Folliculitis keloidalis nuchae severity score: development and reliability assessment

by

DENNIAS TONDERAI NYIKA

NYKDEN001

SUBMITTED TO THE UNIVERSISTY OF CAPE TOWN

In fulfilment of the requirements for the degree

Master of Medicine in Dermatology

(MMed. Dermatology)

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

Date of submission: 18 May 2020

Supervisor: Professor NP Khumalo

Co-supervisor: Dr T Isaacs

Division of Dermatology

University of Cape Town

DECLARATION

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

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

Signature: 

Date: 18 May 2020

TABLE OF CONTENTS

ABSTRACT

Background: Folliculitis keloidalis nuchae (FKN) is a chronic inflammatory condition that targets the hair follicle, leading to keloidal scarring and alopecia. The absence of a severity scoring tool for FKN limits objective assessment of disease progression and response to treatment.

Objectives: To develop and test the reliability of a severity scoring tool for FKN.

Methods: The tool was developed based on lesion type, number, size and distribution on the scalp. An initial pilot period with 2 assessors was followed by the main study that used 78 anonymised and standardised clinical photographs of the back of the scalp. The participants were selected from an ongoing case control study of FKN. The assessors could allocate disease severity in one of 14 categories (with/without inflammation). However, inflammation (especially erythema) can be missed in photographs of pigmented skin. Thus, two groups of analysis were conducted first with all 14 and again with 8 categories (i.e. excluding inflammation).

Assessors were 4 dermatology consultants and 7 registrars, who all independently scored the same anonymised and standardised photographs on two separate occasions, 2 weeks apart.

Results:

Inter-observer standard errors were higher with the 14-category compared to the 8-category analysis for both consultants and registrars. The intraclass correlation coefficient for registrars improved from poor [0.46 (0.36 -0.56)] to good [0.74 (0.68-0.80)] with 14 compared to 8-categories, but stayed the same for consultants [0.82 (0.76 – 0.88) versus 0.81 (0.75 – 0.87)]. Limitations of the study were the use of clinical photographs instead of live participants and the problem that the signs of inflammation may be particularly difficult to judge in pigmented skin.

Conclusion: We developed a severity scoring tool with poor to good reliability which also highlighted the difficulty of perceiving inflammation from clinical photographs. This improved with the seniority of the observer. The 8-category analysis has good reliability for clinical photographs for both junior and senior staff. For live patient care and clinical trials the 14-category version is likely to be more useful, but requires validation.

ACKNOWLEDGEMENTS AND CONTRIBUTIONS

I would like to thank my supervisor, Prof NP Khumalo, my co-supervisor Dr T Isaacs, Prof F Gumedze and Mr W Bhasera, who helped with the statistics and lastly Dr S Jessop for the editing. The photographs which were used for assessment were from a parent ongoing study: the folliculitis keloidalis, clean shave haircuts and the prevalence of blood-borne infections (SHAKA) Study (Faculty of Health Research Ethics HREC REF Number: 784/2015), of which Prof Khumalo is the principal investigator. This work would not have been possible had it not been for their immeasurable input.

LIST OF TABLES

LIST OF ABBREVIATIONS

FKN: folliculitis keloidalis nuchae

HIV: human immunodeficiency virus

HS: hidradenitis suppurativa

IHS4: International Hidradenitis Suppurativa Severity Score System

SHAKA: the folliculitis keloidalis, clean shave haircuts and the prevalence of blood-borne infections study

DLQI: Dermatology life quality index

Nd: YAG; neodymium-doped yttrium aluminium garnet

ANOVA; analysis of variance

ICC; Intraclass correlation coefficient

SE; standard error

STROBE; The Strengthening in Reporting of Observational studies in Epidemiology

HREC; Health Research Ethics Committee

CHAPTER 1: Introduction and literature review

1.1 Background

Folliculitis keloidalis nuchae (FKN) is a poorly understood chronic disorder involving inflammation and scarring of the hair follicles with the subsequent development of keloid-like papules and plaques and scarring alopecia[1], characteristically on the lower posterior scalp and nape of the neck. Anatomically, the nape, also known as the nucha, is bordered superiorly by the superior nuchal line and external occipital protuberance, and inferiorly by the seventh cervical vertebrae.[2]FKN is an unsightly condition with a variable course of fluctuating symptoms (itch, discomfort and exudation of pus) which can have psychological sequelae. Some authors refer to FKN as acne keloidalis nuchae, although this is a misnomer as it is not secondary to acne vulgaris. FKN may show inconsistent distribution and may involve other areas of the scalp besides the nucha.[3] If it, however, occurs on other sites on the scalp alternative diagnoses should be considered.

FKN is predominantly common in African males with prevalence ranging from 0.7% - 10.5 %[4-7] with a male to female ratio of approximately 20:1.[8] Even more uncommon is the occurrence of FKN in women of non-African lineage[9] specifically in Latin American women.[10] Various factors, including trauma, inflammation, genetics, in growing hairs and aberrant immune responses have been suggested as potential causes of FKN. Keratosis follicularis spinulosa decalvans, a rare X-linked disorder, has been reported to be associated with folliculitis keloidalis nuchae.[11] Treatment of FKN is largely unsatisfactory and focusses on avoidance of exacerbating factors, medical management of inflammation with oral antibiotics, topical steroids, intralesional steroids and sometimes retinoids, in various combinations.[12] Surgery may be used, but, unfortunately, recurrences still occur. Other recently reported treatment modalities include the use of long –pulsed Alexandrite laser, which improved the FKN itself and Dermatology Life Quality Index (DLQI) with no recurrences reported after a follow up of 3 months in 17 male patients. This is comparable to other forms of laser e.g. 1064-nm Nd:YAG and 810-nm diode laser. [13] A case of FKN treated with radiotherapy, with reported excellent cosmetic results, was reported.[14] A prospective, randomised, split-scalp comparison study with 11 patients, with one side using targeted ultraviolet B phototherapy, reported significant clinical improvement and the treatment was tolerated well.[15]

A case of spontaneously resolving FKN was reported, in which triamcinolone and antibiotics were administered in a professional American football player with little improvement, but during the patient's off season, the condition spontaneously resolved.[16]

1.2 Scoring systems for FKN

There is no known severity scoring system for FKN in the published literature as far as the authors have established.

A universally accepted scoring system would be desirable for both research and patient care. In the clinical setting, the scoring system might aid in therapeutic decisions as well as serving as an objective basis for assessing treatment effectiveness.

In the research setting, the scoring system could enhance consistency across study sites in the determination of baseline disease severity and to assess treatment effectiveness. The scoring tool could serve as an objective basis for interpreting and comparing published results from different clinical trials and pooling of results for systematic reviews and meta-analyses.

1.3 An Ideal scoring system

Systemic assessment of the severity of FKN continues to challenge the clinician. An ideal severity scoring system [17] would

- Be accurate and reproducible
- Be capable of documentation for future verification
- Be simple to use over serial clinic visits.
- Take little time to conduct.
- Be inexpensive
- Include subjective criteria, i.e., psychosocial factors e.g. quality of life

Furthermore, the ideal severity scoring system should specifically have[18]

- A limited number of levels so as not to be too cumbersome and impractical for use.
- Levels which are sufficiently described so as to limit intra- and inter-observer variability.
- Levels which indicate when treatment is no longer needed or when maintenance therapy should be undertaken e.g. “clear” (no folliculitis keloidalis) or “almost clear.” i.e. to guide treatment.
- Universality for clinical and investigational use.
- A high degree of correlation with lesion counts.[19]

1.4 Methods of assessment of FKN

In simple terms, an ideal grading system should be accurate, easy to use and quick. Since there is no known scoring system for FKN, the clinical knowledge of FKN presentation and an insight into various scoring systems for diseases which are folliculo-occlusive and seem to have the same pleomorphic properties was done, even though FKN is not a folliculo-occlusive disorder. The conditions which were considered are acne[18, 20-22] and hidradenitis suppurativa (HS). [23-25] The possible methods of assessment of FKN include:

a) Lesion counting

This is considered a precise and objective method. This method distinguishes small differences in therapeutic responses. It, however, has drawbacks of being very slow, and not accounting for distribution, size, erythema or underlying skin quality. In addition, it depends on good lighting as well as the visual acuity of the assessor.[19, 26, 27]

Lesion counting has been modified in some scoring systems e.g. the International Hidradenitis Suppurativa Severity Score System(IHS4).[25] In this modification specific types of lesions are given certain severity index scores, which are then multiplied by the number of lesions present. The adding of each product produces a total score. The total score is then graded as mild, moderate or severe. Critics point out that subjective scores assigned to lesions are non-parametric data, while absolute counts are parametric data and it is probably wrong to mix the two types of data.[26] This method is mainly used in clinical trials.

b) Global severity grading

This method entails the comparison of a patient's FKN condition to a text description or photographs by a reviewer or clinician. This method has advantages, particularly in the clinical setting, as it is practical, fast and easy to use. [28]The disadvantages are that it may be subjective and might be too simplistic to provide sufficient detail of the condition in question. A photographic grading system tends to be better than the older text-only grading system since it is more visual.[18] A digital photographic grading scale for assessing FKN would have the advantages of it being simple, inexpensive and reproducible. The disadvantages would be

- It would not allow palpation of lesions
- Small lesions might not be seen
- It might be difficult to maintain constant lighting, distance between patient and camera and developing procedure[24]. [26]
- Cameras come with different specifications and settings depending on their make

A combination of photographs with explanatory text could be used. If the photographs are from white skinned individual it might be difficult to apply them to dark skinned individuals or vice versa i.e. the measuring tool might be insensitive to change[27]. This method is mainly used in clinical settings.

c) Self-assessment

This is largely an unreliable method of assessment though it tends to positively correlate with the quality of life and is desirable for clinical trials.[29]

d) Multi-modal digital imaging

This is widely considered to be an objective method of assessment and may rely on purpose-built equipment which might be fixed in nature.[30] Computer algorithms and specialist photographic equipment are used to capture and analyse lesion types, skin texture distribution, erythema and pigment changes.[31] Examples incorporate the use of fluorescence photography, polarised light, ultraviolet A lamps and digital cameras. Multi-modal digital imaging has the potential to eliminate tedious manual lesion counting. In the setting of Africa, where FKN is prevalent, cost could be a prohibiting factor for such advanced multi modal imaging at the expense of other health needs, which may be deemed to be more pressing.

Lesion counting is useful in an investigational setting but is of questionable value and practicality in the clinical setting, when compared with grading.[20, 32] On the other hand, the efficacy of treatment on small individual lesions cannot be ascertained using grading and grading is less accurate overall than lesion counting. Photographic grading scales run the risk of becoming obsolete with time as digital technology is currently improving remarkably, with a higher resolution of digital images now being produced, even more so with the advent of multi-modal digital imaging.

e) The DLQI is a measure designed to quantify the impact of a skin problem on one's life, used in adults.[33] It could be incorporated into an FKN severity scoring system, but it is cumbersome, particularly in a clinical situation. Until there is a reliable and validated scoring tool, the psychosocial impact of FKN may need to be documented in the patient's record elsewhere.

1.5 Criteria to assess quality of a FKN scoring tool

It is very difficult to develop a scoring tool that can be validated for a condition which is pleomorphic in nature, has variable characteristics of inflammatory lesions and variable course progression, as in this case with FKN. An example of such a condition with similar or even more difficulties is acne, for which there are many published grading scales but none have been satisfactorily validated. To date, a scoring system for FKN has not been published.

The assessment of the quality of a scoring tool will depend on whether it is intended for clinical practice, investigational setting or both. Agnew et al came up with criteria for assessing published acne scoring tools, for use in research. It had 2 broad assessment categories of a) properties of the scale and b) suitability for use in research and evaluation.[27] The authors feel these properties may also be used to develop an ideal FKN scoring tool.

a) Properties of the tool

i. There should be a record of symptoms like itch, pain and psychosocial impact.
ii. Reliability of the scoring tool should be tested in the form of inter-rater and intrarater reliability using analysis of variance (ANOVA). Intra-rater reliability is measured traditionally in a test/ retest design e.g. using an image and then re-assessing the image a week later. Otherwise assessment of a patient occurs on the same day, hours apart such that the same patient is seen and not different presentations of the same condition.[34] This will be done to check if there are any significant differences between the scoring across individuals within each group. Inter-rater and intra-rater reliability may be measured using intraclass correlation coefficient (ICC).There are 10 forms of ICCs that involve distinct assumptions in their calculations leading to different interpretations and that can yield different results when applied to the same set of data.[35] It is therefore imperative that the correct form of ICC for reliability analysis should be specified as it has implications on the results.[36] Standard error of measurement (SEM) can be similarly used as an ICC in cases that include multiple measurements with multiple observers but without the limitation of being sensitive to data range.[37]
iii. The tool should be validated.

iv. The tool should be sensitive to change e.g.one group will be compliant to treatment whilst the other group is not and this should be discernible on the scale.

b) The signs of FKN should be in the tool, that is, the presence of keloid-like papules and nodular plaques with or without inflammation

c) Suitability for use

i. The tool should be easy to use, with raters being asked about their comfort in using the scoring tool.
ii. An appropriate test population should be used e.g. a general random population as opposed to a narrow sample recruited from a clinic. A clinic sample might limit generalisability of findings and transferability of the scale to another clinical or research setting.
iii. The tool should not only be evaluated by an expert dermatologist, but also by a primary healthcare practitioner.

1.6 Summary

FKN affects a significant number of people, especially African males. It is a frustrating condition, both for the clinician and the patient, with the easily available treatment options producing mostly unsatisfactory outcomes. Currently, there is no severity scoring system for FKN. The increased research into molecular, histopathological and clinical aspects of FKN highlights the need for a validated, easy to use severity scoring tool that can be used in both clinical trials and daily clinical practice to accurately classify severity and guide therapeutic strategy. The scoring tool could be used to pool results from different clinical trials for systematic reviews and meta-analyses. This would help to ensure that patients are not over treated or undertreated.

The aim and objectives of this study were to develop a severity scoring system for FKN and to assess its reliability. The FKN scoring system is designed to be used by dermatologists, doctors and nurses.

1.7 References

1. Mackay-Wiggan JM HM. Acne keloidalis nuchae. UpToDate2020.

2. S S. Gray's Anatomy: The anatomical basis of clinical practice. 41st ed: Elsevier; 2016. 1584 p.
3. Wu WY, Otberg N, McElwee KJ, Shapiro J. Diagnosis and management of primary cicatricial alopecia: part II. Skinmed. 2008;7(2):78-83.
4. Ogunbiyi A, Adedokun B. Perceived aetiological factors of folliculitis keloidalis nuchae (acne keloidalis) and treatment options among Nigerian men. Br J Dermatol. 2015;173 Suppl 2:22-5.
5. Khumalo NP. Folliculitis keloidalis nuchea, bleeding from haircuts, and potential HIV transmission. International Journal of Dermatology. 2012;51:21-3.
6. Ogunbiyi A. Acne keloidalis nuchae: prevalence, impact, and management challenges. Clin Cosmet Investig Dermatol. 2016;9:483-9.
7. Khumalo NP JS, Gumedze F, Ehrlich R. Hair dressing is associated with scalp disease in African school children. Br J Dermatol. 2007(157):106-10.
8. Al Aboud DM, Badri T. Acne, Keloidalis Nuchae. StatPearls. Treasure Island (FL)2017.
9. Loayza E, Cazar T, Uraga V, Lubkov A, Garces JC. Acne keloidalis nuchae in Latin American women. Int J Dermatol. 2015;54(5):e183-5.

10. Loayza E, Vanegas E, Cherrez A, Cherrez Ojeda I. Acne keloidalis nuchae in Latin America: is there a different phenotype? Int J Dermatol. 2017;56(12):1469-70.
11. Goh MS, Magee J, Chong AH. Keratosis follicularis spinulosa decalvans and acne keloidalis nuchae. Australas J Dermatol. 2005;46(4):257-60.
12. Maranda EL, Simmons BJ, Nguyen AH, Lim VM, Keri JE. Treatment of Acne Keloidalis Nuchae: A Systematic Review of the Literature. Dermatol Ther (Heidelb). 2016;6(3):363-78.
13. Tawfik A, Osman MA, Rashwan I. A Novel Treatment of Acne Keloidalis Nuchae by Long-Pulsed Alexandrite Laser. Dermatol Surg. 2017.
14. Millan-Cayetano JF, Repiso-Jimenez JB, Del Boz J, de Troya-Martin M. Refractory acne keloidalis nuchae treated with radiotherapy. Australas J Dermatol. 2017;58(1):e11-e3.
15. Okoye GA, Rainer BM, Leung SG, Suh HS, Kim JH, Nelson AM, et al. Improving acne keloidalis nuchae with targeted ultraviolet B treatment: a prospective, randomized, split-scalp comparison study. Br J Dermatol. 2014;171(5):1156-63.
16. Harris H. Acne keloidalis aggravated by football helmets. Cutis. 1992;50(2):154.
17. Adityan B KR, Thappa DM. Scoring systems in acne vulgaris. Indian J Dermatol Venereol Leprol. 2009:323-6.
18. Dreno B, Poli F, Pawin H, Beylot C, Faure M, Chivot M, et al. Development and evaluation of a Global Acne Severity Scale (GEA Scale) suitable for France and Europe. J Eur Acad Dermatol Venereol. 2011;25(1):43-8.
19. Lehmann HP, Robinson KA, Andrews JS, Holloway V, Goodman SN. Acne therapy: A methodologic review. Journal of the American Academy of Dermatology. 2002;47(2):231-40.
20. Cook CH, Centner RL, Michaels SE. An Acne Grading Method Using Photographic Standards. JAMA Dermatology. 1979;115(5):571-5.
21. Samuelson JS. An accurate photographic method for grading acne: initial use in a double-blind clinical comparison of minocycline and tetracycline. J Am Acad Dermatol. 1985;12(3):461-7.
22. Allen BS, Smith JG, Jr. Various Parameters for Grading Acne Vulgaris. JAMA Dermatology. 1982;118(1):23-5.
23. Hurley H. Hidradenitis suppurativa. Roenigk & Roenigk's Dermatologic surgery: Principles and practice. 2nd ed. New york: Marcel Dekker, Inc; 1996. p. 623-45.
24. Sartorius K, Lapins J, Emtestam L, Jemec GBE. Suggestions for uniform outcome variables when reporting treatment effects in hidradenitis suppurativa. 2003;149(1):211-3.
25. Zouboulis CC, Tzellos T, Kyrgidis A, Jemec GBE, Bechara FG, Giamarellos-Bourboulis EJ, et al. Development and validation of the International Hidradenitis Suppurativa Severity Score System (IHS4), a novel dynamic scoring system to assess HS severity 2017 [1401-9]. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/bjd.15748

https://onlinelibrary.wiley.com/doi/full/10.1111/bjd.15748.
26. Adityan B, Kumari R, Thappa DM. Scoring systems in acne vulgaris. Indian J Dermatol Venereol Leprol. 2009;75(3):323-6.
27. Agnew T, Furber G, Leach M, Segal L. A Comprehensive Critique and Review of Published Measures of Acne Severity. J Clin Aesthet Dermatol. 2016;9(7):40-52.
28. Tan JKL. Current measures for the evaluation of acne severity. Expert Review of Dermatology. 2008;3(5):595-603.
29. Menon C, Gipson K, Bowe WP, Hoffstad OJ, Margolis DJ. Validity of subject self-report for acne. Dermatology. 2008;217(2):164-8.
30. Bae Y, Nelson JS, Jung B. Multimodal facial color imaging modality for objective analysis of skin lesions. J Biomed Opt. 2008;13(6):064007-.
31. Patwardhan SV, Kaczvinsky JR, Joa JF, Canfield D. Auto-classification of acne lesions using multimodal imaging. J Drugs Dermatol. 2013;12(7):746-56.
32. Witkowski JA, Parish LC. The assessment of acne: an evaluation of grading and lesion counting in the measurement of acne. Clin Dermatol. 2004;22(5):394-7.

33. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. Clin Exp Dermatol. 1994;19(3):210-6.
34. Lucky AW, Barber BL, Girman CJ, Williams J, Ratterman J, Waldstreicher J. A multirater validation study to assess the reliability of acne lesion counting. J Am Acad Dermatol. 1996;35(4):559-65.
35. McGraw KO WS. Forming inferences about some intraclass correlation coefficients. Psychol Methods. 1996;1:30-46.
36. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of chiropractic medicine. 2016;15(2):155-63.
37. Popovic ZB, Thomas JD. Assessing observer variability: a user's guide. Cardiovasc Diagn Ther. 2017;7(3):317-24.

CHAPTER 2: Publication ready manuscript

FOLLICULITIS KELOIDALIS NUCHAE SEVERITY SCORE: DEVELOPMENT AND RELIABILITY ASSESSMENT

Nyika DT, Isaacs T, Basera W, Gumedze F; Khumalo NP

The Division of Dermatology, Department of Medicine, and the Department of Statistical Sciences, Faculty of Health Sciences, University of Cape Town, South Africa

Corresponding author:
Professor N P Khumalo
Ward G23
Groote Schuur Hospital
Observatory
7925
(n.khumalo@uct.ac.za)

2.1 Background

Folliculitis keloidalis nuchae (FKN) is a chronic disorder involving inflammation and scarring of the hair follicles with the subsequent development of keloid-like papules, plaques and scarring alopecia[1], characteristically on the lower posterior scalp. It predominantly affects African males, with a prevalence ranging from 0.7% to 9.4% in Nigerian clinics[2] increasing to 10.5% in males >18 years old in a Cape Town population study.[3] The disease is less common in other population groups with a prevalence of 0.007% in a single institutional study in the Republic of Korea.[4]

The cause of FKN has not yet been definitely established, though there are theories, with variable levels of evidence, suggesting mechanisms related to skin injury and aberrant immune responses to common antigens.[5] Factors commonly implicated are androgens, inflammation, trauma, genetics and ingrowing hairs.[6]

A positive association between FKN and some aspects of metabolic syndrome, especially hypertension and diabetes, has been reported. [7-9] The clean-shave 'chiskop' hair style, popular among South African black males, was noted to increase the risk of scalp bleeding with 32% of participants reporting episodes of bleeding.[10]

There are no universal treatment guidelines for FKN. Treatment is usually guided by clinical experience modified by clinical presentations and patient preferences and availability of medication. Unfortunately, recurrences and incomplete resolution of FKN are very common.[11]

To date, there is no published severity scoring system for FKN. A universally accepted FKN severity scoring system would be useful for clinical practice. Further, the increased research in this field highlights the need for a validated, easy to use severity scoring tool that can be used in clinical trials to enhance consistency across study sites in the determination of baseline disease severity and to assess treatment efficacy. The scoring system would also serve as an objective measure for interpreting and comparing published results from different clinical trials.

A scoring system might be valuable in elucidating the contribution of individual variables to disease pathogenesis. It would allow more accurate clinical follow up of such patients. This would also help to ensure that patients are not subjected to often painful and uncomfortable treatment options (such as intralesional steroid injections) if no clinical improvement is noted, could reduce over treatment and under treatment.

The absence of a FKN severity scoring tool currently prohibits secondary trial data analysis, complicates interpretation of study results and may be compromising patient care.

2.2 Objectives

To develop and test the reliability of a severity scoring tool for FKN.

2.3 Methods and participants

The Strengthening the Reporting of Observational studies in Epidemiology (STROBE)[12] statement was used in the preparation of this manuscript.

The study was conducted at the Division of Dermatology, Groote Schuur Hospital, an academic tertiary institution in Cape Town, South Africa.

2.3.1 FKN scoring system

Scoring systems for acne [13-16] and hidradenitis suppurativa (HS) [17, 18], which are both follicular-occlusive conditions, were reviewed **(Table 1a** and **Table 1b)**. These scoring systems were chosen for review because, like acne and HS, FKN presents with pustules, papules and nodular plaques of different sizes and distribution. However, there is no single universally accepted gold-standard scoring system for either acne or HS. Although FKN is not a follicular-occlusive disorder, reviewing these scoring systems is likely to be useful. Finally, lesion types, size, their number and distribution on the scalp were used to develop the FKN severity scoring tool. In this study, photographs rather than patients were assessed. It was therefore not possible to include a psychometric component, or a measure of quality of life, such as the Dermatology Life Quality Index (DLQI). .

Keloidal papules were defined as circumscribed raised lesions less than 1cm in diameter, with keloidal plaques being greater than 1cm in diameter. A pustule was defined as a pus-filled lesion less than 1cm in diameter. Inflammation was defined as the presence of pustules and/or perifollicular erythema and/or scale. The presence of transient looking tiny papules, pustules or macular (flat) nuchal scalp lesions like erythema, scale, and pigmentary changes were noted as abnormal but not FKN. After discussion and a pilot study on 22 photographs, done independently by 2 of the authors (TI and NPK), the FKN severity scoring tool was subjected to formal evaluation.

Seventy-eight anonymised and standardised clinical photographs of the back of the scalp on a black background were chosen from an already existing parent study: the folliculitis keloidalis, clean shave haircuts and the prevalence of blood-borne infections (SHAKA) Study (Faculty of Health Research Ethics (HREC) REF Number: 784/2015). Additional ethics committee permission was granted for use of clinical photographs for the current study (HREC REF Number: 781/2019).

2.3.2 Panellists

We asked all medical members of our department to participate as panellists. The panel consisted of dermatology consultants and registrars from the Division of Dermatology, Groote Schuur Hospital and the University of Cape Town. The panellists had to be available on 2 separate days, a fortnight apart, on which the scoring of the exact photographs in the same order was done. The panel consisted of 2 groups of participants. The first group was made up of 7 dermatology registrars with dermatology experience ranging from 1 to 3 years. The second group was composed of 4 consultants in the first and 3 in the second meeting. The lead author was not part of the panel but he collated the data for statistical analysis. The senior author, who has experience in developing alopecia tools, attended the first meeting but could not attend the second meeting. Panellists were all given a cover letter, background information about FKN and instructions on how to use the tool before the scoring of the photographs. Questions about the scoring were clarified to their satisfaction. The training process at the first meeting lasted about 30minutes.The panellists were informed that the forms were blinded for names and that their confidentiality was assured. The photographs were projected onto one big screen and the panellists scored them independently, using the proposed FKN scoring system. At the second scoring meeting the same number of photographs were projected in the same manner. Panellists sat in the same positions relative to the projected screen that they had occupied a fortnight before. The scoring of the photographs was done on 31 October 2019 and 14 November 2019.

2.3.3 Statistical analysis

Sample size calculation in variability studies has 2 variables in the equation i.e. the number of subjects and the number of observations per subjects.[19] In this study we set a 95% confidence interval, with 11 panellists to come up with the sample size of the 78 photographs to be scored.

Two-way random-effects model was used to calculate the ICC. This model was appropriate to evaluate the reliability of this particular scoring system as it had two groups of panellists with varying degrees of dermatology experience. Inter-observer and intra-observer standard error (SE) were estimated using analysis of variance (ANOVA)

The FKN scoring tool was used to grade the observations from 1 to 14, with the 1st category being normal (0-N), the 2nd Abnormal but not FKN; the rest were based on lesion count, size, distribution and evidence of inflammation to give a 14-category score (**Table 2a)**. A second analysis, which excluded inflammation scores, was done. It reduced the tool from a 14 to an 8-category tool (**Table 2b)**.

The ICC estimates were grouped as follows:[20] poor reliability :< 0.5, moderate reliability: 0.5-0.75, good reliability: 0.75-0.90 and excellent reliability :> 0.90. Statistical analysis was done using Genstat (20th edition)[21] and STATA (Version15)[22] software using the user-written packages.

2.4 Results

Eleven panellists had the opportunity to score all 78 photographs at the first meeting and 10 scored for the second time 2 weeks later. 8 assessments, which had missing or illegible scores were excluded. We observed that within each study analysis there was a higher inter-observer SE but a relatively small intra observer SE. Higher inter-observer standard errors were noted with the 14-category compared to the 8-category analysis for both consultants and registrars. Overall standard error values were 0.35 (0.20) and 0.11 (0.05) for the 14-category and 8-category analysis respectively. Similarly, the registrar, consultant and overall intra-observer SE were lower with the 8-category analysis. Although the between subject SE were high, the trend was similarly maintained with much lower values for the 8-category analysis (Table 3a). The inter-observer agreement based on the ICC was poor for registrars with the 14-category analysis at 0.46 (0.36 -0.56) and improved to very good at 0.74 (0.68-0.80) with the 8-category analysis. Interestingly the consultants ICC did not differ with the 14-category at 0.82 (0.76 – 0.88) compared to the 8-category analysis at 0.81 (0.75 – 0.87).

Table 3a: Inter-observer, intra-observer and between subject variability of scores

Inter-observer (Standard Error)				Intra-observer (Standard Error)			Between subject (Standard Error)		
Analysis	**Overall**	**Consultants**	**Registrars**	**Overall**	**Consultants**	**Registrars**	**Overall**	**Consultants**	**Registrars**
14-category	0.35 (0.20)	0.44 (0.38)	0.36 (0.27)	0.01 (0.01)	0.03 (0.02)	0.01 (0.01)	14.84 (2.49)	14.16 (2.34)	15.42 (2.69)
8-category	0.11 (0.05)	0.15 (0.13)	0.10 (0.06)	0.02 (0.01)	0.03 (0.02)	0.02 (0.01)	4.01 (0.66)	4.09 (0.68)	4.00 (0.66)

Inter-observer and intra-observer variabilities were estimated using: ANOVA, Standard Error in parenthesis

Table 3b: Measurements of agreement by rater category

ICC (95% CI) *			
Analysis	**Overall**	**Consultants**	**Registrars**
14-category	0.53 (0.45, 0.61)	0.82 (0.76, 0.88)	0.46 (0.36, 0.56)
8-category	0.76 (0.70, 0.82)	0.81 (0.75, 0.87)	0.74 (0.68, 0.80)

*ICC, intraclass correlation coefficients were calculated to assess the agreement of evaluators in the classification of FKN using a scoring system

2.5 Discussion

There is no known universally accepted FKN scoring system. The authors have developed and tested the reliability of a FKN scoring tool and have done a secondary analysis of the tool. This tool could be vital in guiding therapeutic decisions as well as for use in clinical research studies. We developed a 14 category FKN scoring tool which had poor and good reliability amongst dermatology registrars and consultants respectively, which, when adapted to a 8-category scoring tool, had good reliability for both dermatology registrars and consultants. This possibly highlighted the difficulty in perceiving inflammation on pictures of dark skinned individuals. It should be noted that this scoring system is not a diagnostic tool and a clinician will still be required to make a diagnosis of FKN.

In people with higher Fitzpatrick skin types (4-6), in which FKN is most prevalent, inflammatory components, particularly erythema, may be difficult to appreciate in photographs, especially for those with little dermatological experience or exposure to such patients. This working knowledge prompted us to analyse the FKN scoring tool as 8-categories, which excluded inflammation. The 8-category tool had better inter-observer variability and ultimately reliability, regardless of the level of experience of the rater. Not only does skin type affect the appreciation of inflammation but other factors like visual acuity of the rater, distance between projected picture and rater and the general lighting of the room may also contribute. We tried to minimise the effects of these confounders by making sure the ratings were done at the same time of the day and the raters sat on exactly the same spot on the 2 separate scoring occasions.

In as much as there is no test to compare the ICCs of the 14-category and 8-category scores, a pragmatic approach was to check whether the 95% CI of the overall ICCs overlapped. The 95% CI for the overall ICC for the 2 separate scores did not overlap and this meant that there is a difference between the 2 scores. Had the CI overlapped, it would have meant that the 2 scoring systems do not differ.

Problems which were expected and were encountered arose from the pleomorphic nature of FKN in the context of lesion count, the variable characteristics of inflammatory lesions, and the site of involvement on the nuchal area. Lesion counting is good in an investigational setting but may have questionable value and practicality in the clinical setting.[14]Delineating which area of the scalp was nuchal and picking out the presence of inflammation proved difficult in some of the photographs used.

The Dermatology Life Quality Index (DLQI) seeks to quantify how much a skin problem has affected one's life in a time period in adults.[23] It is a questionnaire which can be handed to patients and they can complete it in 1-2 minutes. It would have helped quantify the experiences of people living with FKN. It could, however, not be incorporated into the FKN severity scoring tools since photographs were assessed, and not live patients.

Reliability reflects the extent to which measurements may be replicated as well as the degree of correlation and agreement between measurements.[24] Based on the reliability of the FKN severity scoring tool amongst the categories tested, the authors suggest that its reliability increases with an increase in the level of dermatology experience, although this finding was not statistically significant. This was not the case with the 8-category scoring.

.

There are other historical ways of measuring reliability e.g. Pearson correlation coefficient, paired t-test and Bland-Altman plots. The Pearson correlation coefficient measures correlation only, while the paired t-test and Bland-Altman plots analyse agreement only. ICC is more reliable since it reflects both the degree of correlation and agreement between measurements.[20] Although ICC is frequently reported in reliability studies it has a significant weakness of being sensitive to data range. It is lower in measurements with a narrow range higher in measurements with a wide range of measurements.[25]

These statistical methods have been transferred into medicine by using a 2 way ANOVA to calculate intra and inter-observer standard error of measurement (SE).[26] The SE can be similarly used as an ICC in cases that include multiple measurements with multiple observers but without the limitation of being sensitive to data range. For our study, we reported both the ICC and the SE .The SE may be used particularly with regards to it as having a reportedly lower sensitive to data range. The inter-observer SE was higher in comparison to the intra observer SE. It was also lower in the adapted 8 point scoring system in comparison to the 14 point scoring system. This was in keeping with the calculated ICC for the same data set.

This FKN severity scoring tool is fairly easy to use and only concentrates on the clinical features present on the nuchal area. Such a tool could be invaluable in the treatment of FKN as it tries to avoid harm due to the underestimation of the severity of FKN patients or over treating FKN patients if the severity is overestimated.

As with other skin conditions like acne and hidradenitis suppurativa, whose scoring tools we looked at, the polymorphic nature of FKN makes it difficult to assess the severity. Acne has no universal severity assessment.[27] Hidradenitis suppurativa has, however, newer severity assessment tools, which have been validated.[17]

It would have been ideal to have the scoring system used by patients' first point of care health personnel, including nurses and general practitioners. As its reliability seems to improve with dermatology experience, it may not be transferral to the Primary care setting. The inter-rater variability measures agreement among a certain group of raters and the intra-rater variability measures the consistency of the rating. It has been argued that consistency is better than agreement in reliability studies.[20] In our study the intra-rater variability was lower than the inter-rater variability. The global assessment of a patient's condition, the DQLI, will also be an essential addition to our current severity scoring system.

There are limitations to the study:

- Photographs were used for scoring instead of actual patients. The disadvantages of photography are that small lesions may not be visualised and maintaining constant lighting, distance between patient and camera may be difficult.[27] Photographs were, however, standardised for this study.
- It would have been ideal to have the scoring system used by patients' first point of care health personnel including nurses and general practitioners, but its reliability seems to improve with dermatology experience and it can be criticised for that.
- The global assessment of a patient's condition, the DQLI, will also be an essential addition to our current severity scoring system.

In summary we have developed and tested the reliability of the first FKN severity scoring tool. Although the 8-category scores has better reliability in comparison to the 14-category scores, clinical notes of inflammation from the 14-category version are more likely to be useful and has application in both clinical and investigational settings. It is fairly easy to use, requiring lesion counting and clinical assessment of pustules, keloidal papules and nodular plaques. The reliability of the FKN severity scoring system was tested in a single dermatology centre. Further validation studies are needed to assess and improve the utility of this tool in the clinical setting.

Table 1 – scoring systems that were reviewed prior to development of the current one

Table 1a Summary of acne grading systems

Acne grading system	Method	Special equipment needed
Global acne grading system[13]	Comedones , papules, pustules and nodules graded from 1-4 respectively, then multiplied by predetermined severity factors of the anatomical location of the face and back with a global score of maximum of 44	none
Cook et al[14]	Grading from 0-8 using photographic standards of comedones, papules, pustules and cysts/sinuses	Photography
Samuelson[15]	Requires both the patient and physician to assess the severity based on a set of reference photographs on a nine grade scale	Photography
Allen and Smith[16]	A photonumeric method-both grading using photographic standards and lesion counting done	Photography

Table 1b Summary of hidradenitis scoring systems

Scoring system	Method
Hidradenitis Suppurativa severity score system(IHS4)[17]	Papules, Nodules, abscesses and fistulae/sinuses given values of 0, 1, 2 and 4 respectively which gets multiplied by the number of lesions present. Mild =<3, mod 4-10, severe>11 points
Sartorius score[18]	Counting involved regions, nodules and sinus tracts(3 points/region).Number and scores of lesions 'other' =1, nodule =2, fistula =4.The distance between relevant lesions is also scored from 2-8(<5cm=2,<10cm=4,>10cm = 10), and if the lesions are clearly separated by normal skin(yes=0,no=6)
Hurley staging[28]	Abscesses, tunnelling/ sinus tract formation and scarring given score from 1-3 respectively

Table 2a. Folliculitis Keloidalis Nuchae (FKN) Severity Grading – 14 category analysis

14-catag.	Grade	Description
1	0 - N	**Normal scalp**
2	0 - AbN	**Nuchal scalp changes not FKN** Erythema, scale, pigment change, transient papules/pustules, up to 3 keloidal papules
	FKN1	**>3 keloidal papules**
3	A	i)Non inflamed nuchal papules
4		ii)Non inflamed nuchal papules +beyond nuchal area
5	B	i)Inflamed nuchal papules+/-pustules
6		ii) Inflamed nuchal papules+/-pustules/ + beyond nuchal area
	FKN2	**Small nodular plaques(<3cm)**
7	A	i)Non inflamed nuchal nodular plaques (n≤3)
8		ii)Non inflamed multiple (n>3)nodules +/- papules/nodules beyond the nuchal area
9	B	i)Inflamed nuchal nodular plaques (n≤3)
10		ii)Inflamed nuchal nodular plaques (n>3) +/- papules/nodules beyond the nuchal area
	FKN3	**Large nodular plaques(>3cm)**
11	A	i)Non inflamed nodular plaques confined to nucha
12		ii)Non inflamed nuchal nodular plaques + papules/nodules beyond nucha
13	B	i)Inflamed nodular plaques confined to the nucha
14		ii)Inflamed nodular plaques +papules/nodules beyond the nucha

*Keloidal papules = circumscribed palpable lesions less than 1cm in dimension,

*Nodular plaques = circumscribed palpable lesions greater than 1cm in dimension.

*Pustule = a pus-filled lesion less than 1cm in dimension.

*Inflammation = the presence of erythema and scaling.

Table 2b Folliculitis Keloidalis Nuchae (FKN) Severity Grading – 8-category analysis

1	**0 – N**	**Normal scalp**
2	**0 - AbN**	**Nuchal scalp changes not FKN:** Scale, pigment change, transient papules/pustules, up to 3 keloidal papules
	FKN1	**>3 keloidal papules**
3		i)Nuchal papules
4		ii)Nuchal papules +beyond nuchal area
	FKN2	**Small nodular plaques(<3cm)**
5		i)Nuchal nodular plaques (n≤3)
6		ii) Nuchal nodular plaques (n>3) +/- papules/nodules beyond the nuchal area
	FKN3	**Large nodular plaques(>3cm)**
7		i)Nodular plaques confined to nucha
8		ii)Nuchal nodular plaques + papules/nodules beyond nucha

*Keloidal papules = circumscribed palpable lesions less than 1cm in dimension,

*Nodular plaques = circumscribed palpable lesions greater than 1cm in dimension.

*Pustule = a pus-filled lesion less than 1cm in dimension.

2.6 References

1. Mackay-Wiggan JM HM. Acne keloidalis nuchae. UpToDate2020.
2. Ogunbiyi A. Acne keloidalis nuchae: prevalence, impact, and management challenges. Clin Cosmet Investig Dermatol. 2016;9:483-9.
3. Khumalo NP JS, Gumedze F, Ehrlich R. Hair dressing is associated with scalp disease in African school children. Br J Dermatol. 2007(157):106-10.
4. Na K, Oh SH, Kim SK. Acne keloidalis nuchae in Asian: A single institutional experience. PLoS One. 2017;12(12):e0189790.
5. Kelly AP. Pseudofolliculitis barbae and acne keloidalis nuchae. Dermatol Clin. 2003;21(4):645-53.
6. Ogunbiyi A, Adedokun B. Perceived aetiological factors of folliculitis keloidalis nuchae (acne keloidalis) and treatment options among Nigerian men. Br J Dermatol. 2015;173 Suppl 2:22-5.
7. Loayza E, Vanegas E, Cherrez A, Cherrez Ojeda I. Acne keloidalis nuchae in Latin America: is there a different phenotype? Int J Dermatol. 2017;56(12):1469-70.
8. East-Innis ADC, Stylianou K, Paolino A, Ho JD. Acne keloidalis nuchae: risk factors and associated disorders - a retrospective study. Int J Dermatol. 2017;56(8):828-32.
9. Adotama P, Rutherford A, Glass DA, 2nd. Association of keloids with systemic medical conditions: a retrospective analysis. Int J Dermatol. 2016;55(1):e38-40.
10. Khumalo NP. Folliculitis keloidalis nuchea, bleeding from haircuts, and potential HIV transmission. International Journal of Dermatology. 2012;51:21-3.
11. Maranda EL, Simmons BJ, Nguyen AH, Lim VM, Keri JE. Treatment of Acne Keloidalis Nuchae: A Systematic Review of the Literature. Dermatol Ther (Heidelb). 2016;6(3):363-78.
12. Hollestein LM NT. Guidelines for statistical reporting in the *British Journal Of Dermatology*. Br J Dermatol. 2015(173):3-5.
13. Doshi A, Zaheer A, Stiller MJ. A comparison of current acne grading systems and proposal of a novel system. 1997;36(6):416-8.
14. Cook CH, Centner RL, Michaels SE. An Acne Grading Method Using Photographic Standards. JAMA Dermatology. 1979;115(5):571-5.
15. Samuelson JS. An accurate photographic method for grading acne: initial use in a double-blind clinical comparison of minocycline and tetracycline. J Am Acad Dermatol. 1985;12(3):461-7.
16. Allen BS, Smith JG, Jr. Various Parameters for Grading Acne Vulgaris. JAMA Dermatology. 1982;118(1):23-5.
17. Zouboulis CC, Tzellos T, Kyrgidis A, Jemec GBE, Bechara FG, Giamarellos-Bourboulis EJ, et al. Development and validation of the International Hidradenitis Suppurativa Severity Score System (IHS4), a novel dynamic scoring system to assess HS severity 2017 [1401-9]. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/bjd.15748

https://onlinelibrary.wiley.com/doi/full/10.1111/bjd.15748.
18. Sartorius K, Lapins J, Emtestam L, Jemec GBE. Suggestions for uniform outcome variables when reporting treatment effects in hidradenitis suppurativa. 2003;149(1):211-3.
19. Bland. How can I decide the sample size for a repeatability study United Kingdom2010 [updated 17.05.2010; cited 2020 07.01]. Available from: https://www-users.york.ac.uk/~mb55/meas/sizerep.htm.
20. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of chiropractic medicine. 2016;15(2):155-63.
21. International V. Genstat for Windows Hemel Hempstead,UK: VSN International; 2019 [20:[
22. StataCorp. Stata statistical software: Release 15. College Station,TX: StataCorp LLC; 2017.
23. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. Clin Exp Dermatol. 1994;19(3):210-6.
24. Bias and Measurement Error. Interpretation and Uses of Medical Statistics. p. 381-421.
25. Popovic ZB, Thomas JD. Assessing observer variability: a user's guide. Cardiovasc Diagn Ther. 2017;7(3):317-24.

26. Eliasziw M, Young SL, Woodbury MG, Fryday-Field K. Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. Physical therapy. 1994;74(8):777-88.
27. Adityan B, Kumari R, Thappa DM. Scoring systems in acne vulgaris. Indian J Dermatol Venereol Leprol. 2009;75(3):323-6.
28. Hurley H. Hidradenitis suppurativa. Roenigk & Roenigk's Dermatologic surgery: Principles and practice. 2nd ed. New york: Marcel Dekker, Inc; 1996. p. 623-45.

2.7 Appendices

2.7.1 Raw data supplements

Supplement 1. First seating scoring

A(R3)	B(R3)	C(R2)	D(CON)	E(R2)	F(CON)	G(R3)	H(R2)	I(CON)	J(CON)	K(R3)
6	6	4	4	4	6	6	4	6	6	6
7	11	6		10	5	10	6	7	10	5
1	1	1	2	2	1	2	1	2	1	3
3	10	5		10	5	5	10	9	10	6
8	12	10		14	10	10	8	11		8
8	14	14		12	8	10	12	13	14	8
4	6	10	6	14	4	6	2	5	6	2
1	1	1	1	1	1	1	1	2	1	1
8	12	14			8	10	12	12	14	6
3	5	5	5	9	3	5	6	6	6	2
13	13	11	11	14	8	9	12	11	13	10
2	2	2	2	2	2	3	4	3	2	3
4	4	8	8	10	3	4	8	9	6	6
14	14	14	8	14	14	14	14	14	14	14
4	4	8	8	4	6	6	8	6	4	6
2	2	8	2	2	2	4	2	4	2	3
1	1	1	2	2	1	2	1	1	1	1
12	4	12	6	11	5	7	12	14	11	4
1	1	2	1	7	1	1	12	2	1	2
3	7	7	3	6	5	6	10	7	6	5
1	2	12	2	4	2	2	1	3	2	1
8	13	10	5	10	5	7	10	13	10	5
7	11	9	7	3	5	9	5	10	9	5
1	1	2	2	2	1	2	1	2	1	1
1	2	1	2	1	1	2	2	2	2	1

3	1	2	2	1	2	1	1	2	1	1
1	1	2	1	2	2	1	1	1	2	4
1	1	1	1	1	1	1	1	2	1	1
1	1	2	1	1	1	1	1	1	1	1
1	1	2	1	1	1	1	1	1	1	1
2	2	1	2	2	2	2	2	2	2	3
6	9	10	12	5	5	5	10	9	10	5
1	1	1	1	2	1	1	2	2	1	2
1	2	1	2	2	7	1	2	2	2	5
1	2	2	1	6	1	1	2	2	1	2
3	3	2	4	8	6	3	3	3	3	6
4	2	12	4	1	2	4	4	2	2	6
2	2	2	2	2	1	2	4	2	2	3
2	2	4	4	2	2	4	2	2	2	6
2	3	3	3	1	3	3	4	2	2	3
1	1	1	1	9	1	1	1	2	1	2
5	10	10	5	12	5	5	10	5	10	6
2	3	12	2	2	6	2	12	5	2	8
1	1	2	1	6	1	1	1	1	2	1
2	3	2	5	4	2	2	4	1	2	3
3	8	10	6	14	5	3	10	13	10	7
2	2	2	2	2	5	3	6	5	2	5
3	6	8	10	4	5	5	3	6	6	10
1	1	2	1	1	1	1	1	2	1	1
11	11	13	13	5	5	9	11	11	10	9
1	1	2	2	2	2	1	7	3	1	1
7	8	9	5	5	5	3	8	5	5	5
13	12	14	10	14	11	13	12	13	11	135
2	2	2	2	2	2	2	2	9	2	13
7	11	13	11	5	7	7	10	11	13	6
11	12	13	11	6	7	7	11	13	9	2

1	1	1	1	1	1	1	1	2	1	9
6	10	10	10	8	6	6	10	9	9	6
2	2	2	3	2	2	3	6	3	2	2
2	2	2	7	2	2	2	2	3	2	1
1	1	1	1	1	1	1	1	3	1	6
2	2	4	3	2	2	4	2	11	2	10
13	13	13	5	10	5	11	12	1	9	1
1	1	1	1	1	2	1	2	1	1	1
1	1	1	1	2	1	1	1	1	1	2
1	2	2	2	2	2	1	2	1	1	1
2	2	2	2	1	2	1	2	2	2	1
1	1	12	1	1	2	1	1	2	1	1
14	14	14	14	14	14	14	14	14	13	14
3	3	8	7	4	3	3	3	5	2	3
11	2	2	2		2	2	2	2	2	11
2	2	2	3	2	5	2	4	3	2	2
10	13	14	13	13	9	13	12	14	13	9
7	11	12	13	11	7	7	12	14	13	7
3	8	4	9	10	3	3	10	10	9	8
12	14	14	13	14	14	13	14	14	14	14
7	13	7	11	5	9	9	9	13	7	11
1	1	1	1	1	1	1	1	1	1	3

Supplement 2.Second seating scoring

A(R3)	B(R3)	C(R2)	D(CON)	E(R2)	F (CON)	G(R3)	H(R2)	I(CON)	J(CON)	K(R3)
6	6	10	5	4	6	6	6	10		6
3	11	7	9	3	5	7	9	13		5
1	1	1	1	1	1	1	2	1		3
3	8	9	5	5	5	3	8	13		8
8	8	12	12	8	10	8	12	12		8
8	14	12	12	14	8	14	14	14		10
4	6	10	6	6	2	6	8	6		6
1	1	1	1	1	1	1	1	1		1
8	14	12	14	14	8	12	14	14		10
5	5	3	5	2	5	3	7	3		5
13	13	12	14	13	10	13	14	12		9
2	2	1	3	1	2	2	2	4		3
3	4	8	8	6	5	3	6	4		4
14	14	14	11	4	12	14	2	12		14
4	4	6	4	2	6	4	9	3		4
2	2	4	4	2	2	3	8	3		3
1	1	1	1	1	1	1	1	2		2
11	14	11	11	12	3	7	9	12		3
1	1	2	1	1	1	1	2	1		3
3	3	7	8	3	5	3	5	3		5
1	2	2	1	3	1	1	2	2		2
7	11	8	9	2	5	7	8	9		6
7	7	5	8	3	5	7	7	5		7
1	1	1	1	1	1	1	1	1		2
2	1	1	1	1	2	3	1	2		2
1	1	1	2	2	2	1	2	1		3
1	1	1	1	2	1	1	2	2		1
1	1	1	1	1	1	1	1	2		2
1	1	1	1	1	1	1	1	1		1

1	1	1	1	1	1	1	1	1	2
1	1	1	1	2	1	1	2	2	2
7	5	8	9	3	5	5	7	6	9
1	1	1	1	1	1	1	1	2	2
2	2	1	2	2	2	1	1	2	2
1	1	1	1	1	1	1	1	2	2
3	2	4	3	3	5	2	2	3	5
2	2	2	4	2	2	2	2	2	4
2	2	1	1	1	2	1	2	2	3
2	2	4	4	2	2	2	2	2	2
2	2	1	1	1	2	2	1	2	3
1	1	2	1	1	1	1	1	2	1
9	13	10	5	6	5	5	8	5	6
5	4	12	2	2	3	3	2	2	4
1	2	2	1	1	1	1	2	2	3
2	2	2	2	3	2	2	2	5	3
3	8	10	9	5	5	5	10	9	9
2	2	2	2	2	2	2	2	3	3
4	4	8	10	4	6	5	6	10	6
1	1	2	1	1	2	1	1	2	2
9	11	11	13	5	5	13	13	13	7
1	1	1	1	1	2	1	2	3	2
3	3	2	7	3	5	9	5	7	7
13	14	14	12	14	11	11	14	12	14
2	7	2	2	2	2	7	8	5	3
9	11	13	11	6	5	7	10	11	10
7	11	11	11	6	5	9	11	11	9
1	1	1	1	1	1	1	1	1	2
6	5	9	7	6	6	5	10	10	6
2	2	2	3	2	5	3	5	3	2
2	2	2	2	2	3	2	2	3	5

1	1	1	1	1	1	1	1	2	2
3	3	4	2	2	2	2	1	4	5
13	14	13	9	4	5	5	12	11	10
1	1	1	1	2	1	1	1	2	5
1	2	1	1	1	1	1	1	2	2
2	1	2	1	1	2	1	2	1	2
3	1	2	1	2	2	1	1	2	3
1	1	2	1	2	1	1	1	2	14
13	13	14	14	14	14	13	14	14	5
3	3	8	3	4	5	3	4	5	10
3	2	13	1	2	2	2	2	2	2
3	2	2	3	1	5	2	2	3	13
9	14	13	13	13	9	13	10	11	13
7	11	11	11	13	9	11	12	13	12
3	12	10	8	6	3	3	6	7	10
13	13	14	13	14	13	13	13	14	14
7	13	11	11	5	7	9	9	13	12
1	1	1	1	1	1	1	1	2	2

2.7.2 Instructions to authors

British Journal of Dermatology author guidelines (2020)

General guidance

All Original Articles should include:

- Concise and informative title not exceeding 70 characters
- A structured abstract with background, objectives, methods, results and conclusions (maximum 250 words).
- Vancouver consecutive unbracketed superscript referencing style
- Up to 3000 words of body text.
- Around 6 total figures and tables.
- No limit for the number of references