptide that corresponded with the *emm* type of GAS recovered from the pharynx. Based on results from our previous longitudinal study [8], one would predict that >60% of participants with recent infection would mount an antibody response to the homologous M peptide and that the antibody levels would be sustained for some time. Although these results are speculative, they provide important information regarding the human immune responses to shared GAS antigens that may provide the framework for future clinical vaccine studies.

Strengths and Limitations

The longitudinal study design provided an opportunity to assess the immune response in sequential samples, mitigating the issues that arise when relying on the upper limit normal (ULN) to describe an increase or decrease to a particular antigen or M-peptide as described by Johnson[19] and Hysmith[8]. We also documented a comprehensive profile for all participants throughout their duration in the study, including the acquisition of GAS, symptoms and administration of antibiotics. All samples were processed according to standard protocols and ELISA assays were conducted in two separate laboratories (CT and UT), to increase the confidence in the results. To the best of our knowledge, this study is the most comprehensive longitudinal analysis of immune responses to GAS-specific antigens.

There were several limitations regarding the overall study that could potentially affect the conclusions derived from the results. While the sequential sampling provided important insight into the immunological patterns to GAS antigens for many of the participants, not all demonstrated immune responses thought to be consistent with a new GAS infection. For most of the participants included in the ELISA subset of the study, the serologic status prior to the positive GAS culture at enrolment was unknown. Thus, for those with a GAS-positive culture on enrolment and elevated levels of M antibodies, it is not possible to determine if these were due to an earlier or the current infection with concomitant increases in antibody levels against one, or more, shared antigens. Secondly, the 2-monthly follow up period may have been too long between visitations as the data suggest that there may have been missed opportunities to detect an immune response and/or recover GAS isolates from the throat. Additionally, there may well have been M peptide responses to alternative emm types; as the panel of M peptides selected for this study was based on the recovered isolates only. Lastly, due to the COVID-19 pandemic, we were not able to recruit our original target of 300 participants and in turn, reducing the number of visits per participant. Moreover, only a subset of the participants was subjected to ELISA assays; these were selected based on the acquisition of GAS, number of visits and the adherence to the

2-monthly follow-up interval. Thus, the completion of the more samples in the future may shed some additional insight to the hosts' immune responses following natural infection with GAS.

Conclusions

This longitudinal study provided insight into the human immune response following pharyngeal GAS acquisitions in children. Some participants with throat cultures positive for GAS but no new immune responses to any GAS antigens were shown to have pre-existing high levels of antibodies against some shared antigens. We describe the complex nature of GAS-specific serum immune responses, including the observation of immunological responses to GAS antigens in the absence of symptomatic GAS infections. This may partly explain the disproportionately high burden of acute rheumatic fever and rheumatic heart disease in low-income populations where GAS acquisition is common but not amenable to antibiotic prevention. Lastly, although the scope of this work was insufficient to confirm the potential of putative vaccine candidates, all antigens tested were in fact immunogenic. This is supportive of future vaccine clinical studies where true human immune correlates of protection may be determined.

References

1. Walker MJ, Barnett TC, McArthur JD, Cole JN, Gillen CM, Henningham A, Sriprakash KS, Sanderson-Smith ML, Nizet V. Disease manifestations and pathogenic mechanisms of Group A Streptococcus. Clin Microbiol Rev 2014;**27**(2):264-301.
2. Lancefield RC. Current knowledge of type-specific M antigens of group A streptococci. The Journal of Immunology 1962;**89**(3):307-313.
3. Penfound TA, Chiang EY, Ahmed EA, Dale JB. Protective efficacy of group A streptococcal vaccines containing type-specific and conserved M protein epitopes. Vaccine 2010;**28**(31):5017-22.
4. Wannamaker L, Denny F, Perry W, Siegel A, Rammelkamp Jr CH. Studies on immunity to streptococcal infections in man. AMA American journal of diseases of children 1953;**86**(3):347-348.
5. Dale JB, Penfound TA, Chiang EY, Walton WJ. New 30-valent M protein-based vaccine evokes cross-opsonic antibodies against non-vaccine serotypes of group A streptococci. Vaccine 2011;**29**:8175-8.
6. Pastural É, McNeil SA, MacKinnon-Cameron D, Ye L, Langley JM, Stewart R, Martin LH, Hurley GJ, Salehi S, Penfound TA. Safety and immunogenicity of a 30-valent M protein-based group a streptococcal vaccine in healthy adult volunteers: A randomized, controlled phase I study. Vaccine 2020;**38**(6):1384-1392.
7. Dale JB, Walker MJ. Update on group A streptococcal vaccine development. Current opinion in infectious diseases 2020;**33**(3):244.
8. Hysmith ND, Kaplan EL, Cleary PP, Johnson DR, Penfound TA, Dale JB. Prospective Longitudinal Analysis of Immune Responses in Pediatric Subjects After Pharyngeal Acquisition of Group A Streptococci. J Pediatric Infect Dis Soc 2017;**6**(2):187-196.
9. Mayosi BM, Gamra H, Dangou J-M, Kasonde J. Rheumatic heart disease in Africa: the Mosi-o-Tunya call to action. The Lancet Global Health 2014;**2**(8):e438-e439.
10. Zuhlke LJ, Engel ME, Watkins D, Mayosi BM. Incidence, prevalence and outcome of rheumatic heart disease in South Africa: a systematic review of contemporary studies. Int J Cardiol 2015;**199**:375-83.
11. Engel ME, Haileamlak A, Zühlke L, Lemmer CE, Nkepu S, van de Wall M, Daniel W, King MS, Mayosi BM. Prevalence of rheumatic heart disease in 4720 asymptomatic scholars from South Africa and Ethiopia. Heart 2015;**101**(17):1389-1394.
12. Engel ME, Muhamed B, Whitelaw AC, Musvosvi M, Mayosi BM, Dale JB. Group A streptococcal emm type prevalence among symptomatic children in Cape Town and potential vaccine coverage. The Pediatric infectious disease journal 2014;**33**(2):208.
13. Beall B, Facklam R, Thompson T. Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci. Journal of clinical microbiology 1996;**34**(4):953-958.
14. Barth D, Naicker P, Engel K, Muhamed B, Basera W, Mayosi B, Dale J, Engel M. Molecular Epidemiology of Noninvasive and Invasive Group A Streptococcal Infections in Cape Town. Msphere 2019;**4**(5).
15. Engel ME. A study of determinants and prevalence of Rheumatic Heart Disease in Cape Town. 2012.
16. CDC. *Protocol for emm typing*. https://www.cdc.gov/streplab/groupa-strep/emm-typing-protocol.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fstreplab%2Fprotocol-emm-type.html (29/04/2020).
17. Frost H, Davies MR, Velusamy S, Delforge V, Erhart A, Darboe S, Steer A, Walker MJ, Beall B, Botteaux A. Updated emm-typing protocol for Streptococcus pyogenes. Clinical Microbiology and Infection 2020;**26**(7):946. e5-946. e8.
18. CDC. *emm- Designations and Clusters*. https://www.cdc.gov/streplab/groupa-strep/resources.html (29/04/2020).
19. Johnson DR, Kurlan R, Leckman J, Kaplan EL. The human immune response to streptococcal extracellular antigens: clinical, diagnostic, and potential pathogenetic implications. Clin Infect Dis 2010;**50**(4):481-90.
20. Frost HR, Sanderson-Smith M, Walker M, Botteaux A, Smeesters PR. Group A streptococcal M-like proteins: from pathogenesis to vaccine potential. FEMS microbiology reviews 2018;**42**(2):193-204.

21. Dale JB, Niedermeyer SE, Agbaosi T, Hysmith ND, Penfound TA, Hohn CM, Pullen M, Bright MI, Murrell DS, Shenep LE, Courtney HS. Protective immunogenicity of group A streptococcal M-related proteins. Clin Vaccine Immunol 2015;**22**(3):344-50.
22. Muhamed B. The molecular characterisation of Group A Streptococcus among children with phyryngitis in the Vanguard community (Bonteheuwel). University of Cape Town 2012.
23. Salie T, Engel K, Moloi A, Muhamed B, Dale JB, Engel ME. Systematic Review and Meta-analysis of the Prevalence of Group A Streptococcal emm Clusters in Africa To Inform Vaccine Development. Msphere 2020;**5**(4):e00429-20.
24. Shulman ST, Tanz RR, Dale JB, Beall B, Kabat W, Kabat K, Cederlund E, Patel D, Rippe J, Li Z. Seven-year surveillance of North American pediatric group A streptococcal pharyngitis isolates. Clinical infectious diseases 2009;**49**(1):78-84.
25. Kaplan EL, Wotton JT, Johnson DR. Dynamic epidemiology of group A streptococcal serotypes associated with pharyngitis. The Lancet 2001;**358**(9290):1334-1337.
26. Martin JM, Green M, Barbadora KA, Wald ER. Group A streptococci among school-aged children: clinical characteristics and the carrier state. Pediatrics 2004;**114**(5):1212-9.
27. Quinn RW, Vander Zwaag R, Lowry PN. Acquisition of group A streptococcal M protein antibodies. Pediatric infectious disease 1985;**4**(4):374-378.
28. Karthikeyan G, Mayosi BM. Is primary prevention of rheumatic fever the missing link in the control of rheumatic heart disease in Africa? Circulation 2009;**120**(8):709-713.
29. Salie MT, Rampersadh K, Muhamed B, Engel KC, Zühlke LJ, Dale JB, Engel ME. Utility of Human Immune Responses to GAS Antigens as a Diagnostic Indicator for ARF: A Systematic Review. Frontiers in Cardiovascular Medicine 2021;**8**.
30. Frost H, Excler J-L, Sriskandan S, Fulurija A. Correlates of immunity to Group A Streptococcus: a pathway to vaccine development. npj Vaccines 2023;**8**(1):1.
31. Osowicki J, Azzopardi KI, Fabri L, Frost HR, Rivera-Hernandez T, Neeland MR, Whitcombe AL, Grobler A, Gutman SJ, Baker C. A controlled human infection model of Streptococcus pyogenes pharyngitis (CHIVAS-M75): an observational, dose-finding study. The Lancet Microbe 2021;**2**(7):e291-e299.
32. Hu MC, Stroop SD, Reddish MA, Beall B, Dale JB. Immunogenicity of a 26-Valent Group A Streptococcal Vaccine. Infection and Immunity 2002;**70**(4):2171-2177.
33. Park HS, Cleary PP. Active and passive intranasal immunizations with streptococcal surface protein C5a peptidase prevent infection of murine nasal mucosa-associated lymphoid tissue, a functional homologue of human tonsils. Infect Immun 2005;**73**(12):7878-86.
34. Ji Y, Carlson B, Kondagunta A, Cleary PP. Intranasal immunization with C5a peptidase prevents nasopharyngeal colonization of mice by the group A Streptococcus. Infection and immunity 1997;**65**(6):2080-2087.
35. Cleary PP, Prahbu U, Dale J, Wexler D, Handley J. Streptococcal C5a peptidase is a highly specific endopeptidase. Infection and immunity 1992;**60**(12):5219-5223.
36. Zingaretti C, Falugi F, Nardi-Dei V, Pietrocola G, Mariani M, Liberatori S, Gallotta M, Tontini M, Tani C, Speziale P, Grandi G, Margarit I. Streptococcus pyogenes SpyCEP: a chemokine-inactivating protease with unique structural and biochemical features. FASEB J 2010;**24**(8):2839-48.
37. Turner CE, Kurupati P, Jones MD, Edwards RJ, Sriskandan S. Emerging role of the interleukin-8 cleaving enzyme SpyCEP in clinical Streptococcus pyogenes infection. J Infect Dis 2009;**200**(4):555-63.
38. Liu M, Zhu H, Zhang J, Lei B. Active and passive immunizations with the streptococcal esterase Sse protect mice against subcutaneous infection with group A streptococci. Infect Immun 2007;**75**(7):3651-7.
39. Gallotta M, Gancitano G, Pietrocola G, Mora M, Pezzicoli A, Tuscano G, Chiarot E, Nardi-Dei V, Taddei AR, Rindi S, Speziale P, Soriani M, Grandi G, Margarit I, Bensi G. SpyAD, a moonlighting protein of group A Streptococcus contributing to bacterial division and host cell adhesion. Infect Immun 2014;**82**(7):2890-901.

472 Table 1. Group A streptococcus Antigens Evaluated

Antigen	Bacterium Location	Function	Reference(s)
Synthetic N-terminal M peptides (25 M types)	Cell surface	Opsonic epitopes	32
Mrp2, 4, and 49 (N-terminal) and Mrp4B (intact protein)	Cell surface	Opsonic epitopes	21
SLO	Secreted	Hemolysin	19
DNaseB	Secreted	Degrades neutrophil nets	19
SCPA	Cell surface and secreted	Cleaves C5a	33 34 35
SpyCEP	Cell surface and secreted	Cleaves IL-8	36 37
SSE	Secreted	Tissue invasion	38
SpyAD	Cell surface	Cell division and adhesion	39

473 Abbreviations: DNaseB, deoxyribonuclease B; IL-8, interleukin 8; Mrp, M-related proteins; SCPA, C5a peptidase; SLO,
474 streptolysin O; SpyCEP, serine protease; SSE, serine esterase.

475

476

Table 2. Clinical events according to throat culture results and immune responses*

Clinical Event number	Throat culture result and observed immune responses						
			Immune Response			Immune Response	
	Culture	All participants	Shared Ags + M peptide,	Shared Ags	Symptomatics only	Shared Ags + M peptide,	Shared Ags
1	+	16	11	16	7	3	6
2	+	28	0	0	22	0	0
3	-	21	9	21	2	0	2

*The 2 participants with negative cultures and no immune responses were not included

477
478

Table 3. Throat culture results and immune responses to shared GAS antigens and M peptides within three defined clinical event groups.

Subj	Culture/M type	Event category	Visit	Sxs	Antibiotic	Antibody level	SSE	SpyAD	MRP4U	MRP2U	MRP4B	SCPA	SLO	DNaseB	SpyCEP	M Peptide
95	89	1	3	-	-	New res.										M92
						Pre-exist										
97	1	1	3	-	-	New res.										M1, M58, M116
						Pre-exist										
124	92	1	3	-	-	New res.										M49, M92
						Pre-exist										
156	GAS	1	7	-	-	New res.										M192
						Pre-exist										
163	75	1	5	-	-	New res.										M75
						Pre-exist										
163	18	1	7	-	-	New res.										M18
						Pre-exist										
170	58	1	1	+	-	New res.										M58
						Pre-exist										M2, M8
174	44	1	1	+	-	New res.										M12
						Pre-exist										
178	GAS	1	5	-	-	New res.										M1
						Pre-exist										
181	22	1	1	+	-	New res.										M94
						Pre-exist										
212	12	1	1	+	Y	New res.										M2, M87
						Pre-exist										

Subj	Culture/M type	Event category	Visit	Sxs	Antibiotic	Antibody level	SSE	SpyAD	MRP4U	MRP2U	MRP4B	SCPA	SLO	DNaseB	SpyCEP	M Peptide
10	94	1	1	+	Y	New res.										
						Pre-exist										M58
45	GAS	1	1	+	Y	New res.										
						Pre-exist										
109	12	1	1	+	-	New res.										
						Pre-exist										M22
109	82	1	7	-	-	New res.										
						Pre-exist										M22
135	87	1	3	-	-	New res.										
						Pre-exist										M1, M87
3	22	2	1	+	-	New res.										
						Pre-exist										M22
28	192	2	1	+	Y	New res.										
						Pre-exist										M8, M192
39	94	2	1	+	-	New res.										
						Pre-exist										M22
41	1	2	10	-	-	New res.										
						Pre-exist										M22
43	8	2	1	+	-	New res.										
						Pre-exist										
54	95	2	1	+	Y	New res.										
						Pre-exist										
77	49	2	1	+	-	New res.										
						Pre-exist										
94	44.1	2	1	+	Y	New res.										
						Pre-exist										

Subj	Culture/M type	Event category	Visit	Sxs	Antibiotic	Antibody level	SSE	SpyAD	MRP4U	MRP2U	MRP4B	SCPA	SLO	DNaseB	SpyCEP	M Peptide
112	70	2	1	+	-	New res.										
						Pre-exist										
125	12	2	3	-	-	New res.										
						Pre-exist										M1, M12, M87, M89
126	75	2	1	+	Y	New res.										
						Pre-exist										
139	GAS	2	2	-	-	New res.										
						Pre-exist										
139	GAS	2	5	-	-	New res.										
						Pre-exist										
143	231	2	1	+	Y	New res.										
						Pre-exist										
149	75	2	1	+	-	New res.										
						Pre-exist										
160	95	2	1	+	Y	New res.										
						Pre-exist										M22
164	80	2	1	+	Y	New res.										
						Pre-exist										
167	44	2	1	+	-	New res.										
						Pre-exist										M75

Subj	Culture/M type	Event category	Visit	Sxs	Antibiotic	Antibody level	SSE	SpyAD	MRP4U	MRP2U	MRP4B	SCPA	SLO	DNaseB	SpyCEP	M Peptide
176	GAS	2	1	+	-	New res.										
						Pre-exist										M58
178	58	2	1	+	Y	New res.										
						Pre-exist										M89
182	GAS	2	1	+	Y	New res.										
						Pre-exist										M1
183	1	2	1	+	Y	New res.										
						Pre-exist										
194	184	2	1	+	-	New res.										
						Pre-exist										M58
194	89	2	3	-	-	New res.										
						Pre-exist										M58
200	1	2	1	+	-	New res.										
						Pre-exist										M116
215	GAS	2	1	+	-	New res.										
						Pre-exist										
217	25	2	3	-	-	New res.										
						Pre-exist										
231	1	2	1	+	-	New res.										
						Pre-exist										M49
6	-	3	7	-	-	New res.										M49
						Pre-exist										
10	-	3	3	-	-	New res.										M75
						Pre-exist										M58, M70
41	-	3	8	-	-	New res.										M6
						Pre-exist										M22

Subj	Culture/M type	Event category	Visit	Sxs	Antibiotic	Antibody level	SSE	SpyAD	MRP4U	MRP2U	MRP4B	SCPA	SLO	DNaseB	SpyCEP	M Peptide
54	-	3	2	-	-	New res.										M82, M89
						Pre-exist										
97	-	3	2	-	-	New res.										M87
						Pre-exist										
124		3	2	-	Y	New res.										M49
						Pre-exist										
167	-	3	3	-	-	New res.										M1
						Pre-exist										
178	-	3	3		-	New res.										M58
						Pre-exist										M89
200	-	3	3	-	-	New res.										M75
						Pre-exist										M116
6	-	3	3	-	-	New res.										
						Pre-exist										
39	-	3	9	-	-	New res.										
						Pre-exist										
41	-	3	5	-	-	New res.										
						Pre-exist										M22
77	-	3	4	-	Y	New res.										
						Pre-exist										
77	-	3	6	-	-	New res.										
						Pre-exist										
94	-	3	3	+	-	New res.										
						Pre-exist										
95	-	3	5	-	-	New res.										
						Pre-exist										

Subj	Culture/M type	Event category	Visit	Sxs	Antibiotic	Antibody level	SSE	SpyAD	MRP4U	MRP2U	MRP4B	SCPA	SLO	DNaseB	SpyCEP	M Peptide
143	-	3	4	-	-	New res.										
						Pre-exist										
156	-	3	3	-	-	New res.										
						Pre-exist										M22
183	-	3	3	+	-	New res.										
						Pre-exist										
215	-	3	3	-	-	New res.										
						Pre-exist										
217	-	3	2	-	-	New res.										
						Pre-exist										

Key: Gradient scale, green for significant increases, blue for high pre-existing antibodies. FC – Fold change, OD – Optical density.								
	1.5-2 FC		2-2.5 FC		2.5-3 FC		3-4 FC		>4 FC
	~0.8 OD		0.8-1.2 OD		1.2-1.6 OD		1.6-2 OD		>2 OD

Table 4. New antibody responses and pre-existing high antibody levels against GAS antigens according to clinical event group (n=65).

	SSE	SpyAD	Mrp4U	Mrp2U	Mrp49U	Mrp4B	SCPA	SLO	DNaseB	SpyCEP	M Peptide
New antibody responses (%)											
GRP 1	5 (31.3)	3 (18.8)	0	4 (25)	0	4 (25)	3 (18.8)	6 (37.5)	3 (18.8)	11 (68.8)	11 (68.8)
GRP 2	0	0	0	0	0	0	0	0	0	0	0
GRP 3	3 (14.3)	7 (33.3)	0	1 (4.8)	0	10 (47.6)	5 (23.8)	12 (57.1)	5 (23.8)	9 (42.9)	9 (42.9)
Total	8 (12.3)	10 (15.4)	0	5 (7.7)	0	14 (21.5)	8 (12.3)	18 (27.7)	8 (12.3)	20 (30.8)	
Pre-existing high antibody levels (%)											
GRP 1	1 (6.3)	0	0	1 (6.3)	0	4 (25)	1 (6.3)	1 (6.3)	1 (6.3)	2 (12.5)	5 (31.3)
GRP 2	10 (35.7)	9 (32.1)	0	1 (3.6)	0	11 (39.3)	5 (17.9)	13 (46.4)	5 (17.9)	10 (35.7)	14 (50)
GRP 3	4 (19.1)	3 (14.3)	2 (9.5)	0	0	10 (47.6)	1 (4.8)	3 (14.3)	0	5 (23.8)	6 (28.6)
Total	15 (23.1)	12 (18.5)	2 (3.1)	2 (3.1)	0	25 (38.5)	7 (10.8)	17 (26.2)	6 (9.2)	17 (26.2)	

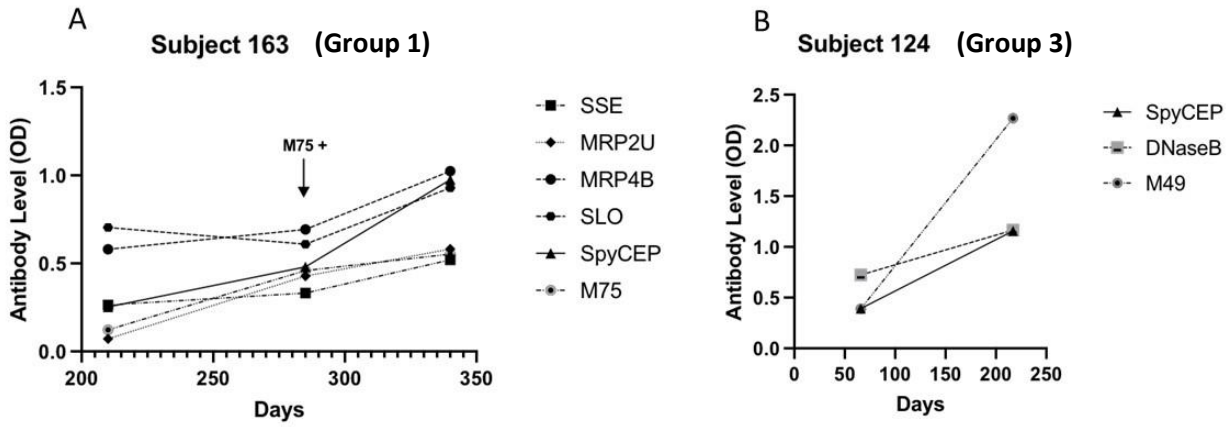


Figure 1. Examples of antibody responses in clinical event groups 1 and 3. The arrow in panel A indicates a throat culture positive for M75 GAS. The throat culture in subject 124 (panel B) was negative for GAS.

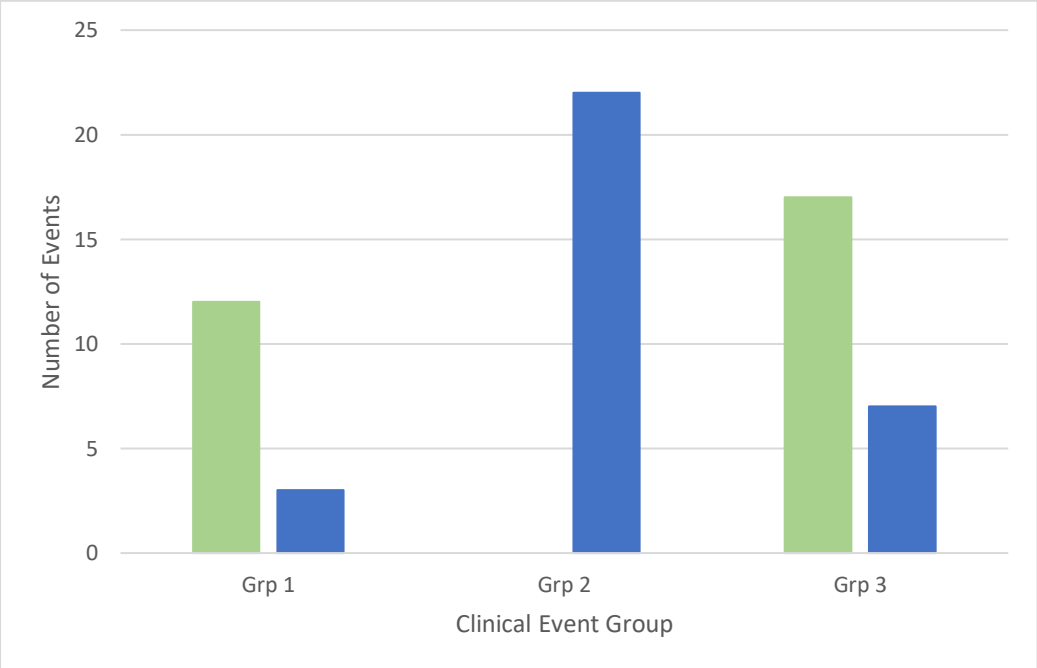


Figure 2. Comparison of clinical events exhibiting pre-existing high levels of antibodies (open bars) against shared GAS antigens (SSE, SpyAD, SCPA, SLO and SpyCEP) versus those without pre-existing high levels of antibodies (grey bars). Grp 1 vs Grp 2: P=8.6E-07, Grps 1 and 3 vs Grp 2: P=4.1E-07, Grp 3 vs Grp 2: P=5.3E-07, Chi Square.

Chapter 6

Highlights and Conclusion

An integral part of vaccine development is the generation of epidemiological data in high disease burden areas. Through the systematic review we provided a succinct summary of the *emm* type distribution in Africa, identifying future research priorities towards vaccine development applicable to the continent. It further demonstrates the possibility of focusing on *emm* clusters as potential vaccine targets, to broaden coverage. The second study had a narrower focus, reporting the epidemiological data pertaining to GAS, in a high-risk population in Cape Town. It also allowed for the testing of the protective coverage afforded by the putative 30-valent vaccine highlighting the uniqueness of regional *emm* type prevalence as well as temporal differences. Study three was a systematic review that summarised the association of GAS shared antigens with ARF. Reporting mainly on SLO and DNase B, I provide support to previous insights questioning the utility of employing the upper-limit-of-normal approach to guide positive immune responses as evidence of a preceding GAS infection for diagnosing ARF. I thus, provide an alternative recommendation for the use of sequential samples subjected to an array of GAS antigens to affirm diagnosis of ARF. The final longitudinal two-year follow-up study integrates clinical evaluations of participants, the characterization of GAS, along with immunological assessment to an array of GAS antigens. The study design with its sequential sampling strategy provides an opportunity to understanding the human immune response to GAS infections; this study, categorizing participants into clinical event groups, provided evidence supporting the use of a panel of GAS antigens in confirming a previous GAS infection.

In conclusion, this thesis serves to aid vaccine development initiatives through four interconnected studies. It further provides a framework for future studies in exploring the human immune response to GAS infection in the quest to provide protective coverage across all serotypes of GAS.

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938) This serves as notification of annual approval, including any documentation described below.			
☑ Approved	Annual progress report	Approved until/next renewal date	30/11/2023
☐ Not approved	See attached comments		
Signature Chairperson of the HREC/ Designee		Date Signed	30/11/2022

Note: Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za.
Please clarify your plan for research-related activities during COVID-19 lockdown.
Please use the latest form found on our website:
http://www.health.uct.ac.za/fhs/research/humanethics/forms

Comments to PI from the HREC
Thank you for the deviation document

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	28 November 2022		
HREC REF Number	461/2018	Current Ethics Approval was granted until	30/07/2022
Protocol title	Towards Rheumatic Heart Disease vaccine development: Defining host immune responses to Group A Streptococcal infection in Cape Town.		
Protocol number (if applicable)			
Are there any sub-studies linked to this study?	☐ Yes	✓ No	
If yes, could you please provide the HREC Reference number for all sub-studies? **Note:** A separate FHS016 must be submitted for each sub-study.			

HUMAN RESEARCH ETHICS COMMITTEE
29 NOV 2022
FACULTY

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Principal Investigator	Dr Mark Engel: M Taariq Salie (Student)
Department / Office Internal Mail Address	J46.43 Old Main Build, GSH Dept of Medicine, UCT

1.1 Does this protocol receive US Federal funding?	☐ Yes	✓ No
1.2 If the study receives US Federal Funding, does the annual report require full committee approval? **Note:** Any annual approvals for **Full Committee** review MUST be submitted on the monthly HREC submission dates. (Please send electronic copy for full committee review to hrec-submission@uct.ac.za)	☐ Yes	☐ No

If yes in 1.2 please complete section 1.3 below for invoicing purposes

1.3 Ethics Renewal Fee

Please **(tick ✓)**appropriate box for billing purposes:

Submission Type	***Description***	***New fee (Vat Incl.)***	***tick ✓***
Research funded solely from UCT departmental/ divisional/group budget	Annual evaluation of research progress report for re-certification	**R0,00**	☑
Non-sponsored student research for degree purposes at UCT/Other Universities & Colleges	Annual evaluation of research progress report for re-certification	**R0,00**	☐
Annual re-certification / Progress report (FHS016 Form)	Clinical Trial & International Grant Funded Research - Annual evaluation of research progress report for re-certification for Full Committee Approval	**R7000,00**	☐
Annual re-certification / Progress report (FHS016 Form)	Clinical Trial & International Grant Funded Research - Annual evaluation of research progress report for re-certification for Expedited review	**R3 710.00**	☐
Annual re-certification / Progress report (FHS016 Form)	National grant funded research - Annual evaluation of research progress report for re-certification for Full Committee Approval	**R6000.00**	☐
Annual re-certification / Progress report (FHS016 Form)	National Grant funded research for Annual evaluation of research progress report for re-certification for Expedited review	**R1 500,00**	☐

NB: Protocols funded by UCT (e.g. departmental funding / student research) and by certain grant funding organizations (e.g. MRC, NRF, CANSA,) are exempt from these charges.

Please provide details for Invoicing, either complete section 1 or 2 :

1. Invoice billing – Directly to Sponsor

Sponsor's name	

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Billing Address of Sponsor:	
Vat Number:	
Contact person	
Telephone number	
Email Address	
2. **Internal Journal Billing:**	
Fund Number:	
Cost Centre Number:	
Account Holder Name:	
Division of Account Holder:	

2. List of documentation for approval

None.

3. Protocol status (tick ✓)

☐	Open Enrolment
✓	Closed to enrolment (tick ✓)
☐	Research-related activities are ongoing
☐	Research-related activities are complete, long-term follow-up only
✓	Research-related activities are complete, data analysis only
☐	Main study is complete but sub-study research-related activities are ongoing
☐	Study is closed → Please submit a Study Closure Form (FHS010)

4. Enrolment

Number of participants enrolled to date	261
Number of participants enrolled, since last HREC Progress report (continuing review)	0
Additional number of participants still required	0

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

5. Refusals

Total number of refusals (participants invited to join the study, but refused to take part)	5

6. Cumulative summary of participants

Total number of participants who provided consent	261
Number of participants determined to be ineligible (i.e. after screening)	3
Number of participants currently active on the study	0
Number of participants completed study (without events leading to withdrawal)	237
Number of participants withdrawn at participants' request (i.e. changed their mind)	24
Number of participants withdrawn by PI due to toxicity or adverse events	0
Number of participants withdrawn by PI for other reasons (e.g. pregnancy, poor compliance)	0
Number of participants lost to follow-up. Please comment below on reasons for loss of follow-up.	
Number of participants no longer taking part for reasons not listed above. Please provide reasons below:	

7. Progress of study

Please provide a brief summary of the research to date including the overall progress and the progress since the last annual report as well as any relevant comments/issues you would like to report to the HREC:
Study recruitment and follow-up was rendered complete at the start of the COVID pandemic. All laboratory activities at the end of semester one and candidate has concentrated on thesis write-up since then. A close-out application will be submitted once thesis is evaluated and corrected.

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

8. Protocol violations and exceptions (tick ✓ all that apply)

✓	No prior violations or exceptions have occurred since the original approval
☐	Prior violations or exceptions have been reported since the last review and have already been acknowledged or approved
☐	Unreported minor violations that have occurred since the last review, as well as significant deviations not yet reported, are attached for review

9. Amendments (tick ✓ all that apply)

✓	No Prior amendments have been made since the original approval
☐	Prior amendments have been reported since the last review and have already been approved
☐	New protocol changes/ amendments are requested as part of this continuing review (See note below)

Note: If new protocol changes are being requested in this review, please complete an amendment form (FHS006).
Specific changes in the amended protocol and consent/assent forms must be **bolded**, *italicised* or tracked and all changes must include a rationale.

10. Adverse events

10.1 Please provide below or attach a narrative summary of serious adverse events and/ or unanticipated problems since the last progress report. Please indicate changes made to the protocol and informed consent document(s) as a result (if not already reported to the HREC). Please comment on whether causality to any study procedure or intervention could be established.
None.

10.2 Have participants received appropriate treatment/ follow-up/ referral when indicated (e.g. in the case of abnormal or incidental clinical findings, distress or anxiety)?		
☐ Yes	☐ No	✓ Not applicable
If yes, please describe:		

11. Summary of Monitoring and Audit Activities (tick ✓)

11.1 Was this study monitored or audited by an external agency (e.g. SAHPRA, FDA)?		
☐ Yes	☐ No	✓ Not applicable

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

11.2 Did a Data and Safety Monitoring Board publish a report?		
☐ Yes	☐ No	✓ Not applicable

11.3 If yes, please identify the agency and attach a summary of the findings.					
Agency Name		Report attached	☐ Yes	☐ No	☐ Not applicable
		DSMB report attached	☐ Yes	☐ No	☐ Not applicable

11.4 Has there been any agency, institutional or other inquiry into non-compliance in this study, or any finding of non-compliance concerning a member of the research team?	
☐ Yes	✓ No
If yes, please explain:	

12. Level of risk (tick ✓)

12.1 In light of your experience of this research, please indicate whether the level of risk to participants has:	
☐	Increased
☐	Decreased
✓	Shown no change
If there has been a change, please explain:	

12.2 Please provide a narrative summary of recent relevant literature that may have a bearing on the level of risk.
None.

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

13. Insurance

Please confirm that valid no fault insurance is still in place? **(tick ✓)**			
☐ Yes	☐ No		✓ Not Applicable – N/A
If yes, please complete the following:			
Insurer's name:			
Policy no.		*Coverage Period:	
For UCT sponsored studies please liaise the Insurance office via fhs.sponsorship@uct.ac.za regarding the required documentation and information required obtain a renewed UCT No-fault Insurance Certificate.			

14. Statement of conflict of interest

Has there been any change in the conflict of interest status of this protocol since the original approval? (tick ✓)	
☐ Yes	✓ No
If yes, please explain and if necessary, attach a revised conflict of interest statement (Section #7 in the New Protocol Application Form FHS013):	

15. Signature

My signature certifies that the above is complete and correct.			
Signature of PI		Date	28 November 2022

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

Form FHS011: Study deviation

HREC office use only (FWA00001637; IRB00001938)			
This serves as acknowledgement of a protocol deviation as described below.			
Chairperson of the HREC signature/ Designee		Date	30/11/2022

Note: Please note that incomplete submissions will not be reviewed.
Please email this form and supporting documents (if applicable) in a combined pdf-file to hrec-enquiries@uct.ac.za.

Please clarify your plan for research-related activities during COVID-19 lockdown

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	29 November 2022
HREC REF Number	461/2018
Project Title	"Towards Rheumatic Heart Disease vaccine development: Defining host immune responses to Group A Streptococcal
Protocol number (if applicable)	
Principal Investigator	Taariq Salie [MSc thesis] / Mark Engel
Department / Office Internal Mail Address	Mark.engel@uct.ac.za

2. Protocol deviation description

Please describe the deviation below, including the reason why the deviation occurred.
We report a failure on our part to submit the annual renewal timeously, as manadated by HREC, by 30/07/2022. Recruitment had stopped as a results of the COVID pandemic, with laboratory assays and analysis having been completed by the 2021-2022 expiry date. Thus, no research activities were undertaken, except for the write-up of the thesis which we wish to submit in the coming weeks, following your approval. I apologize for the delayed submission of the annual report.

HUMAN RESEARCH ETHICS COMMITTEE
29 NOV 2022
HEALTH SCIENCES FACULTY
UNIVERSITY OF CAPE TOWN

3. Follow-up actions

3.1 Please describe any follow-up action(s) taken or planned as a result of this deviation e.g. DSMB reporting, report to sponsor, informing participants.
3.2 Please describe what action(s) have or will be taken to prevent similar deviations in future.
We have delegated the task of Ethics QA to the laboratory coordinator who has set up a calendar for monitoring.

4. Principal Investigator's acknowledgement of responsibility

This signature indicates the PI has reviewed the deviation, taken appropriate follow-up action and implemented or plans to implement preventative steps where possible.			
Signature of PI		Date	29/11/2022

Appendix B – Tables and Figures for Chapter 2: Study I

Table S1. Search Strategy with MeSH terms used for databases

Subject	MeSH terms
Organism	
Group A Streptococcus	"Group A Streptococcus" OR "Group A β-haemolytic Streptococcus" OR "Streptococcus pyogenes" OR "GAS"
Emm type/cluster	"Streptococcus pyogenes gene" OR "Group A Streptococcus typing" OR "GAS strains" OR "sequence types" OR "emm cluster typing system" OR "emm cluster" OR "emm typing" OR "emm sequences" OR "M protein" OR "M protein gene" OR "emm"
Infections	
Superficial Infections	"Non-invasive GAS infections" OR "Pharyngitis" OR "sore throat" OR "strep throat" OR "Impetigo" OR "Group A Strep skin infections" OR "Scarlet fever" OR "scarlatina"
Invasive Infections	"Invasive GAS" OR "iGAS" OR "sepsis" OR "septicaemia" OR "Streptococcal toxic shock syndrome" OR "Necrotizing fasciitis" OR "Strep bacteremia" OR "GAS blood infections" or Group A Streptococcal postpartum metritis" OR "Streptococcal postpartum infections"
Region	
Africa	africa[tw] OR africa'[tw] OR africa's[tw] OR africa1[tw] OR africa2[tw] OR africaans[tw] OR africacollaborations[tw] OR africae[tw] OR africaeaustralis[tw] OR africahiv[tw] OR africaid[tw] OR africaid's[tw] OR africain[tw] OR africaine[tw] OR africaine's[tw] OR africaines[tw] OR africains[tw] OR africal[tw] OR africam[tw] OR africamum[tw] OR african[tw] OR african'[tw] OR african"[tw] OR african's[tw] OR african1[tw] OR african2[tw] OR africana[tw] OR africanae[tw] OR africanalleles[tw] OR africanamerican[tw] OR africanan[tw] OR africanane[tw] OR africananes[tw] OR africanasian[tw] OR africanastrongylus[tw] OR africancalotropis[tw] OR africander[tw] OR africanders[tw] OR africane[tw] OR africanendemic[tw] OR africanene[tw] OR africanenes[tw] OR africanensis[tw] OR africanenvironment[tw] OR africaner[tw] OR africanes[tw] OR africani[tw] OR africanised[tw] OR africanism[tw] OR africanist[tw] OR africanists[tw] OR africanity[tw] OR africanium[tw] OR africanizada[tw] OR africanization[tw] OR africanization'[tw] OR africanize[tw] OR africanized[tw] OR africanized'[tw] OR africanizing[tw] OR africanjournal[tw] OR africannum[tw] OR africano[tw] OR africanoides[tw] OR africanol[tw] OR africanos[tw] OR africanoside[tw] OR africanpatients[tw] OR africanpiper[tw] OR africans[tw] OR africans'[tw] OR africanton[tw] OR africantrinervitermes[tw] OR africantriol[tw] OR africanum[tw] OR africanum'[tw] OR africanumsp[tw] OR africanumt[tw] OR africanus[tw] OR africanus'[tw] OR africanusgen[tw] OR africanz[tw] OR africare[tw] OR africarice[tw] OR africas[tw] OR africasia[tw] OR africative[tw]) OR Algeria[tw] OR Angola[tw] OR Benin[tw] OR Botswana[tw] OR Burundi[tw] OR Cameroon[tw] OR Chad[tw] OR Comoros[tw] OR Congo[tw] OR Djibouti[tw] OR Egypt[tw] OR Eritrea[tw] OR Ethiopia[tw] OR Gabon[tw] OR Gambia[tw] OR Ghana[tw] OR Guinea[tw] OR Jamahiriya[tw] OR Jamahiriya[tw] OR Kenya[tw] OR Lesotho[tw] OR Liberia[tw] OR Libya[tw] OR Libya[tw] OR Madagascar[tw] OR Malawi[tw] OR Mali[tw] OR Mauritania[tw] OR Mauritius[tw] OR Mayotte[tw] OR Morocco[tw] OR Mozambique[tw] OR Mozambique[tw] OR Namibia[tw] OR Niger[tw] OR Nigeria[tw] OR Principe[tw] OR Reunion[tw] OR Rwanda[tw] OR Senegal[tw] OR Seychelles[tw] OR Somalia[tw] OR Sudan[tw] OR Swaziland[tw] OR Tanzania[tw] OR Togo[tw] OR Tunisia[tw] OR Uganda[tw] OR Zaire[tw] OR Zambia[tw] OR Zimbabwe[tw]

Table S2. Inclusion criteria

Inclusion Criteria	
Study aims	Studies that describes the distribution of *emm* types or clusters in a population
Detection method	All study designs in which a GAS *emm* typing scheme was performed, PCR amplification of the *emm* gene or whole genome sequencing
Participants	All participants with confirmed GAS infection at the time of enrolment
Region	Countries situated within the African continent
Publication types	Published and unpublished studies
Language	All, with full English abstracts

Table S3. Study appraisal for risk of bias for included studies. Adapted from Hoy et al. (2012) and Werfalli et al. (2014)

Items	**Quality score**
1. Was the study's target population a close representation of the national population in relation to relevant variables?	1 point
2. Was the sampling frame a true or close representation of the target population?	1 point
3. Were data collected directly from the subjects (as opposed to a proxy)?	1 point
4. Was an acceptable case definition used in the study?	1 point
5. Was the study instrument that measured the parameter of interest shown to have validity and reliability?	1 point
6. Was the same mode of data collection used for all subjects?	1 point
7. Authors' reported limitations of study's methods/results	1 point
Risk of assessment **Total**	7 points
Quality	**Overall score**
Low risk: Further research is very unlikely to change our confidence in the estimate	6 – 7
Moderate risk: Further research is likely to have an important impact on our confidence in the estimate and may change the estimate	4 - 5
High risk: Further research is very likely to have an important impact on our confidence in the estimate and is like to change the estimate	1 - 3

Table S4. Characteristics of excluded studies

Author	Year	Title	Reason for exclusion
Mnif	2019	A report on the first outbreak of emm89 group A streptococcus invasive infections in a burns unit in Tunisia	Investigation report
Maalej	2018	Post-streptococcal glomerulonephritis in the south of Tunisia: A 12-year retrospective review	Not reporting on prevalence
Abd El-Ghany	2015	Group A beta-hemolytic streptococcal pharyngitis and carriage rate among Egyptian children: a case-control study	Not reporting on prevalence
Gonsu	2015	A comparative study of the diagnostic methods for Group A streptococcal sore throat in two reference hospitals in Yaounde, Cameroon	Not reporting on prevalence
Dale	2013	Potential coverage of a multivalent M protein-based Strep A vaccine	Duplicated dataset as in Tapia 2015
Muhamed	2012	The molecular characterisation of Group A Streptococcus among children with pharyngitis in the Vangaurd Community (Bonteheuwel/Langa), Cape Town, South Africa	Duplicated dataset as in Engel 2014
Abdissa	2011	Throat carriage rate and antimicrobial susceptibility pattern of group A Streptococci (GAS) in healthy Ethiopian school children	Not reporting on prevalence
Braito	2004	Epidemiology of streptococcus group A in school aged children in Pemb	Not reporting on prevalence
Tamburlini	1998	Streptococcal pharyngitis in Egyptian children	Case report

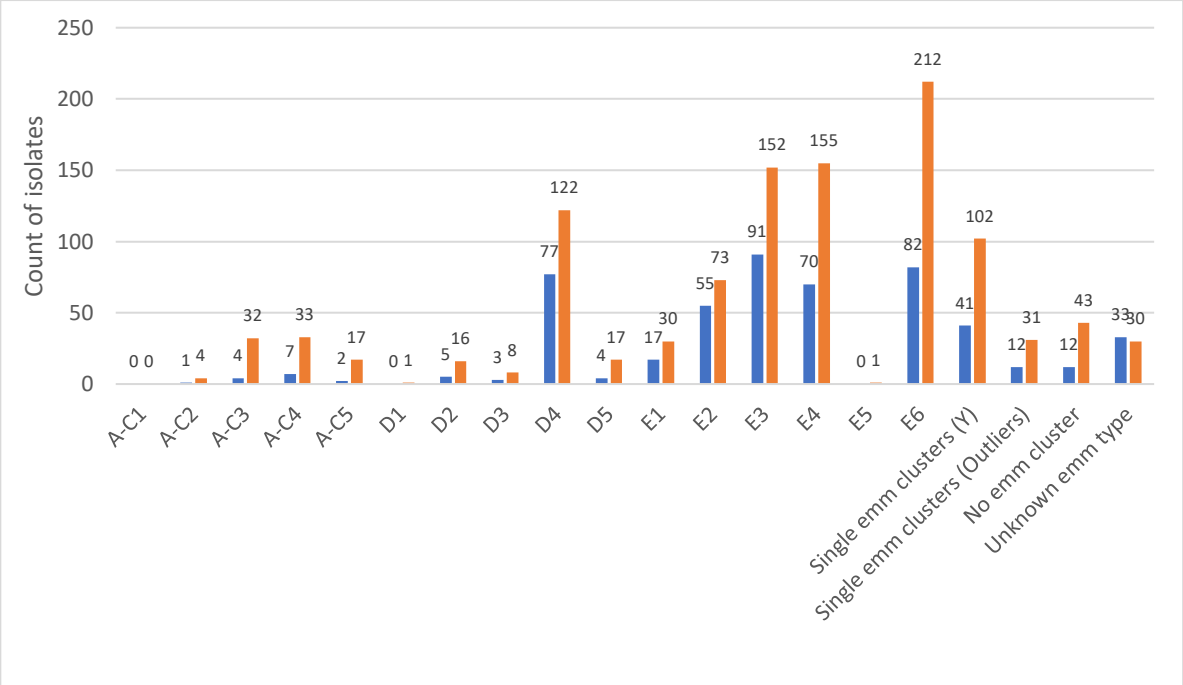


Figure S1. Total prevalence of *emm* clusters categorized on clinical manifestation.

Blue bars, Invasive isolates; orange bars, non-invasive isolates

References

Hoy, D., Brooks, P., Woolf, A., Blyth, F., March, L., Bain, C., Baker, P., Smith, E. et al. 2012. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. *Journal of clinical epidemiology.* 65(9):934-939.

Werfalli, M., Musekiwa, A., Engel, M.E., Ross, I., Kengne, A.P. & Levitt, N.S. 2014. The prevalence of type 2 diabetes mellitus among older people in Africa: a systematic review study protocol. *BMJ open.* 4(6):e004747.

Appendix C – Tables and Figures for Chapter 3: Study II

Supplementary Table 1. Participants with GAS acquisitions and those administered antibiotics with corresponding symptoms

PID	Visit of GAS acquisition	BHS	Antibiotic dosage		Symptoms of pharyngitis (sore throat)											
			500 mg	250 mg	Coug	Hoar	Temp	Rhin	Rash	Tons. swell	Tons. eryth	Tend. ant	Exud. tons	Exud. phar	Ant. cerv	Orop. can
1	Visit 1	N	Y	-	Y	Y	-	-	-	Y	-	-	-	Y	-	-
1[$]	Visit 2	Y(1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	Visit 1	Y(22)	-	-	-	-	Y	-	-	Y	-	-	-	-	-	-
3	Visit 2	N	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-	-
3[$]	Visit 3	Y(1)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
6	Visit 8	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	Visit 1	Y(94)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
10	Visit 1	Y(94)	-	Y[b]	-	-	-	-	-	Y	-	-	-	-	-	-
11[$]	Visit 4	Y(12)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
14	Visit 2	N	Y[a]	-	-	-	-	-	-	Y	Y	-	Y	-	-	-
23	Visit 1	Y(25)	Y	-	-	Y	Y	-	-	Y	-	-	-	-	-	-
23[$]	Visit 4	Y(95)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
28	Visit 1	Y(192)	-	Y[b]	Y	-	-	-	Y	-	Y	-	-	-	-	-
30	Visit 1	Y(244)	-	Y[b]	-	Y	Y	-	-	Y	Y	Y	Y	-	-	-
35	-	-	-	Y[b]	-	-	-	-	-	-	-	-	-	-	-	-
36	Visit 1	N	Y	-	Y	-	Y	-	-	Y	-	-	-	-	-	-
39	Visit 1	Y(94)	-	-	Y	Y	Y	-	-	Y	-	-	-	-	-	-
41	Visit 3	N	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-	-
41[$]	Visit 10	Y(1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
43	Visit 1	Y(8)	-	-	Y	Y	Y	-	-	Y	-	-	-	-	-	-
45	Visit 1	Y*	-	Y	-	-	-	-	-	-	-	-	-	-	-	-
51	Visit 1	GCS	-	Y	Y	Y	-	Y	-	Y	-	-	-	-	-	-
54	Visit 1	Y(95)	-	Y	-	-	Y	-	-	Y	-	-	-	-	-	-
55	Visit 1	GCS	-	Y	-	Y	-	-	-	-	-	-	-	-	-	-
60	Visit 1	N	Y	-	Y	-	-	-	-	Y	-	-	-	Y	-	-
60[$]	Visit 2	Y(49)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
63	Visit 1	N	-	Y	-	Y	-	-	-	Y	Y	-	-	-	-	-
70	Visit 1	Y(1)	Y	-	Y	-	-	Y	-	Y	-	-	-	-	Y	-
71	Visit 1	Y(1)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
75	Visit 1	GCS	-	Y	-	Y	-	-	Y	Y	-	-	-	-	-	-

77	Visit 1	Y(49)	-	-	Y	Y	-	-	Y	Y	-	-	-	-	-	-
77	Visit 4	N	Y[a]	-	N	-	-	-	-	-	-	-	-	-	-	-
83	-	-	Y[b]	-	-	-	-	-	-	-	-	-	-	-	-	-
89	Visit 1	N	-	Y[a]	Y	-	-	-	-	Y	-	-	-	-	-	-
94	Visit 1	Y(44)	Y[a]	-	-	Y	Y	-	-	Y		Y	Y	Y	Y	-
95	Visit 1	N	-	Y[a]	Y	-	Y	Y	-	Y	Y	-	-	Y	-	-
95[$]	Visit 3	Y(187)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
96	Visit 1	N	Y[a]	-	Y	-	-	Y	-	Y	-	-	-	-	-	-
97	Visit 1	N	-	Y	Y	-	-	-	Y	Y	Y	Y	-	Y	-	-
97[$]	Visit 2	Y(1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
99	Visit 1	Y(8)	-	Y	-	-	Y	-	-	Y	Y	Y	-	Y	Y	-
106	Visit 3	GCS	-	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-
107	Visit 1	GCS	-	Y	Y	Y	-	Y	-	Y	-	-	Y	-	-	-
109	Visit 1	Y(12)	-	-	Y	-	-	Y	-	Y	Y	Y	-	-	-	-
109	Visit 3	N	-	Y	-	-	-	-	-	Y	Y	-	-	-	-	-
109[$]	Visit 7	Y(82)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
112	Visit 1	Y(70)	-	-	Y	Y	-	Y	-	Y	-	-	-	-	-	-
117	Visit 1	N	-	Y	Y	Y	-	Y	-	Y	-	-	-	-	-	-
118	Visit 1	Y(116)	-	Y	-	-	-	-	-	Y	Y	-	-	-	-	-
124	Visit 1	N	-	Y[a]	Y	-	Y	Y	-	Y	Y	-	-	-	-	-
124[$]	Visit 3	Y(92)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
125	Visit 1	N	-	Y	-	-	-	-	-	Y	Y	-	Y	-	Y	Y
125	Visit 2	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
125[$]	Visit 3	Y(12)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
126	Visit 1	Y(75)	-	Y[a]	-	Y	-	-	-	Y	-	Y	-	-	Y	-
131	Visit 1	GCS	-	Y	Y	-	-	-	-	Y	Y	Y	Y	-	-	-
135	Visit 1	N	-	Y[c]	-	-	-	-	-	Y	-	-	-	-	-	-
135[$]	Visit 3	Y(87)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
136	Visit 1	Y(70)	-	-	-	Y	Y	-	-	Y	-	-	-	Y	Y	-
136	Visit 3	N	-	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-
137	Visit 1	Y(49)	-	Y	-	-	Y	-	Y	Y	Y	-	-	-	-	-
139	Visit 1	Y*	Y[b]	-	-	-	-	-	-	Y	-	-	Y	-	Y	-
139	Visit 5	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
142	Visit 1	N	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
142[$]	Visit 2	Y(82)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
142	Visit 4	Y(82)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
143	Visit 1	Y(231)	-	Y	-	-	-	Y	-	Y	-	-	-	-	-	-

145	Visit 1	N	Y	-	-	-	-	-	-	Y	-	-	-	-	Y	-
146	Visit 1	N	Y	-	-	-	-	-	-	Y	-	Y	-	-	-	-
146[$]	Visit 6	Y(22)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
149	Visit 1	Y(75)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
149	Visit 3	N	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-	-
150	Visit 1	Y*	Y	-	-	-	-	-	-	Y	Y	-	Y	Y	Y	-
154	Visit 1	N	Y	-	Y	-	-	-	-	Y	-	-	-	-	Y	-
154[$]	Visit 3	Y(12)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
156	Visit 1	N	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
156	Visit 6	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
156	Visit 7	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
158	Visit 1	Y(80)	Y	-	-	Y	-	Y	-	Y	-	Y	-	-	Y	-
158[$]	Visit 2	Y(49)	-	-	-	-	-	-	-	Y	-	-	-	Y	-	-
159	Visit 1	Y(95)	Y	-	-	-	-	-	-	Y	-	-	-	-	Y	-
160	Visit 1	Y(95)	Y	-	-	-	-	-	-	Y	-	-	-	-	-	-
161	Visit 1	GCS	-	Y	-	Y	-	Y	-	Y	-	-	-	-	Y	-
161[$]	Visit 3	Y(49)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
161[$]	Visit 4	Y(223)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
163	Visit 1	N	Y	-	-	Y	-	-	-	Y	-	-	-	-	-	-
163[$]	Visit 5	Y(75)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
163[$]	Visit 7	Y(18)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
164	Visit 1	Y(80)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
167	Visit 1	Y(44)	-	Y[a]	-	-	-	-	-	Y	-	Y	Y	-	Y	-
170	Visit 1	Y(58)	-	Y[b]	-	-	-	-	-	Y	-	-	-	Y	-	-
171	Visit 1	GCS	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
173	Visit 1	Y(6)	-	Y	-	Y	-	-	-	Y	-	-	-	-	-	-
174	Visit 1	Y(44)	-	-	-	-	-	-	-	Y	-	-	-	-	Y	-
176	Visit 1	Y*	-	-	-	Y	Y	-	-	Y	-	Y	Y	-	Y	-
178	Visit 1	Y(58)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
180	Visit 1	Y(12)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
181	Visit 1	Y(22)	-	-	-	-	Y	-	-	Y	-	-	-	Y	-	-
182	Visit 1	Y*	-	-	-	-	-	Y	-	Y	-	Y	-	-	Y	-
183	Visit 1	Y(1)	-	Y	-	-	-	-	-	Y	-	-	Y	-	-	-
194	Visit 1	Y(184)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
194[$]	Visit 3	Y(89)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
200	Visit 1	Y(1)	-	-	-	-	-	-	-	Y	-	Y	-	-	Y	-
203	Visit 1	Y(75)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-

205	Visit 1	Y*	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
209	Visit 1	Y(49)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
212	Visit 1	Y(12)	Y	-	-	-	-	-	-	Y	-	-	-	-	-	-
215	Visit 1	Y*	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
217[$]	Visit 3	Y(25)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
220	Visit 1	Y*	-	Y	Y	Y	Y	-	-	Y	Y	-	Y	-	-	-
224	Visit 1	Y(12)	-	-	-	-	-	-	-	Y	-	-	Y	-	-	-
231	Visit 1	Y(1)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
232	Visit 1	Y(2)	-	Y	-	-	-	Y	-	Y	Y	-	Y	-	-	-
239	Visit 1	N	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
241	Visit 1	Y(87)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
241	Visit 2	N	-	Y	-	-	-	-	-	-	-	-	-	-	-	-
248	Visit 1	Y(94)	Y	-	Y	Y	-	-	-	Y	-	-	-	-	-	-
250	Visit 1	Y*	Y	-	-	-	-	-	-	Y	-	-	-	-	-	-
261	Visit 1	Y(1)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
Total GAS infections (received antibiotics on visit)		83 from 72 PID's (24)	11 (9)	21 (15)	9	14	13	7	3	55	9	10	10	7	14	0
Total of PID's administered antibiotics		69	25	44	11	13	6	8	3	43	11	5	7	4	9	1

Coug, Cough; Hoar, Hoarseness, Tons. swell, Tonsillar swelling; Tons. eryth, Tonsillar erythema; Exud phar / tons, Exudate on the pharynx / tonsils; Temp, Temperature > 38C; Rhin, Rhinorrhea; Ant. cerv, Ant. cervical node > 1.5cm in diameter; Tend. ant, Tender anterior cervical node, Orop. can, Oropharyngeal candidiasis.

N, no GAS recovered; *, GAS *emm* type unknown (untypable); [$], new GAS acquisition, a; received antibiotics after visit; b, received antibiotics but date unknown; c, received antibiotics just before entry into study; GCS, Group C Streptococcus.

Appendix D – Tables and Figures for Chapter 4: Study III

Table S1. Guidelines associated with the diagnosis of ARF

Lead Author	Publication Title	Country	Describes
Gewitz et al. (2015)	Revision of the Jones Criteria for the Diagnosis of Acute Rheumatic Fever in the Era of Doppler Echocardiography	USA	Diagnostic guidelines (echo) & includes preceding GAS evidence & antibody response
Gerber et al. (2009)	Prevention of rheumatic fever and diagnosis and treatment of acute Streptococcal pharyngitis: a scientific statement from the American Heart Association Rheumatic Fever, Endocarditis, and Kawasaki Disease Committee of the Council on Cardiovascular Disease in the Young, the Interdisciplinary Council on Functional Genomics and Translational Biology, and the Interdisciplinary Council on Quality of Care and Outcomes Research: endorsed by the American Academy of Pediatrics	USA	Diagnostic guidelines & includes preceding GAS evidence & antibody response
Saxena et al. (2008)	Consensus guidelines on pediatric acute rheumatic fever and rheumatic heart disease	India	Diagnostic guidelines
Marzouk et al. (2020)	New guidelines for diagnosis of rheumatic fever; do they apply to all populations?	Egypt	Diagnostic guidelines & antibody responses
Dajani et al. (1993)	Guidelines for the diagnosis of rheumatic fever. Jones Criteria, 1992 update. Special Writing Group of the Committee on Rheumatic Fever, Endocarditis, and Kawasaki Disease of the Council on Cardiovascular Disease in the Young of the American Heart Association	USA	Diagnostic guidelines (JC)
Atatoa-Carr et al. (2008)	Rheumatic fever diagnosis, management, and secondary prevention: A New Zealand guideline	New Zealand	Includes preceding GAS evidence & antibody response
Jack et al. (2019)	Streptococcal Serology in Acute Rheumatic Fever Patients: Findings From 2 High-income, High-burden Settings	New Zealand / Australia	ULN described for SLO and DNase B titres
Khriesat et al. (2003)	Acute rheumatic fever in Jordanian children	Jordan	Assessing guidelines & antibody responses (SLO)

Wilson et al. (2013)	New Zealand guidelines for the diagnosis of acute rheumatic fever: small increase in the incidence of definite cases compared to the American Heart Association Jones criteria	New Zealand	Diagnostic comparison of guidelines (NZ vs JC)
Ralph et al. (2006)	The challenge of acute rheumatic fever diagnosis in a high-incidence population: a prospective study and proposed guidelines for diagnosis in Australia's Northern Territory	Australia	Diagnostic assessment of JC
Carapetis et al. (2007)	An Australian guideline for rheumatic fever and rheumatic heart disease: an abridged outline	Australia	Diagnostic guidelines
Vijayalakshmi et al. (2008)	The efficacy of echocardiographic criterions for the diagnosis of carditis in acute rheumatic fever	India	Diagnostic assessment (echo)
Sanyahumbi et al. (2019)	Two-year evolution of latent rheumatic heart disease in Malawi	Malawi	Diagnostic assessment (echo)
Shiffman (1995)	Guideline maintenance and revision. 50 years of the Jones criteria for diagnosis of rheumatic fever	-	Revisions of JC
Clark et al. (2016)	Using a Low-Risk Population to Estimate the Specificity of the World Heart Federation Criteria for the Diagnosis of Rheumatic Heart Disease	Columbia	Diagnostic assessment (WHF)
Kumar et al. (2016)	Evaluation of the American Heart Association 2015 revised Jones criteria versus existing guidelines	India	Diagnostic comparison of guidelines (IND vs JC)
Alqanatish et al. (2019)	Acute rheumatic fever diagnosis and management: Review of the global implications of the new revised diagnostic criteria with a focus on Saudi Arabia	Saudi Arabia	Implication of revised JC
Boyarchuk, Boytsanyuk & Hariyan (2017)	Acute rheumatic fever: clinical profile in children in western Ukraine	Ukraine	Diagnostic assessment of JC
Pereira et al. (2007)	Jones criteria and underdiagnosis of rheumatic fever	Brazil	Diagnostic assessment of JC
Grassi et al. (2009)	Clinical characteristics and cardiac outcome of acute rheumatic fever in Italy in the last 15 years	Italy	Diagnostic assessment of JC
Olgunturk et al. (2006)	Review of 609 patients with rheumatic fever in terms of revised and updated Jones criteria	Turkey	Diagnostic assessment of JC
Chen et al. (2009)	Changes of manifestations of 122 patients with rheumatic fever in South China during last decade	China	Diagnostic comparison of guidelines (WHO vs JC)
Remenyi et al. (2013)	World Heart Federation criteria for echocardiographic diagnosis of rheumatic heart disease—an evidence-based guideline	Global	Diagnostic guidelines (echo)

Table S2. Search strategy with MeSH terms used for Databases (PubMed, Scopus & Google Scholar)

Subject	MeSH terms
Group A Streptococcus	"Group A Streptococcus" OR "Group A β-haemolytic Streptococcus" OR "Streptococcus pyogenes" OR "GAS" OR "Strep A"
Antigens	"SLO" OR "Streptolysin O" OR "DNase B" OR "SCPA" OR "C5A peptidase" OR "M protein" OR "Streptococcal antigens" OR "GAS antigens" OR "common antigens"
Immunology	"Antibod*" OR "immune response" OR "immunological response" OR "ELISA"
Infection: Post-sequelae	"Rheumatic fever" OR "Acute rheumatic fever" OR "ARF" OR "Rheumatic heart disease" OR "RHD"

3

4

Table S3. Inclusion criteria

Inclusion criteria	
Study aims	Studies that describes the expression of bactericidal antibodies evoked by GAS-specific antigens
Detection method	All study designs in which an appropriate immunological assay was completed using sera of participants, producing an upper limit of normal (ULN) to describe elevation in both cases and controls
Participants	All participants with post-sequelae diseases, ARF & RHD with appropriate guidelines and use of an echo completed with controls within the same population and longitudinal studies assessing patients with new GAS acquisitions
Region	Any region
Publication types	Published and unpublished studies
Language	All, with full English abstracts

5

6

Table S4. Summary of risk of bias assessment in case–control studies included in the systematic review (Wells et al., 2000)

Study	A				B		C			Quality score
	A1	A2	A3	A4	B1	B2	C1	C2	C3	
Saini et al. (2019)	1	1	1	1	1	1	1	1	1	9
Kotby, Habeeb & Ezz El Arab (2012)	1	1	1	1	1	1	1	1	1	9
Ayoub et al. (2003)	1	1	0	1	1	1	1	1	1	8
Gomaa et al. (2018)[c]	1	1	0	0	1	1	1	1	1	7
Hanson-Manful et al. (2018)	1	1	0	1	0	1	1	1	1	7
Julie et al. (2014)	1	1	1	1	0	0	0	1	1	6
Read, Stanley E et al. (1974)[bc]	1	1	0	0	1	0	1	1	1	6
Read, SE et al. (1986)[bc]	1	1	0	0	1	0	1	1	1	6
Das et al. (2017)[c]	1	1	0	0	0	0	1	1	1	5
Sagar et al. (2012)	1	1	0	0	0	0	1	1	1	5
Tewodros, Norgren & Kronvall (1995)[c]	1	1	0	0	1	0	0	1	1	5
Thakur & Prakash (1996)[c]	1	1	0	0	1	0	0	1	1	5
Fujikawa & Ohkuni (1984)[b]	1	0	0	0	0	0	1	1	1	4
Fujikawa & Okuni (1981)[b]	1	0	0	0	0	0	1	1	1	4
Fujikawa et al. (1982)[b]	1	0	0	0	0	0	1	1	1	4
Widdowson et al. (1974)[b]	1	0	0	0	1	0	0	1	1	4
Halbert, Swick & Sonn (1955)	1	0	0	0	1	0	0	1	1	4
Hokonohara, Yoshinaga & Baba (1987)[b]	1	0	0	0	0	0	0	1	1	3
Kawakita et al. (1981)[b]	1	0	0	0	0	0	0	1	1	3
Zainab et al. (2020)[a]	1	0	0	0	0	0	1	0	0	2

A. Selection of the study groups: A1 right case definition (ARF); A2 right controls definition (no prior history of ARF); A3 the representativeness of the cases (general population); A4 the representativeness of controls (general population).
B. Comparability of the groups: B1 control of main confounders (age/ethnicity); B2 control of any additional factor.
C. Ascertainment of exposure: C1 appropriate method of exposure ascertainment (guideline); C2 same method of exposure ascertainment for cases and controls; C3 same non-response rate of case and control groups.
1: study met the criteria; 0: the study did not meet the criteria.
Quality score: < 5 high risk of bias. ≥ 5 low risk of bias
[a], no controls but case measurement data; [b], old article with minimal patient demographics; [c], mean data only

Table S5. Characteristics of excluded studies

Author	Year	Title	Reason for exclusion
Abo-Zenah	2008	The Reactive Nature of Acute Rheumatic Fever: Evidence from Streptococcal Cell Wall Antigen Detection by immunotechnology	No controls
Baba	1973	The effect of Beta-Lipoprotein of Antistreptolysin-O Antibody Titres	Wrong disease
Bisno	1982	Type-Specific Antibodies to Structurally Defined Fragments of Streptococcal M Proteins in Patients with Acute Rheumatic Fever	Not GAS antigens
Blackwell	2005	Antistreptokinase antibodies: implications for thrombolysis in a region with endemic streptococcal infection	Wrong disease
Brandt	2001	Antibody levels to the class I and II epitopes of the M protein and myosin are related to Group A Streptococcal exposure in endemic populations	Not GAS antigens
Carapetis	2001	Rheumatic fever in a high incidence population: the importance of monoarthritis and low grade fever	No controls
Catarino	2021	Ficolin-3 in rheumatic fever and rheumatic heart disease	Not GAS antigens
Ekelund	2005	Variations in emm Type among Group A Streptococcal Isolates Causing Invasive or Non-invasive Infections in a Nationwide Study	Wrong disease
Garcia	2016	Cardiac Myosin Epitopes Recognized by Autoantibody in Acute and Convalescent Rheumatic Fever	Not GAS antigens
Gray	1981	Cellular Immune Responses to Extracellular Streptococcal Products Rheumatic Heart Disease	Wrong disease (RHD)
Gupta	2016	Immune response against M protein conserved region peptides from prevalent group A Streptococcus in a North Indian population	Not GAS antigens
Jones	2000	Reactivity of Rheumatic Fever and Scarlet Fever Patients' Sera with Group A Streptococcal M Protein, Cardiac Myosin, and Cardiac Tropomyosin: a Retrospective Study	Not GAS antigens
Joseph	2017	Immuno-nephelometric determination of group streptococcal anti-streptolysin O titres (ASOT) from dried blood spots: Method for validating a new assay	Wrong disease
Kreikemeyer	2005	Streptococcus pyogenes Collagen Type I-binding Cpa Surface Protein	Wrong disease
Marshall	2015	Group A Streptococcal Carriage and Seroepidemiology in Children up to 10 Years of Age in Australia	Wrong disease
Martins	2008	Comprehensive analysis of antibody responses to streptococcal and tissue antigens in patients with acute rheumatic fever	No controls
McMillan	2004	Immune response to superoxide dismutase in group A Streptococcal infection	Wrong disease (RHD)
Mori	1996	Persistent Elevation of Immunoglobulin G Titer against the CRegion of Recombinant Group A Streptococcal M Protein in Patients with Rheumatic Fever	Not GAS antigens
Oda	1981	Clinical appraisal of the Antideoxyribonuclease-B (ADN-B) by means of Streptonase-B test	No controls / NG
Oner	2016	Parameters indicative of persistence of valvular pathology at initial diagnosis in acute rheumatic carditis: the role of albumin and CD19 expression	No controls
Reid	2002	Postgenomic Analysis of Four Novel Antigens of Group A Streptococcus: Growth Phase-Dependent Gene Transcription and Human Serologic Response	No controls
Steer	2009	Normal Ranges of Streptococcal Antibody Titers Are Similar Whether Streptococci Are Endemic to the Setting or Not	Wrong disease
Wahid	1996	Mitral regurgitation may be related with previous streptococcal infection	Not GAS antigens
Watanabe	1987	Anti-streptopolysaccharide Antibody in Children with Rheumatic Fever and Scarlet Fever	Minimal data
Watanabe	1981	Follow-up study of ASO, ADN-B, and ASK levels in children with Rheumatic fever	No controls / NG
Watanabe	1976	Antihyaluronidase Level in Children with Rheumatic Fever and Other Streptococcal Infection	Testing assay
Watanabe	1979	The Significance of Measurement of Anti-Deoxyribonuclease-B in the Patients with Streptococcal Infection	Testing assay
Widdowson	1971	An M-associated protein antigen (MAP) of group a streptococci	Not GAS antigens
Zegeye	2016	Throat culture positivity rate and antibiotic susceptibility pattern of beta-hemolytic streptococci in children on secondary prophylaxis for rheumatic heart disease	No controls

Table S6. Evaluation of studies reporting average mean titres in ARF

Study ID	GAS antigen	No. controls	Control mean	No. cases	Cases mean
Tewodros, 1995	SK	10	87.99 OD450nm	11	131.5 OD450nm
Read, 1986[b]	SLO	34	102 todd units	44	495 todd units
Read, 1974[a]	SLO	NS	600u		800u
Thakur, 1996[a]	GAC	50	0.26u	50	0.63
Das, 2017	DNase B	20	SEM: 20.4 +- 6.36ug/ml	NS	SEM: 93.5 +- 22.73ug/ml
Gomaa, 2018	SLO	80	MED: 29.0 UI/ml (11.3–97.0)	80	MED: 116 UI/ml (26.3–172)

[a], Measurement units not given; [b], geometric mean; NS, Not stated; SEM, standard error of mean; MED, median (interquartile range)

Table S7. Evaluation of studies reporting on less common GAS antigens

Study ID	GAS antigen	No. controls	No. cases	ULN	Odds Ratio (95% CI)
Hanson-Manful, 2018	GAS nuclease A (Spn A)	4/36	14/16	170µg/ml	56.00 (9.17; 342.13)
Sagar, 2012	collagen-like surface protein (SCI)	7/25	8/24	0.3 OD405nm	1.29 (0.38; 4.34)
Sagar, 2012	putative surface antigen (PSA)	11/25	13/24	1.4 OD405nm	1.50 (0.49; 4.64)
Sagar, 2012	C5a peptidase (SCPA)	11/25	11/24	1.2 OD405nm	1.08 (0.35; 3.32)
Fujikawa, 1984	streptococcal esterase (SE)	84/354	5/8	400units	5.36 (1.25; 22.89)
Kawakita, 1981	Nicotinamide adenine dinucleotidase (NADase)	3/84	0/3	333U/ml	-
McMillan, 2004	superoxide dismutase (SOD)	4/23	6/23	0.18 OD450nm	1.68 (0.40; 6.97)

References

Alqanatish, J., Alfadhel, A., Albelali, A. & Alqahtani, D. 2019. Acute rheumatic fever diagnosis and management: Review of the global implications of the new revised diagnostic criteria with a focus on Saudi Arabia. *Journal of the Saudi Heart Association.* 31(4):273-281.

Atatoa-Carr, P., Lennon, D., Wilson, N. & Group, N.Z.R.F.G.W. 2008. Rheumatic fever diagnosis, management, and secondary prevention: a New Zealand guideline. *The New Zealand Medical Journal (Online).* 121(1271).

Ayoub, E.M., Nelson, B., Shulman, S.T., Barrett, D.J., Campbell, J.D., Armstrong, G., Lovejoy, J., Angoff, G.H. et al. 2003. Group A streptococcal antibodies in subjects with or without rheumatic fever in areas with high or low incidences of rheumatic fever. *Clinical and diagnostic laboratory immunology.* 10(5):886-890.

Boyarchuk, O., Boytsanyuk, S. & Hariyan, T. 2017. Acute rheumatic fever: clinical profile in children in western Ukraine. *Journal of medicine and life.* 10(2):122.

Carapetis, J.R., Brown, A., Wilson, N.J. & Edwards, K.N. 2007. An Australian guideline for rheumatic fever and rheumatic heart disease: an abridged outline. *Medical journal of Australia.* 186(11):581-586.

Chen, L., Xie, X., Gu, J., Xu, L., Yang, X. & Yu, B. 2009. Changes of manifestations of 122 patients with rheumatic fever in South China during last decade. *Rheumatology international.* 30(2):239-243.

Clark, B.C., Krishnan, A., McCarter, R., Scheel, J., Sable, C. & Beaton, A. 2016. Using a low-risk population to estimate the specificity of the World Heart Federation criteria for the diagnosis of rheumatic heart disease. *Journal of the American Society of Echocardiography.* 29(3):253-258.

Dajani, A.S., Ayoub, E., Bierman, F.Z., Bisno, A.L., Denny, F.W., Durack, D., Ferrieri, P., Freed, M. et al. 1993. Guidelines for the diagnosis of rheumatic fever: Jones criteria, updated 1992: special writing group of the committee on rheumatic fever, endocarditis, and Kawasaki disease of the council on cardiovascular disease in the young, American Heart Association. *Circulation.* 87(1).

Das, S., Dileepan, T., Johnson, D., Kaplan, E. & Cleary, P.P. 2017. Enzyme-linked immunosorbent assay for group A Streptococcal anti-DNase B in human sera, using recombinant proteins-Comparison to the DNA methyl green micromethod. *Journal of immunological methods.* 451:111-117.

Fujikawa, S. & Okuni, M. 1981. Diagnosis of Streptococcal Infection: Previous or Recent: Proceedings of the 5th Conference on Prevention for Rheumatic Fever and Rheumatic Heart Disease. *Japanese circulation journal.* 45(12):1382-1383.

Fujikawa, S. & Ohkuni, M. 1984. Clinical Significance of Anti-streptococcal Esterase (ASE) Determination in Rheumatic Fever and Other Streptococcal Diseases: THE 8TH CONFERENCE ON PRVENTION FOR RHEUMATIC FEVER AND RHEUMATIC HEART DISEASE. *Japanese circulation journal.* 48(12):1330-1333.

Fujikawa, S., Kawakita, S., Kosakai, N., Oda, T., Ohkuni, M., Shiokawa, Y., Watanabe, N. & Yamada, T. 1982. Significance of Anti-deoxyribonuclease-B (ADN-B) Determination in Clinical Practice: THE 6th CONFERENCE ON PREVENTION FOR RHEUMATIC FEVER AND RHEUMATIC HEART DISEASE. *Japanese circulation journal.* 46(11):1180-1183.

Gerber, M.A., Baltimore, R.S., Eaton, C.B., Gewitz, M., Rowley, A.H., Shulman, S.T. & Taubert, K.A. 2009. Prevention of rheumatic fever and diagnosis and treatment of acute Streptococcal pharyngitis: a scientific statement from the American Heart Association Rheumatic Fever, Endocarditis, and Kawasaki Disease Committee of the Council on Cardiovascular Disease in the Young, the Interdisciplinary Council on Functional Genomics and Translational Biology, and the Interdisciplinary Council on Quality of Care and Outcomes Research: endorsed by the American Academy of Pediatrics. *Circulation.* 119(11):1541-1551.

Gewitz, M.H., Baltimore, R.S., Tani, L.Y., Sable, C.A., Shulman, S.T., Carapetis, J., Remenyi, B., Taubert, K.A. et al. 2015. Revision of the Jones Criteria for the diagnosis of acute rheumatic fever in the era of Doppler echocardiography: a scientific statement from the American Heart Association. *Circulation.* 131(20):1806-1818.

Gomaa, M.H., Ali, S.S., Fattouh, A.M., Hamza, H.S. & Badr, M.M. 2018. MBL2 gene polymorphism rs1800450 and rheumatic fever with and without rheumatic heart disease: an Egyptian pilot study. *Pediatric Rheumatology.* 16(1):24.

Grassi, A., Fesslova, V., Carnelli, V., Boati, E., Dell'Era, L., Salice, P., Bardare, M. & Corona, F. 2009. Clinical characteristics and cardiac outcome of acute rheumatic fever in Italy in the last 15 years. *Clin Exp Rheumatol.* 27(2):366-372.

Halbert, S.P., Swick, L. & Sonn, C. 1955. The use of precipitin analysis in agar for the study of human streptococcal infections: II. Ouchterlony and Oakley technics. *The Journal of experimental medicine.* 101(5):557-576.

Hanson-Manful, P., Whitcombe, A.L., Young, P.G., Carr, P.E.A., Bell, A., Didsbury, A., Mitchell, E.A., Dunbar, P.R. et al. 2018. The novel Group A Streptococcus antigen SpnA combined with bead-based immunoassay technology improves streptococcal serology for the diagnosis of acute rheumatic fever. *Journal of Infection.* 76(4):361-368.

Hokonohara, M., Yoshinaga, M. & Baba, Y. 1987. Study of Antibody Response to 4 Streptococcal Antigens in Rheumatic Fever and Kawasaki Disease with or without Cardiovascular Lesions: THE 11th CONFERENCE ON PREVENTION FOR RHEUMATIC FEVER AND RHEUMATIC HEART DISEASE. *Japanese circulation journal.* 51(12):1353-1356.

Jack, S., Moreland, N.J., Meagher, J., Fittock, M., Galloway, Y. & Ralph, A.P. 2019. Streptococcal serology in acute rheumatic fever patients: findings from 2 high-income, high-burden settings. *The Pediatric Infectious Disease Journal.* 38(1):e1-e6.

Julie, Z.N., Arivelo, R.Z., Hendriso, R.D., Ramamonjisoa, A. & Andry, R. 2014. Current practice about the evaluation of antibody to streptolysin O (ASO) levels by physicians working in Antananarivo, Madagascar. *African health sciences.* 14(2):384-389.

Kawakita, S., TAKEUCHI, T., INOUE, J., ONISHI, T. & UEMURA, Y. 1981. Infection of Group A Streptococcus and Antibody Response to Extracellular Antigens: Proceedings of the 5th Conference on Prevention for Rheumatic Fever and Rheumatic Heart Disease. *Japanese circulation journal.* 45(12):1384-1390.

Khriesat, I., Najada, A., Al Hakim, F. & Abu Haweleh, A. 2003. Acute rheumatic fever in Jordanian children. *EMHJ-Eastern Mediterranean Health Journal, 9 (5-6), 981-987, 2003.*

Kotby, A.A., Habeeb, N.M. & Ezz El Arab, S. 2012. Antistreptolysin O titer in health and disease: levels and significance. *Pediatric reports.* 4(1):25-29.

Kumar, D., Bhutia, E., Kumar, P., Shankar, B., Juneja, A. & Chandelia, S. 2016. Evaluation of the American Heart Association 2015 revised Jones criteria versus existing guidelines. *Heart Asia.* 8(1):30-35.

Marzouk, P.A.A., Hamza, H., Mosaad, N., Emam, S., Fattouh, A.M. & Hamid, L. 2020. New guidelines for diagnosis of rheumatic fever; do they apply to all populations? *Turkish Journal of Pediatrics.* 62(3).

Olgunturk, R., Canter, B., Tunaoglu, F.S. & Kula, S. 2006. Review of 609 patients with rheumatic fever in terms of revised and updated Jones criteria. *International journal of cardiology.* 112(1):91-98.

Pereira, B.A., da Silva, N.A., Andrade, L.E., Lima, F.S., Gurian, F.C. & de Almeida Netto, J.C. 2007. Jones criteria and underdiagnosis of rheumatic fever. *The Indian Journal of Pediatrics.* 74(2):117-121.

Ralph, A., Jacups, S., McGough, K., McDonald, M. & Currie, B.J. 2006. The challenge of acute rheumatic fever diagnosis in a high-incidence population: a prospective study and proposed guidelines for diagnosis in Australia's Northern Territory. *Heart, Lung and Circulation.* 15(2):113-118.

Read, S., Reid, H., Fischetti, V., Poon-King, T., Ramkissoon, R., McDowell, M. & Zabriskie, J. 1986. Serial studies on the cellular immune response to streptococcal antigens in acute and convalescent rheumatic fever patients in Trinidad. *Journal of clinical immunology.* 6(6):433-441.

Read, S.E., Fischetti, V.A., Utermohlen, V., Falk, R.E. & Zabriskie, J.B. 1974. Cellular Reactivity Studies to Streptococcal Antigens MIGRATION INHIBITION STUDIES IN PATIENTS WITH STREPTOCOCCAL INFECTIONS AND REHEUMATIC FEVER. *The Journal of clinical investigation.* 54(2):439-450.

Remenyi, B., Carapetis, J., Wyber, R., Taubert, K. & Mayosi, B.M. 2013. Position statement of the World Heart Federation on the prevention and control of rheumatic heart disease. *Nature Reviews Cardiology.* 10(5):284.
Sagar, V., Bergmann, R., Nerlich, A., McMillan, D.J., Schmitz, D.P.N. & Chhatwal, G.S. 2012. Variability in the distribution of genes encoding virulence factors and putative extracellular proteins of Streptococcus pyogenes in India, a region with high streptococcal disease burden, and implication for development of a regional multisubunit vaccine. *Clinical and Vaccine Immunology.* 19(11):1818-1825.
Saini, N., Kumar, D., Swarnim, S., Bhatt, D. & Kishore, S. 2019. Comparison of antistreptolysin O and anti-deoxyribonucleic B titers in healthy children to those with acute pharyngitis, acute rheumatic fever, and rheumatic heart disease aged 5–15 years. *Annals of pediatric cardiology.* 12(3):195.
Sanyahumbi, A., Beaton, A., Guffey, D., Hosseinipour, M.C., Karlsten, M., Minard, C.G., Penny, D.J., Sable, C.A. et al. 2019. Two-year evolution of latent rheumatic heart disease in Malawi. *Congenital heart disease.* 14(4):614-618.
Saxena, A., Kumar, R.K., Gera, R., Radhakrishnan, S., Mishra, S. & Ahmed, Z. 2008. Consensus guidelines on pediatric acute rheumatic fever and rheumatic heart disease. *Indian pediatrics.* 45(7):565-573.
Shiffman, R.N. 1995. Guideline maintenance and revision: 50 years of the Jones criteria for diagnosis of rheumatic fever. *Archives of pediatrics & adolescent medicine.* 149(7):727-732.
Tewodros, W., Norgren, M. & Kronvall, G. 1995. Streptokinase activity among group A streptococci in relation to streptokinase genotype, plasminogen binding, and disease manifestations. *Microbial pathogenesis.* 18(1):53-65.
Thakur, A. & Prakash, K. 1996. Detection of antibody to C-carbohydrate of group A streptococci with enzyme-treated whole bacterial cells as antigen for ELISA. *Journal of medical microbiology.* 45(3):214-218.
Vijayalakshmi, I.B., Vishnuprabhu, R.O., Chitra, N., Rajasri, R. & Anuradha, T.V. 2008. The efficacy of echocardiographic criterions for the diagnosis of carditis in acute rheumatic fever. *Cardiology in the Young.* 18(6):586-592.
Wells, G.A., Shea, B., O'Connell, D.a., Peterson, J., Welch, V., Losos, M. & Tugwell, P. 2000. *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses.* Oxford.
Widdowson, J.P., Maxted, W., Newrick, C. & Parkin, D. 1974. An outbreak of streptococcal sore throat and rheumatic fever in a Royal Air Force Training camp; significance of serum antibody to M-associated protein. *Epidemiology & Infection.* 72(1):1-12.
Wilson, N.J., Voss, L., Morreau, J., Stewart, J. & Lennon, D. 2013. New Zealand guidelines for the diagnosis of acute rheumatic fever: small increase in the incidence of definite cases compared to the American Heart Association Jones criteria. *The New Zealand Medical Journal (Online).* 126(1379).
Zainab, S., Saleem, N., Manzoor, A., Khaliq, S., Wasim, A., Khaliq, S. & Shehzad, F. 2020. Antistreptolysin O titer and C-reactive protein levels in pediatric patients of acute rheumatic fever. *The Professional Medical Journal.* 27(07):1335-1339.

Appendix E – Tables and Figures for Chapter 5: Study IV

Supplementary Table 1. Participants with GAS acquisitions and those administered antibiotics with corresponding symptoms

PID	Visit of GAS acquisition	BHS	Antibiotic dosage		Symptoms of pharyngitis (sore throat)											
			500 mg	250 mg	Coug	Hoar	Temp	Rhin	Rash	Tons. swell	Tons. eryth	Tend. ant	Exud. tons	Exud. phar	Ant. cerv	Orop. can
1	Visit 1	N	Y	-	Y	Y	-	-	-	Y	-	-	-	Y	-	-
1[$]	Visit 2	Y(1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	Visit 1	Y(22)	-	-	-	-	Y	-	-	Y	-	-	-	-	-	-
3	Visit 2	N	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-	-
3[$]	Visit 3	Y(1)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
6	Visit 8	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9	Visit 1	Y(94)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
10	Visit 1	Y(94)	-	Y[b]	-	-	-	-	-	Y	-	-	-	-	-	-
11[$]	Visit 4	Y(12)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
14	Visit 2	N	Y[a]	-	-	-	-	-	-	Y	Y	-	Y	-	-	-
23	Visit 1	Y(25)	Y	-	-	Y	Y	-	-	Y	-	-	-	-	-	-
23[$]	Visit 4	Y(95)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
28	Visit 1	Y(192)	-	Y[b]	Y	-	-	-	Y	-	Y	-	-	-	-	-
30	Visit 1	Y(244)	-	Y[b]	-	Y	Y	-	-	Y	Y	Y	Y	-	-	-
35	-	-	-	Y[b]	-	-	-	-	-	-	-	-	-	-	-	-
36	Visit 1	N	Y	-	Y	-	Y	-	-	Y	-	-	-	-	-	-
39	Visit 1	Y(94)	-	-	Y	Y	Y	-	-	Y	-	-	-	-	-	-
41	Visit 3	N	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-	-
41[$]	Visit 10	Y(1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
43	Visit 1	Y(8)	-	-	Y	Y	Y	-	-	Y	-	-	-	-	-	-
45	Visit 1	Y*	-	Y	-	-	-	-	-	-	-	-	-	-	-	-
51	Visit 1	GCS	-	Y	Y	Y	-	Y	-	Y	-	-	-	-	-	-
54	Visit 1	Y(95)	-	Y	-	-	Y	-	-	Y	-	-	-	-	-	-
55	Visit 1	GCS	-	Y	-	Y	-	-	-	-	-	-	-	-	-	-
60	Visit 1	N	Y	-	Y	-	-	-	-	Y	-	-	-	Y	-	-
60[$]	Visit 2	Y(49)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
63	Visit 1	N	-	Y	-	Y	-	-	-	Y	Y	-	-	-	-	-
70	Visit 1	Y(1)	Y	-	Y	-	-	Y	-	Y	-	-	-	-	Y	-
71	Visit 1	Y(1)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
75	Visit 1	GCS	-	Y	-	Y	-	-	Y	Y	-	-	-	-	-	-

77	Visit 1	Y(49)	-	-	Y	Y	-	-	Y	Y	-	-	-	-	-	-
77	Visit 4	N	Y[a]	-	N	-	-	-	-	-	-	-	-	-	-	-
83	-	-	Y[b]	-	-	-	-	-	-	-	-	-	-	-	-	-
89	Visit 1	N	-	Y[a]	Y	-	-	-	-	Y	-	-	-	-	-	-
94	Visit 1	Y(44)	Y[a]	-	-	Y	Y	-	-	Y		Y	Y	Y	Y	-
95	Visit 1	N	-	Y[a]	Y	-	Y	Y	-	Y	Y	-	-	Y	-	-
95[$]	Visit 3	Y(187)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
96	Visit 1	N	Y[a]	-	Y	-	-	Y	-	Y	-	-	-	-	-	-
97	Visit 1	N	-	Y	Y	-	-	-	Y	Y	Y	Y	-	Y	-	-
97[$]	Visit 2	Y(1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
99	Visit 1	Y(8)	-	Y	-	-	Y	-	-	Y	Y	Y	-	Y	Y	-
106	Visit 3	GCS	-	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-
107	Visit 1	GCS	-	Y	Y	Y	-	Y	-	Y	-	-	Y	-	-	-
109	Visit 1	Y(12)	-	-	Y	-	-	Y	-	Y	Y	Y	-	-	-	-
109	Visit 3	N	-	Y	-	-	-	-	-	Y	Y	-	-	-	-	-
109[$]	Visit 7	Y(82)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
112	Visit 1	Y(70)	-	-	Y	Y	-	Y	-	Y	-	-	-	-	-	-
117	Visit 1	N	-	Y	Y	Y	-	Y	-	Y	-	-	-	-	-	-
118	Visit 1	Y(116)	-	Y	-	-	-	-	-	Y	Y	-	-	-	-	-
124	Visit 1	N	-	Y[a]	Y	-	Y	Y	-	Y	Y	-	-	-	-	-
124[$]	Visit 3	Y(92)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
125	Visit 1	N	-	Y	-	-	-	-	-	Y	Y	-	Y	-	Y	Y
125	Visit 2	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
125[$]	Visit 3	Y(12)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
126	Visit 1	Y(75)	-	Y[a]	-	Y	-	-	-	Y	-	Y	-	-	Y	-
131	Visit 1	GCS	-	Y	Y	-	-	-	-	Y	Y	Y	Y	-	-	-
135	Visit 1	N	-	Y[c]	-	-	-	-	-	Y	-	-	-	-	-	-
135[$]	Visit 3	Y(87)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
136	Visit 1	Y(70)	-	-	-	Y	Y	-	-	Y	-	-	-	Y	Y	-
136	Visit 3	N	-	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-
137	Visit 1	Y(49)	-	Y	-	-	Y	-	Y	Y	Y	-	-	-	-	-
139	Visit 1	Y*	Y[b]	-	-	-	-	-	-	Y	-	-	Y	-	Y	-
139	Visit 5	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
142	Visit 1	N	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
142[$]	Visit 2	Y(82)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
142	Visit 4	Y(82)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
143	Visit 1	Y(231)	-	Y	-	-	-	Y	-	Y	-	-	-	-	-	-

145	Visit 1	N	Y	-	-	-	-	-	-	Y	-	-	-	-	Y	-
146	Visit 1	N	Y	-	-	-	-	-	-	Y	-	Y	-	-	-	-
146[$]	Visit 6	Y(22)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
149	Visit 1	Y(75)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
149	Visit 3	N	Y[a]	-	-	-	-	-	-	-	-	-	-	-	-	-
150	Visit 1	Y*	Y	-	-	-	-	-	-	Y	Y	-	Y	Y	Y	-
154	Visit 1	N	Y	-	Y	-	-	-	-	Y	-	-	-	-	Y	-
154[$]	Visit 3	Y(12)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
156	Visit 1	N	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
156	Visit 6	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
156	Visit 7	Y*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
158	Visit 1	Y(80)	Y	-	-	Y	-	Y	-	Y	-	Y	-	-	Y	-
158[$]	Visit 2	Y(49)	-	-	-	-	-	-	-	Y	-	-	-	Y	-	-
159	Visit 1	Y(95)	Y	-	-	-	-	-	-	Y	-	-	-	-	Y	-
160	Visit 1	Y(95)	Y	-	-	-	-	-	-	Y	-	-	-	-	-	-
161	Visit 1	GCS	-	Y	-	Y	-	Y	-	Y	-	-	-	-	Y	-
161[$]	Visit 3	Y(49)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
161[$]	Visit 4	Y(223)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
163	Visit 1	N	Y	-	-	Y	-	-	-	Y	-	-	-	-	-	-
163[$]	Visit 5	Y(75)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
163[$]	Visit 7	Y(18)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
164	Visit 1	Y(80)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
167	Visit 1	Y(44)	-	Y[a]	-	-	-	-	-	Y	-	Y	Y	-	Y	-
170	Visit 1	Y(58)	-	Y[b]	-	-	-	-	-	Y	-	-	-	Y	-	-
171	Visit 1	GCS	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
173	Visit 1	Y(6)	-	Y	-	Y	-	-	-	Y	-	-	-	-	-	-
174	Visit 1	Y(44)	-	-	-	-	-	-	-	Y	-	-	-	-	Y	-
176	Visit 1	Y*	-	-	-	Y	Y	-	-	Y	-	Y	Y	-	Y	-
178	Visit 1	Y(58)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
180	Visit 1	Y(12)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
181	Visit 1	Y(22)	-	-	-	-	Y	-	-	Y	-	-	-	Y	-	-
182	Visit 1	Y*	-	-	-	-	-	Y	-	Y	-	Y	-	-	Y	-
183	Visit 1	Y(1)	-	Y	-	-	-	-	-	Y	-	-	Y	-	-	-
194	Visit 1	Y(184)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
194[$]	Visit 3	Y(89)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
200	Visit 1	Y(1)	-	-	-	-	-	-	-	Y	-	Y	-	-	Y	-
203	Visit 1	Y(75)	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-

205	Visit 1	Y*	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
209	Visit 1	Y(49)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
212	Visit 1	Y(12)	Y	-	-	-	-	-	-	Y	-	-	-	-	-	-
215	Visit 1	Y*	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
217[$]	Visit 3	Y(25)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
220	Visit 1	Y*	-	Y	Y	Y	Y	-	-	Y	Y	-	Y	-	-	-
224	Visit 1	Y(12)	-	-	-	-	-	-	-	Y	-	-	Y	-	-	-
231	Visit 1	Y(1)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
232	Visit 1	Y(2)	-	Y	-	-	-	Y	-	Y	Y	-	Y	-	-	-
239	Visit 1	N	-	Y	-	-	-	-	-	Y	-	-	-	-	-	-
241	Visit 1	Y(87)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
241	Visit 2	N	-	Y	-	-	-	-	-	-	-	-	-	-	-	-
248	Visit 1	Y(94)	Y	-	Y	Y	-	-	-	Y	-	-	-	-	-	-
250	Visit 1	Y*	Y	-	-	-	-	-	-	Y	-	-	-	-	-	-
261	Visit 1	Y(1)	-	-	-	-	-	-	-	Y	-	-	-	-	-	-
Total GAS infections (received antibiotics on visit)		83 from 72 PID's (24)	11 (9)	21 (15)	9	14	13	7	3	55	9	10	10	7	14	0
Total of PID's administered antibiotics		69	25	44	11	13	6	8	3	43	11	5	7	4	9	1

Coug, Cough; Hoar, Hoarseness, Tons. swell, Tonsillar swelling; Tons. eryth, Tonsillar erythema; Exud phar / tons, Exudate on the pharynx / tonsils; Temp, Temperature > 38C; Rhin, Rhinorrhea; Ant. cerv, Ant. cervical node > 1.5cm in diameter; Tend. ant, Tender anterior cervical node, Orop. can, Oropharyngeal candidiasis.

N, no GAS recovered; *, GAS *emm* type unknown (untypable); [$], new GAS acquisition, a; received antibiotics after visit; b, received antibiotics but date unknown; c, received antibiotics just before entry into study; GCS, Group C Streptococcus.

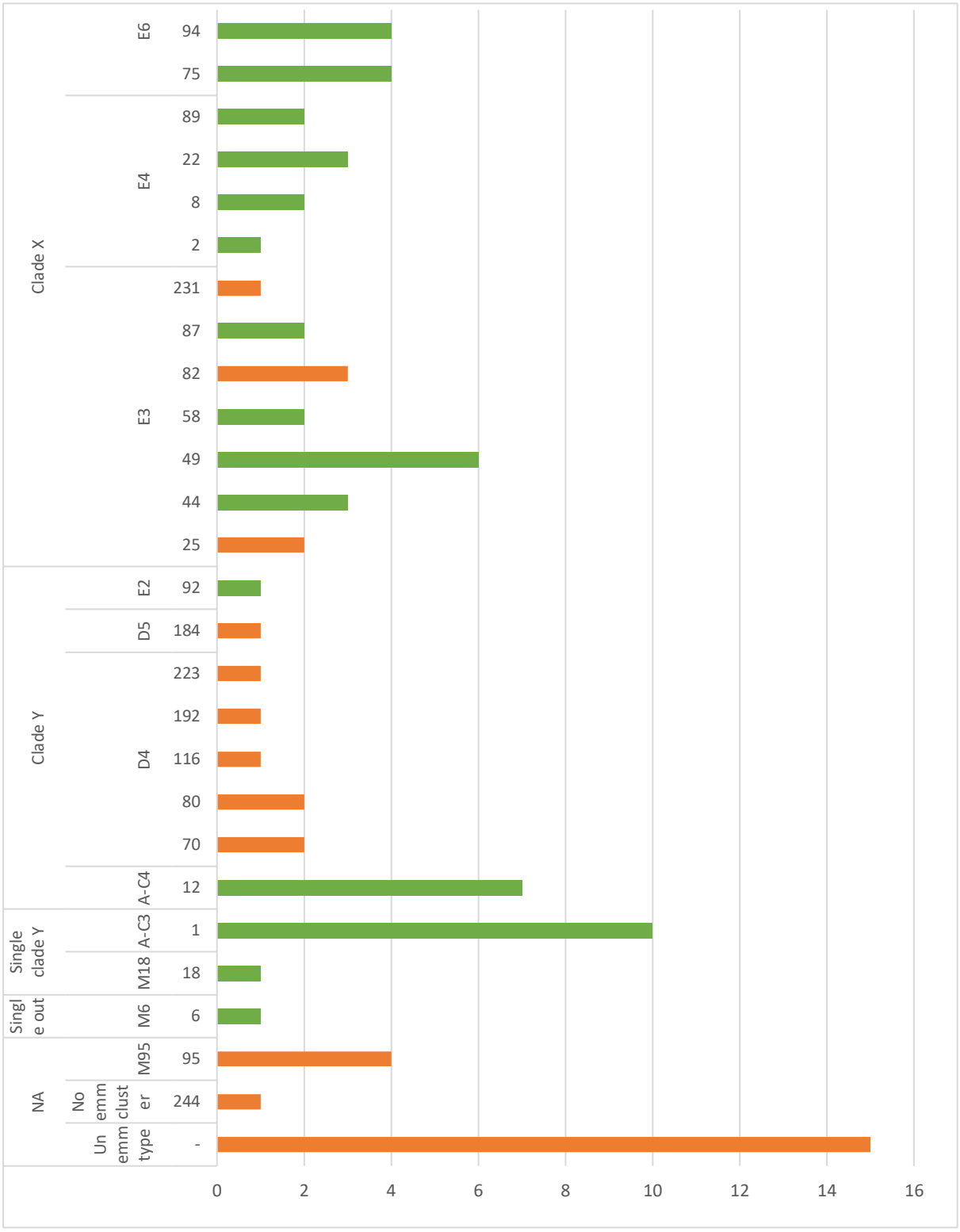
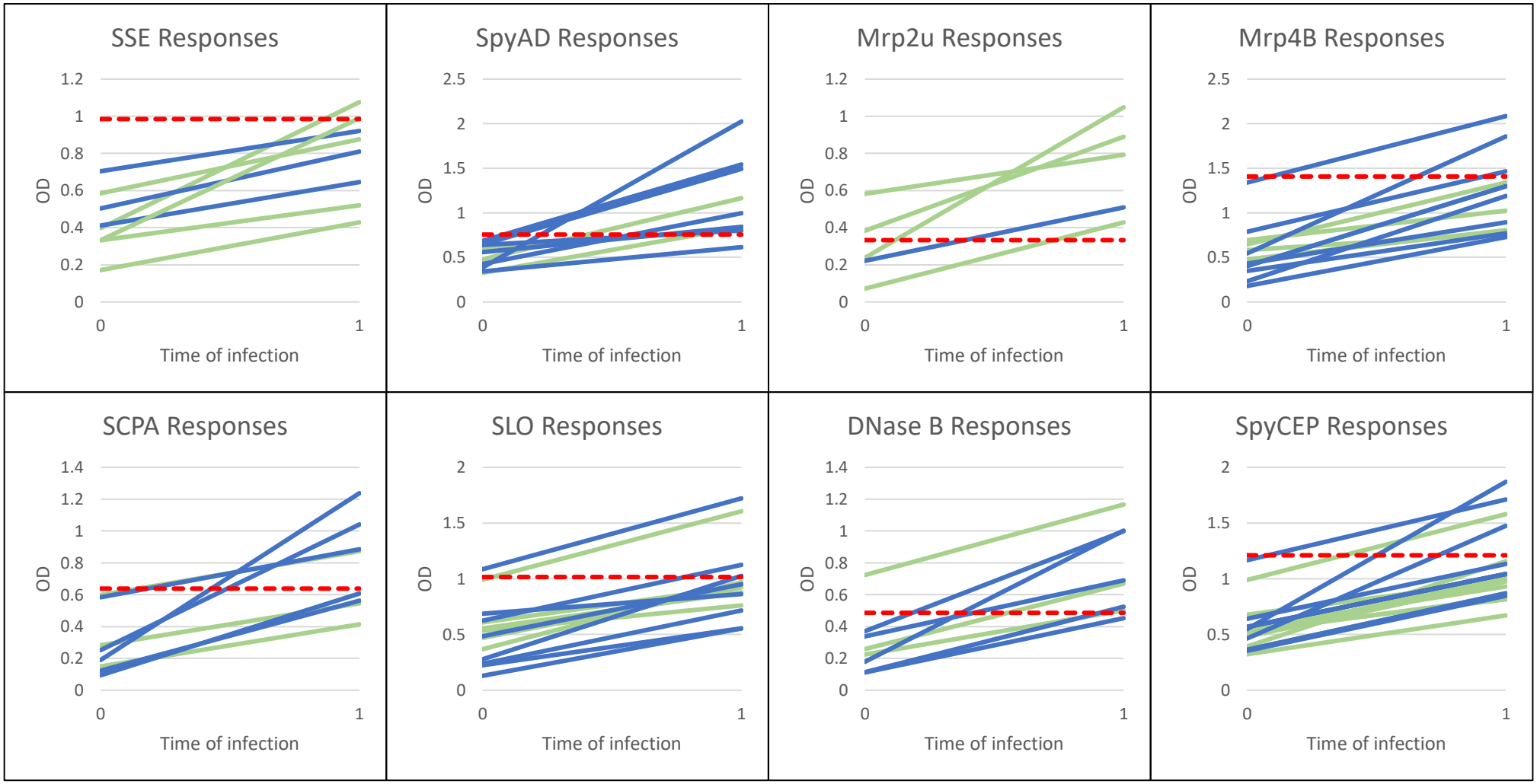


Figure S1. *emm* type distribution with corresponding *emm* cluster and clade. Green bars represent the *emm* types included in the 30-valent vaccine (n=43, 63.2% coverage), with cross-protection afforded to *emm*8 and *emm*94 (n=49, 72.1% coverage). Un *emm* type, unknown *emm* types that includes: untypable (no *emm* sequence) = 9, failed cultures = 4, missing sample = 2; No *emm* cluster, has not been classified according to Sanderson-Smith [2].



Supplementary Figure 2 . Antibody kinetics of participants included in group 1 and group 3, covering the range of increasing titres between the time of infection (a visit and the sequential visit). Group 1 in green and group 3 in blue, ULN in red.