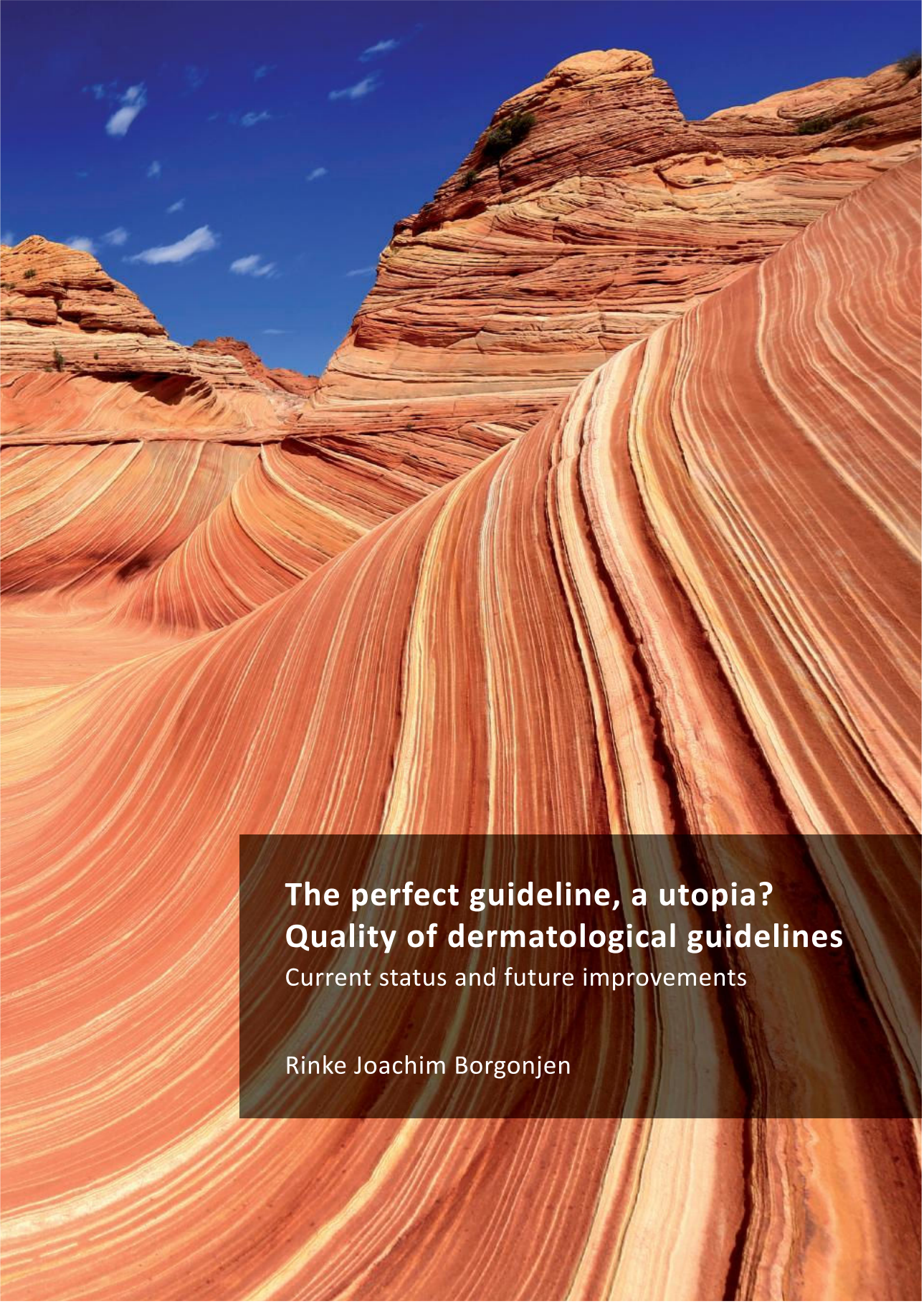


The perfect guideline, a utopia?
Quality of dermatological guidelines

Current status and future improvements

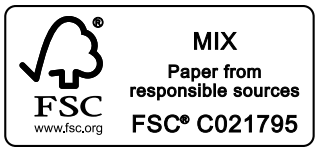
Rinke Joachim Borgonjen

The perfect guideline, a utopia?

Quality of dermatological guidelines

Current status and future improvements

Rinke Joachim Borgonjen



Cover design: Rinke Borgonjen and Willy Cremers
Layout: Rinke Borgonjen and Willy Cremers
Printing: GVO drukkers & vormgevers
© Rinke Borgonjen, 2018. All rights reserved. No part of this thesis may be reproduced in any form or by any means without permission of the holder of the copyright.
ISBN: 978-94-6332-324-6
This thesis was not financially supported by any company.

The perfect guideline, a utopia?

Quality of dermatological guidelines

Current status and future improvements

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op woensdag 9 mei 2018, te 12.00 uur

door
Rinke Joachim Borgonjen
geboren te Brummen

PROMOTIECOMMISSIE

Promotores	Prof. dr. P.C.M. van de Kerkhof	Radboud Universiteit Nijmegen
	Prof. dr. P.I. Spuls	Universiteit van Amsterdam
Copromotor	Dr. J.J.E. van Everdingen	Nederlandse Vereniging voor Dermatologie en Venereologie
Overige leden	Prof. dr. E.M.G.J. de Jong	Radboud Universiteit Nijmegen
	Prof. dr. G.D.E.M. van der Weijden	Maastricht University
	Prof. dr. M.J.M.H. Lombarts	Universiteit van Amsterdam
	Prof. dr. N.S. Klazinga	Universiteit van Amsterdam
	Prof. dr. P.M.M. Bossuyt	Universiteit van Amsterdam
	Prof. dr. T.E.C. Nijsten	Erasmus Universiteit Rotterdam
Faculteit	Geneeskunde	

“You wouldn't abandon ship in a storm
just because you couldn't control the winds.”
—Sir Thomas More, Utopia

TABLE OF CONTENTS

CHAPTER 1
GENERAL INTRODUCTION AND THESIS OUTLINE 9

CHAPTER 2
RATIONALLY CHOOSING GUIDELINE TOPICS
Prioritizing dermatoses; rationally choosing guideline topics 25

CHAPTER 3
QUALITY OF DERMATOLOGICAL GUIDELINES
Quality of Dutch dermatological guidelines according to AGREE II 37

CHAPTER 4
TO A GUIDELINE FOR OFF-LABEL CYCLOSPORIN USE; A GRADE APPROACH
Off-label use of cyclosporin in dermatology: a GRADE review 49

CHAPTER 5
COMPARISON OF THREE DIFFERENT METHODS OF GUIDELINE DEVELOPMENT
Prospective comparison of three different methods of guideline development for treatment of actinic keratosis 129

CHAPTER 6
GUIDELINE-BASED CLINICAL DECISION SUPPORT
Guideline-based clinical decision support in acne patients receiving isotretinoin; improving adherence and cost-effectiveness 157

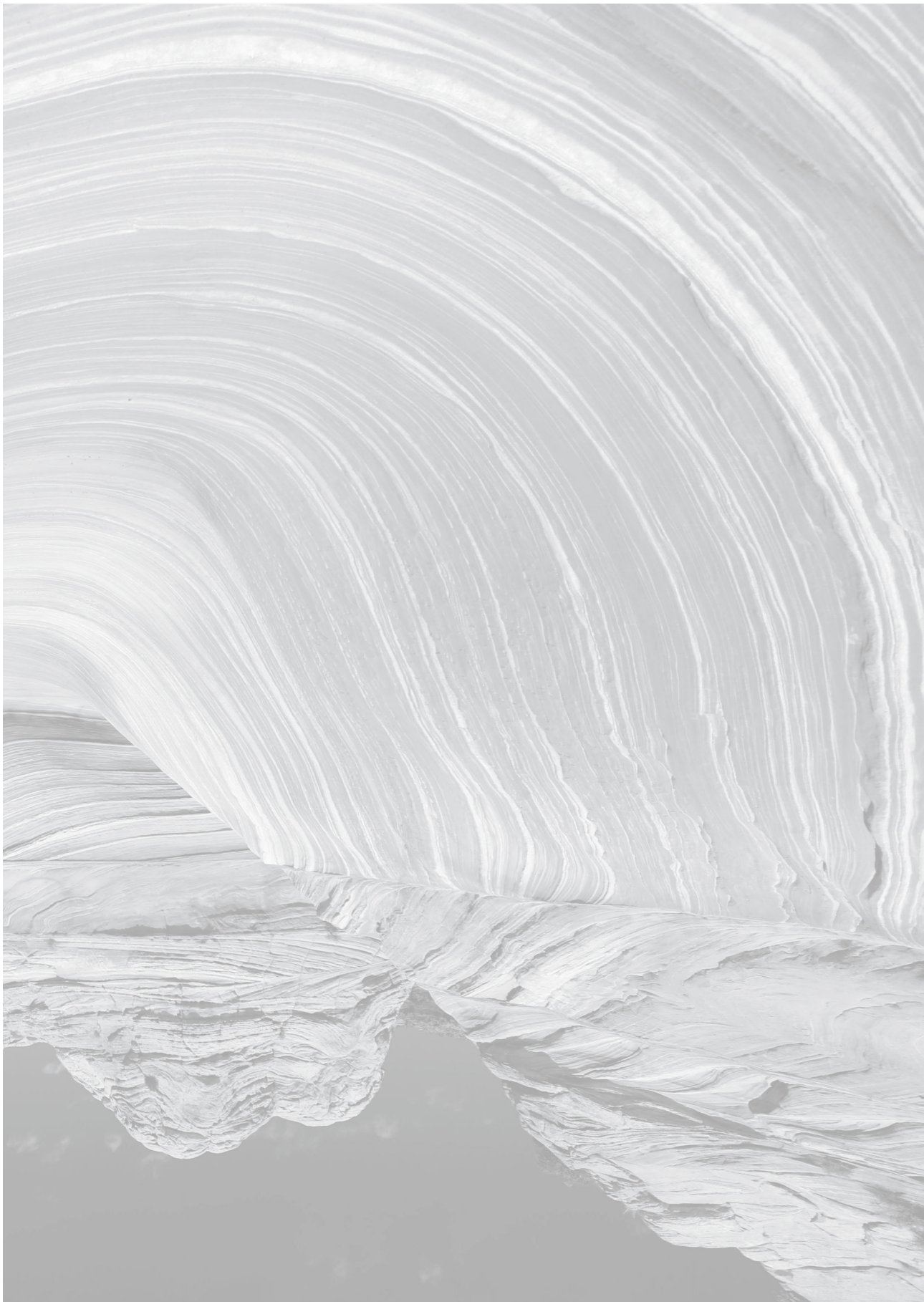
CHAPTER 7
IMPROVING THE APPLICABILITY OF GUIDELINES
Improving the applicability of guidelines on nonmelanoma skin cancer in frail older adults: a multidisciplinary expert consensus and systematic review of current guidelines 167

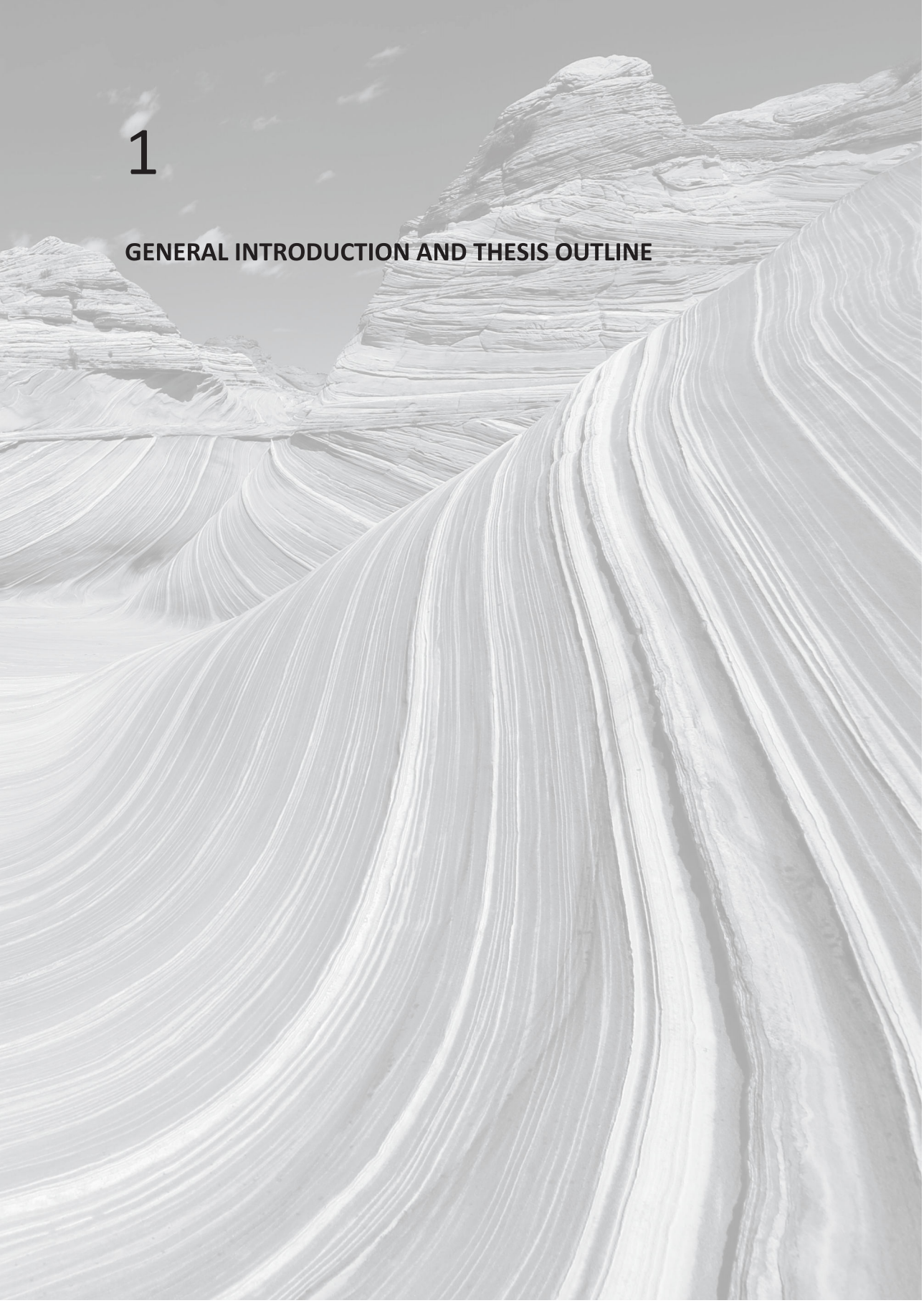
CHAPTER 8
DEVELOPMENT OF A TOOL TO ASSESS GUIDELINE ADHERENCE
A national study on adherence to a basal cell carcinoma guideline; development of a tool to asses guideline adherence 189

CHAPTER 9
DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS 203

CHAPTER 10
SUMMARY 233
NEDERLANDSE SAMENVATTING (DUTCH SUMMARY) 237

CHAPTER 11
LIST OF PUBLICATIONS 243
CURRICULUM VITAE 245
ACKNOWLEDGEMENTS 247





1

GENERAL INTRODUCTION AND THESIS OUTLINE

INTRODUCTION

1

Humanity has probably always used frameworks or models in trying to understand and control a complex environment. Nevertheless, nowadays the oldest known guidance is the Code of Ur-Nammu, a set of laws on a clay tablet from the Middle East around 2100 BC.[1] Clay tablets from later dates showed that all around the Middle East people documented medical laws and cases. Ancient Indian and Chinese societies made more systematic texts about surgical procedures, treatments and hygienic measures, that still form the basis for traditional medicine.[2,3] Yet, truly organized methodical development of clinical practice guidelines started in the seventies and early eighties of the last century (Figure 1) because of concerns about rising healthcare costs and an increasing demand for transparency from healthcare users.[4] The Netherlands were one of the first in Europe with a national policy about medical guideline development. The first national Dutch clinical practice guideline, about blood transfusion policy, was published in 1982.[5] Initially Dutch national guidelines were solely developed by the so-called 'Centraal Begeleidingsorgaan voor de Intercollegiale Toetsing (CBO)'; a national organisation supervising guideline development.[4,5] Dutch dermatologists took part in CBO guideline development from the beginning with a guideline about melanoma in 1984 and diabetic ulcers in 1985.[6] Internationally, dermatological societies such as the British Association of Dermatologists and the American Academy of Dermatology started developing guidelines in the 1990's and Pan-European guidelines followed after the foundation of the European Dermatology Forum in 1997.[7]

At first clinical practice guidelines were mainly consensus based; the product of an expert meeting in which experts proposed literature that they considered important. This is sometimes a bit derogatorily called the GOBSAT method (Good Old Boys Sitting Around The Table). Controversial topics concerning multiple medical subspecialties suited a consensus based guideline. The goal of those guidelines was to describe adequate healthcare and to divide responsibilities. However clinicians, the intended audience for a guideline, progressively demanded to know why a recommendation was expressed the way it was.[8] This call for explanation gave rise to a more formalized explicit way to develop recommendations including a scientific rationale. Thus, reproducible, more evidence based, guidelines began to emerge. Systematic searches in scientific literature and rating of scientific papers turned a guideline into an up to date scientific summary with subsequent (multidisciplinary) judgment. Consequently, a guideline tries to reflect appropriate medical care and give explicit evidence based recommendations to deliver optimal healthcare.

More and more medical professional groups began to develop guidelines and guidelines became important documents in the quality of care policy of those professional groups.[9] Additionally clinicians and policy makers identified guidelines as a tool to condense the increasing amount of medical literature. For a dermatologist until 1987 it was possible to keep abreast of all the dermatological articles enlisted in MEDLINE® by reading approximately one article every other day. However, a dermatologist in 2015 would need to read 36 articles a day to stay up to date (Figure 1) and extensive textbooks or systemic reviews often became outdated even before they were published.[10] Systematic reviews also tried to summarize the growing available scientific literature, but failed to deliver a translation into daily practice.[11]

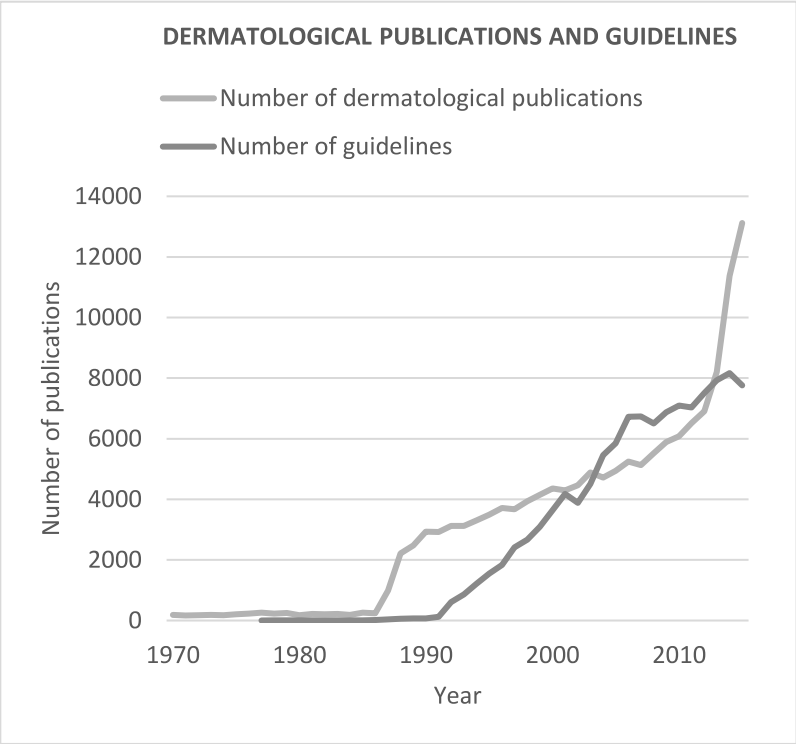


Figure 1: Number of dermatological publications and guidelines enlisted in MEDLINE per year. Source: PubMed. Accessed 20.03.2017.

Guidelines attempt to deliver meaningful guidance for everyday life. This was reflected in the attention to involve patients in the guideline development process, to consider costs when formulating a recommendation with financial consequences and to give attention to implementation after a guideline was made.[12] But, a few problems arose with the continuing production of more guidelines and the broadening scope of guidelines. In the last decades guideline development has focused on further refining the systematic and evidence-based process by harmonising the grading of evidence and formulation of recommendations (www.gradeworkinggroup.org) as well as the evaluation and ease of use of guidelines.[13] Although the potential application of an evidence based guideline in daily care is huge, a number of difficulties and downsides exist concerning the development and implementation.[14]

Definition and objective

First of all, are guidelines doing what they intend to do? From early on guidelines were planned to raise the quality of care and reduce unwanted variation in care.[15] There are many definitions of a guideline, but perhaps the most fitting Dutch definition of a guideline is the Brummen-definition: a guideline is a document with recommendations to support healthcare professionals and healthcare users, aiming to improve quality of care, build on scientific research complemented with expertise and experience of both healthcare professionals and healthcare users.[16] Several international reviews showed that the majority of guideline studies contributed to significant improvements with respect to the process of care. Fewer studies have focused on the effects of guidelines on patient health outcomes.[17] In the field of dermatology Nast et al. studied if prescription behaviour altered after the German psoriasis guideline was introduced and they found an increase from 31 to 35% in the average proportion of prescribed systemic treatments for patients with moderate to severe psoriasis. 80% of the study participants stated that the guideline had led to changes in their treatment decisions, whereas only 20% stated that the guideline had no impact.[18] Wakkee et al. studied the awareness and acceptance of the Dutch evidence-based guidelines for the treatment of moderate to severe plaque psoriasis among Dutch dermatologists. Almost all respondents were aware of the guidelines and 60% reported to have a decent knowledge of their content. Enhancing the familiarity with the guidelines led to an improved attitude towards them and consequently to increased use of the guidelines.[19] Still, guideline utilization is complex and many factors may influence the impact of guidelines on care.[20] Effective implementation should ensure guideline adherence in practice and subsequently lead to improved patient outcomes.[21]

Implementation and adherence

The influence of a guideline on clinical practice depends to a great extent upon following the actual recommendations. In the last decades implementation and adherence became hot topics. It namely turned out that flaws in an initial stage had a negative impact on adherence afterwards and often were irreversible at a later stage.[4] As said before guidelines should be a means to an end. Implementation and evaluation are therefore essential parts of the guideline development process. Although the importance of guideline implementation is progressively recognized, various healthcare institutions still pay more attention to guideline development than to implementation. Many guidelines do not include features that promote its use.[22] Several factors can form a barrier to successful implementation.

- User related barriers

There are barriers intrinsic to the guideline user, such as lack of awareness or a compulsion to treat.[19] The expanding literature (Figure 1) makes it difficult for any healthcare professional to be aware of every applicable guideline and critically apply it to practice. For 78% of the guidelines, more than 10% of physicians are not aware of its existence.[22] Furthermore healthcare users may be resistant or perceive no need for guideline recommendations.

- Guideline related barriers

Also, barriers relating to the guideline itself can be critical. An example is the wording used in recommendations. Recommendations in guidelines are not always clear and sometimes vague wording is used (e.g. start methotrexate when the patient has *serious* itch for a *long* period). The wording of a behavioural instruction affects the likelihood that it will be followed.[23] In a study of 10 Dutch national clinical guidelines, general practitioners followed the guidelines on 67% of occasions if they were concrete and precise but on only 36% of occasions when they were vague and non-specific.[24]

Another problem was that with attention for healthcare users, an extensive structured method and a desire to be transparent, guidelines became all-embracing documents. It is still not uncommon that guidelines are written down as long texts of several hundred pages, which can be cumbersome to read and difficult to integrate in the care process. A vital revolution in this case is the relative ease today with which a recommendation might be incorporated into electronic health systems to deliver computerized clinical decision support. Electronic implementations are successful in

influencing clinician behavior, however wide variation in electronic healthcare systems often hamper applicability.[25]

There are different implementation strategies to overcome aforementioned barriers. An overview of 41 systematic reviews found that the most promising approach was to use a variety of interventions including audit and feedback, reminders, and educational outreach.[26] Tools were designed to help developers and implementers better understand and anticipate barriers such as the GuideLine Implementability Appraisal (GLIA). After usage of the tool guideline developers might choose to make modifications to the guideline before dissemination. However the GLIA tool mainly deals with factors related to the guideline itself.[27] Future extensions to such tools will possibly help with extrinsic barriers to implementation.

Other stakeholders and legal issues

The primary end users of a guideline are healthcare professionals and healthcare users. As increasingly valuable tools to improve healthcare and support clinicians, other stakeholders started to take interest in and use a clinical practice guideline as well. Examples of these in the Netherlands are health insurers, Boards of Medical Examiners, the Dutch Health Care Inspectorate and other governmental organisations. As such, the reach of a guideline is widening even more. Accordingly guidelines can play an important part in the budgeting of healthcare, in legal matters and in making agreements between factions. The application of guidelines in practice is the responsibility of the healthcare professional. Since guidelines are not mandatory, professionals can deviate from or even ignore a recommendation. Facts or circumstances may occur in which case it may be advisable to deviate from the guidelines in the interest of the patient. Yet, documentation of the reasons and considerations is essential to avoid misinterpretations.

Quality and validity

An essential matter, bearing in mind the above-mentioned issues, is how to guarantee the quality of a guideline. A good quality guideline deals with potential sources of bias to warrant internal and external validity and achievable recommendations in daily practice. In the Netherlands the responsibility for developing good quality guidelines rests with the professional societies. In this case the Dutch Society of Dermatology and Venereology. Important was the international development and validation of the AGREE (Appraisal of Guidelines for Research and Evaluation) instrument.[28] The AGREE instrument is a tool for guideline developers and users to ensure the methodological

quality and reporting of a clinical practice guideline. The original AGREE instrument, as well as the updated second version of the instrument (AGREE II), uses criteria, grouped into 6 domains (scope and purpose, stakeholder involvement, methodology, clarity and presentation, applicability, and editorial independence). A limitation of the instrument is a lack of criteria assessing the clinical content of recommendations. Also, a description of the search strategy is enough to achieve a high domain score and the instrument thus does not test the validity of the strategy. Therefore the AGREE instrument is a great instrument to achieve methodological transparency and reproducibility, but not of the underlying quality of supporting evidence and guideline content itself.

Regarding the underlying scientific quality of recommendations a lot of progress has been made with the introduction of the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) system for grading the quality of evidence and the strength of recommendations.[29] GRADE tries to enhance transparancy by rating the available evidence in a uniform way and ultimately rate the strength of a recommendation as well. The strength of a recommendation reflects the extent to which one can, across the range of healthcare users for whom the recommendations are intended, be confident that desirable effects of a management strategy outweigh undesirable effects. Traditionally GRADE was developed for therapeutic questions where a blinded randomized controlled trial is the highest form of evidence. Increasingly guidance becomes available as well for other questions or cases of lacking evidence.[30,31]

Cost-effectiveness and revision

As mentioned before concerns about rising healthcare costs were one of the driving forces behind guideline development. Cost-effectiveness should not only apply to healthcare questions covered in the guideline itself though. The whole process of guideline development should be cost-effective and manageable too.[32] A disadvantage of the increasing complexity of guidelines is that a guideline development team should include members with skills in guideline implementation and critical literature appraisal as well as members with specific understanding of the clinical domain. And with the continuing production of guidelines (Figure 1) the question soon arose who is going to keep those guidelines up to date? When developing a guideline the scientific literature is sometimes outdated as soon as the guideline is implemented. Creating and updating guidelines that keep abreast with expanding knowledge requires a huge amount of time and resources.[33] In Germany,

for example, solely the national psoriasis guideline has cost approximately 250,000 euro to create.[34] International collaboration and adaptation of existing guidelines could save time and valuable resources in creating guidelines.[35,36] However modifications might be imperative to reflect the way in which medical care is delivered within a particular setting.

Yet, a guideline can only remain valid in case it is maintained on a regular base, with systematic monitoring of medical scientific literature as well as data from practice and comments given by users of the guideline.[37,38] Clearly a guideline needs updating if the majority of recommendations are out-of-date, with new evidence demonstrating that the recommended interventions are inappropriate, ineffective, or superseded by new interventions.[39] Modular updates might be another way to save time and resources.

THE EVIDENCE BASED GUIDELINE CYCLE

Eventually, all the different parts of evidence based guideline development form a cycle. The cycle depicts different parts of the process (Figure 2). The first part consists of the analysis of bottlenecks experienced in daily practice. A guideline is not the solution for every bottleneck. Subsequently, potential subjects are selected and prioritized. Depending on the topic, a balanced guideline development group that includes representatives from all relevant disciplines is composed. The composition of a group and group dynamics can influence the outcome of discussions and lead to variation among the eventual recommendations.[4,40,41] Addressing group dynamics requires careful discussion within a guideline development group.

Questions relating to the problem are specified according to the PICO format (Patient, Intervention, Comparison, Outcome) and appropriate outcome measures selected. The latter is part of the GRADE method to rank potential outcome measures and only consider important or critical outcomes. Then, a systematic literature search with the appropriate criteria to answer the raised questions is carried out. After that relevant literature is selected according to the predefined selection criteria. Next the literature is summarized and rated according to the prevailing method. What follows is the translation of evidence summaries into clinical recommendations. A guideline after all has the most potential impact when evidence is limited, controversial or contradictory. Judgment beyond the evidence is necessary, taking into consideration the balance between benefits and harms, healthcare users views and preferences, the availability of a technique, a device or certain medication and potential organizational, social and financial barriers.[14]

The final formulated clinical recommendations are the result of the available evidence in combination with these practical issues. The aim of following this procedure is to

enhance the unambiguousness for the user of the guideline. Also, systematic and explicit approaches can help to protect against errors and resolve disagreements. The end products of a guideline are then implemented and disseminated among end users. Over time feedback from users and perhaps innovations bring up new bottlenecks and the cycle is complete.

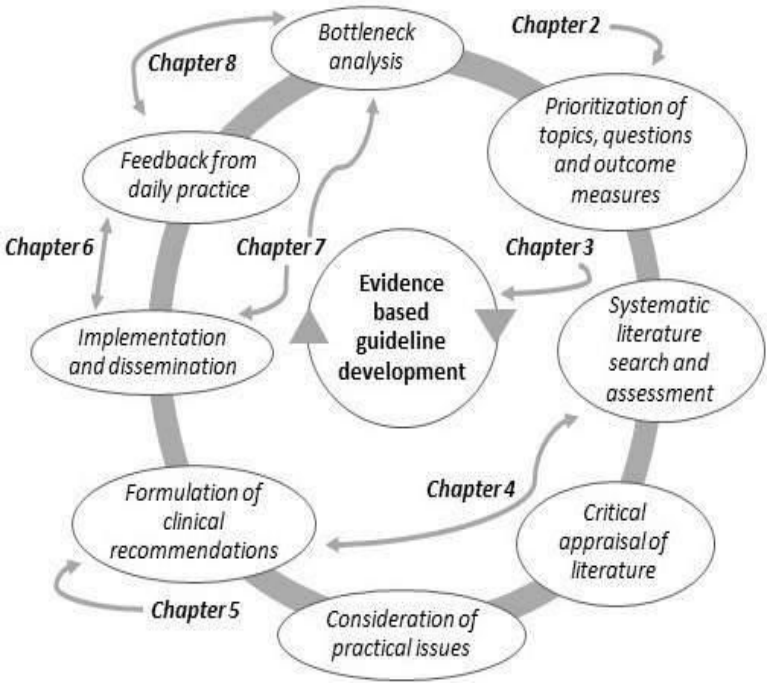


Figure 2: Evidence based guideline development cycle. Arrows represent the parts of the cycle covered in that particular chapter.

THESIS OUTLINE

The aim of this thesis is to investigate the quality of dermatological guidelines. As said before, guideline development can be and should be depicted as a cycle that shows the various aspects of its process.[42] The studies presented in this thesis each highlight parts of this cycle to explore possible ways to further improve dermatological guidelines (Figure 2). Major steps in setting a high-quality standard have already been made with the introduction of the AGREE instrument, the formation of national and international platforms (such as G-I-N, the Guideline International Network and GENEVER, the Dutch alliance for excellent guidelines (GEzamenlijk NEderlands Verbond voor Excellente Richtlijnen)) and in promoting transparent guideline development and implementation methods (GRADE).[13, 28, 29,43] However much remains to be elucidated.

Chapter 2 addresses the question: which subject has the most priority? Resources to develop a guideline are limited and not every subject needs a clinical practice guideline. Hence guideline development starts with selecting and prioritizing. However, a standardized method for topic selection is frequently missing and the way different criteria are used to prioritize is indistinct. In chapter 2 dermatological topics for guideline development are selected and prioritized. An additional objective was to elucidate criteria dermatologists find important in selecting guideline topics. By involving future guideline users in the process of selecting and prioritizing topics, support for an upcoming guideline can be created. This in turn might promote successful implementation and adherence of a guideline.

Chapter 3 investigates the current methodological quality of Dutch dermatological guidelines with AGREE II. What is essential for a guideline to be a high quality guideline? It is known that guideline quality is highly variable across countries and between specialty societies.[44,45] AGREE II domain scores highlight existing good quality guideline areas as well as domains with room for future methodological improvements.

How to translate the evidence from scientific literature to a clear recommendation fit for practice? **Chapter 4** describes the whole process of making a guideline for off-label cyclosporine use in dermatology according to GRADE. GRADE is at present the golden standard in explicating considerations in formulating recommendations and appraisal of literature.[29]

Chapter 5 focusses on the most efficient way to develop a high quality guideline. The early consensus based guidelines had the big advantage of being relatively cheap and quick to make. The systematic identification and critical analysis of evidence are the most costly and time-consuming components in the guideline development process.[46] In chapter 5 three different methods of guideline development are tested to see whether using alternative evidence-based methods result in variation of recommendations for treating actinic keratosis. Equally substantial outcomes were total costs and time period necessary to make a draft guideline.

How to improve the implementation of a guideline? In **Chapter 6** it is recognized that guidelines are often published as long texts and implementation in daily practice is challenging. However clinical practice guidelines are intended to improve quality of care. Chapter 6 tests the impact of active and passive guideline based clinical decision support on laboratory monitoring of acne patients receiving isotretinoin. Its aim is to increase guideline adherence and achieve a reduction in clinical practice variation and costs.

Chapter 7 seeks to what extent existing bottlenecks experienced in everyday practice are incorporated in clinical practice guidelines. Moreover an integrated and structured way of retrieving bottlenecks is tried. The underlying problem is that balancing treatment decisions in frail older adults with nonmelanoma skin cancer can be challenging. Clinical practice guidelines could provide assistance. The goal was to collect and prioritize items related to frail older adults with nonmelanoma skin cancer for integration into clinical practice guidelines and to assess the current extent of this integration.

How to receive feedback and convert it into a revision of a guideline? **Chapter 8** describes the development of a tool for obtaining and giving nationwide feedback regarding adherence. Feedback regarding adherence is essential to monitor the effect that a guideline has in clinical practice and whether or not the quality of care is raised by implementation. Clinical practice guideline implementation may be at variance with actual daily practice, as guideline adherence is a complex process depending on many actors and factors.

This thesis concludes with **Chapter 9**, which contains a general discussion, conclusions, recommendations and suggestions for further research.

REFERENCES

1 Kramer, S. Ur-Nammu Law Code. Orientalia 1954;23(1):40-51.
2 Bhishagratna, Kaviraj KL. An English Translation of the Sushruta Samhita in Three Volumes. Calcutta, 1907.
3 Harper, D. Early Chinese Medical Literature: The Mawangdui Medical Manuscripts. Kegan Paul International. London and New York, 1998.
4 Van Everdingen JJE, Dreesens DHH, Burgers JS, Swinkels JA, van Barneveld TA, van der Weijden (red). Evidence-Based Richtlijnontwikkeling. Een leidraad voor de praktijk. Bohn Stafleu van Loghum. Houten, 2014.
5 Van Everdingen JJE. Van consensus naar CBO-richtlijn. Ned Tijdschr Geneeskd 1999;143:2086-8.
6 Hukkelhoven CWPM, Burgers JS, van Everdingen JJE et al. De kwaliteit van dermatologische CBO-richtlijnen beoordeeld met het AGREE Instrument. Ned Tijdschr Dermatol 2007;17:22-7.
7 Peter Fritsch, Walter Burgdorf (eds.) EDF White Book: Skin diseases in Europe. ABW Wissenschaftsverlag. Berlin, 2005 (2nd edition).
8 Grol RPTM, Everdingen JJE van, Kuipers F, Casparie AF. Consensus over consensus. Een kritische beschouwing van de procedure van de CBO-consensusontwikkeling. Ned Tijdschr Geneeskd 1990;134:1186-9.
9 Buiting AMJ, van Barneveld T, Bartels PCM. Richtlijnen als basis voor kwaliteitsontwikkeling. Ned Tijdschr Klin Chem 2003;28:192-196.
10 Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med 2007;147:224-233.
11 Offringa M, Craen AJM de. De praktijk van systematische reviews. I. Inleiding. Ned Tijdschr Geneeskd 1999;143:653-6.
12 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003;362(9391):1225-30.
13 Burgers JS, Van Everdingen JJE. Evidence-based richtlijnontwikkeling: het EBRO-platform. Ned Tijdschr Geneeskd 2004;148:2057-9.
14 Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 1999;318(7182):527-30.
15 Field MJ, Lohr KN. Clinical Practice Guidelines: Directions for a New Program. National Academy Press. Washington D.C., 1990.
16 Werkgroep Richtlijn voor Richtlijnen; in opdracht: Regieraad Kwaliteit van Zorg. Richtlijn voor Richtlijnen. Den Haag, 2010
17 Lugtenberg M, Burgers JS, Westert GP. Effects of evidence-based clinical practice guidelines on quality of care: a systematic review. Qual Saf Health Care. 2009;18(5):385-92.
18 Nast A, Erdmann R, Hofelich V, Reytan N, Orawa H, Sterry W et al. Do guidelines change the way we treat? Studying prescription behaviour among private practitioners before and after the publication of the German Psoriasis Guidelines. Arch Dermatol Res 2009;301:553–559.
19 Wakkee M, Lugtenberg M, Spuls PI, de Jong EM, Thio HB, Westert GP et al. Knowledge, attitudes and use of the guidelines for the treatment of moderate to severe plaque psoriasis among Dutch dermatologists. Br J Dermatol 2008;159(2):426–432.
20 Sheldon TA, Cullum N, Dawson D et al. What's the evidence that nice guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients, notes, and interviews. BMJ 2004;329:999-1006.

21 Grol, R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Medical Care 2001;39(2):46-54.

22 Cabana MD. Why Don't Physicians Follow Clinical Practice Guidelines?: A Framework for Improvement. JAMA 1999;282(15):1458-65.

23 Shekelle PG, Kravitz RL, Beart J, Marger M, Wang M, Lee M. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. Health Serv Res. 2000;34(7):1429-48.

24 Grol R, Dalhuijsen J, Thomas S, Veld C, Rutten G, Mokkink H. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ 1998;317:858-61.

25 Gagliardi et al. How can we improve guideline use? A conceptual framework of implementability. Implementation Science 2011;6:26.

26 Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, et al. Changing provider behaviour: an overview of systematic reviews of interventions. Med Care 2001;39(2):2-45.

27 Shiffman RN, Dixon J, Brandt C, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak 2005;5:23.

28 Cluzeau FA, Burgers JS, Brouwers M, Grol R, Mäkelä M, Littlejohns P et al.; The AGREE Collaboration Writing Group. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 2003;12:18-23.

29 Guyatt GH, Oxman AD, Schunemann HJ, et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clinical Epidemiol. 2011;64(4):380-2.

30 Alexander PE, Gionfriddo MR, Li SA, Bero L, Stoltzfus RJ, Neumann I et al. A number of factors explain why WHO guideline developers make strong recommendations inconsistent with GRADE guidance. J Clin Epidemiol. 2016;70:111-22.

31 Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al.; GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ. 2016;353:i2089.

32 Shekelle P, Woolf S, Grimshaw JM, Schunemann HJ, Eccles MP. Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development. Implementation science 2012;7:62.

33 Shekelle P, Eccles MP, Grimshaw JM, Woolf SH: When should clinical guidelines be updated? BMJ 2001;323:155-157.

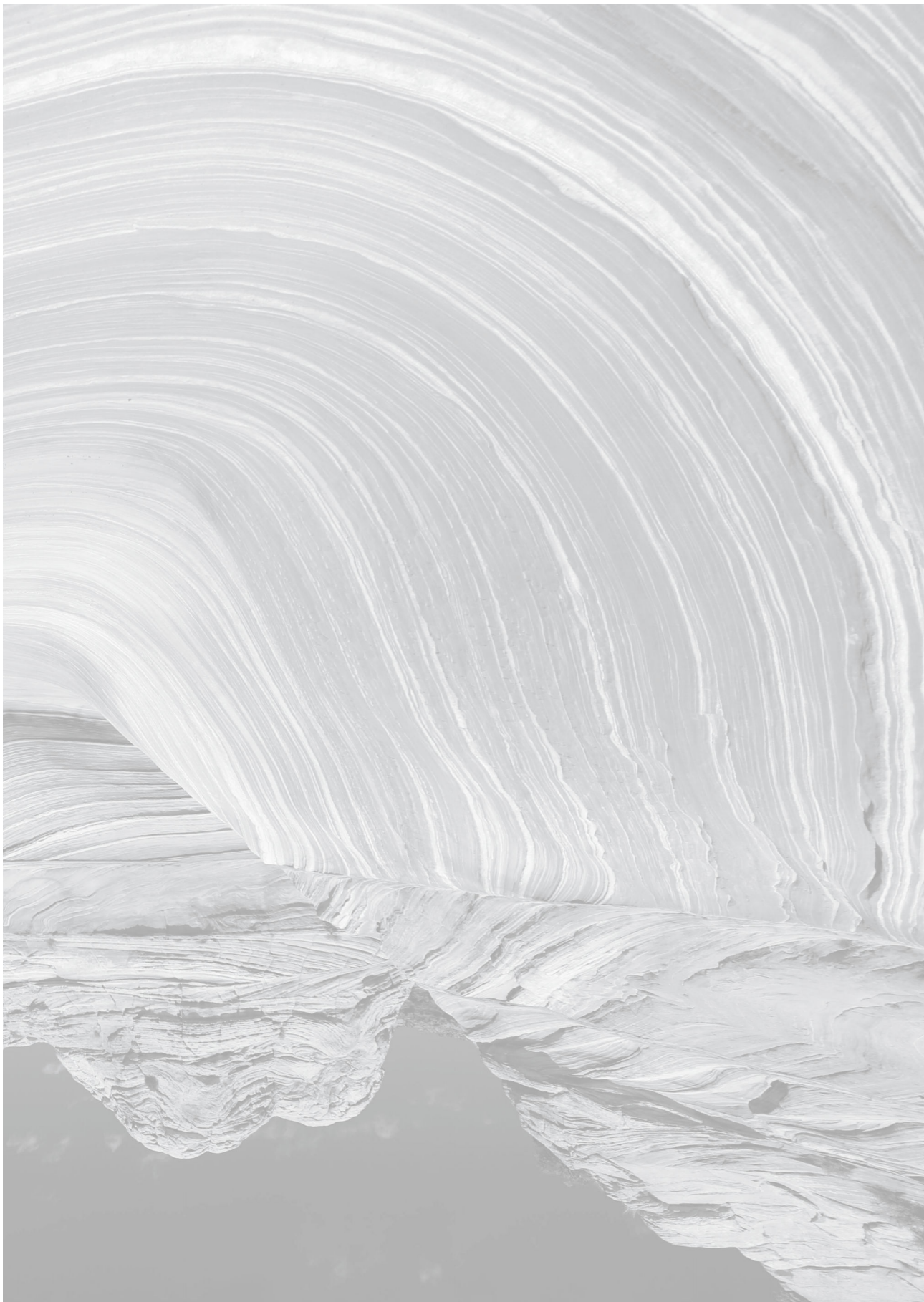
34 Nast A, Reytan N, Rosumeck S et al. Zeitaufwand und Kosten bei der Erstellung der S3-Leitlinie zur Therapie der Psoriasis vulgaris. J Dtsch Dermatol Ges 2007;5(s2):198

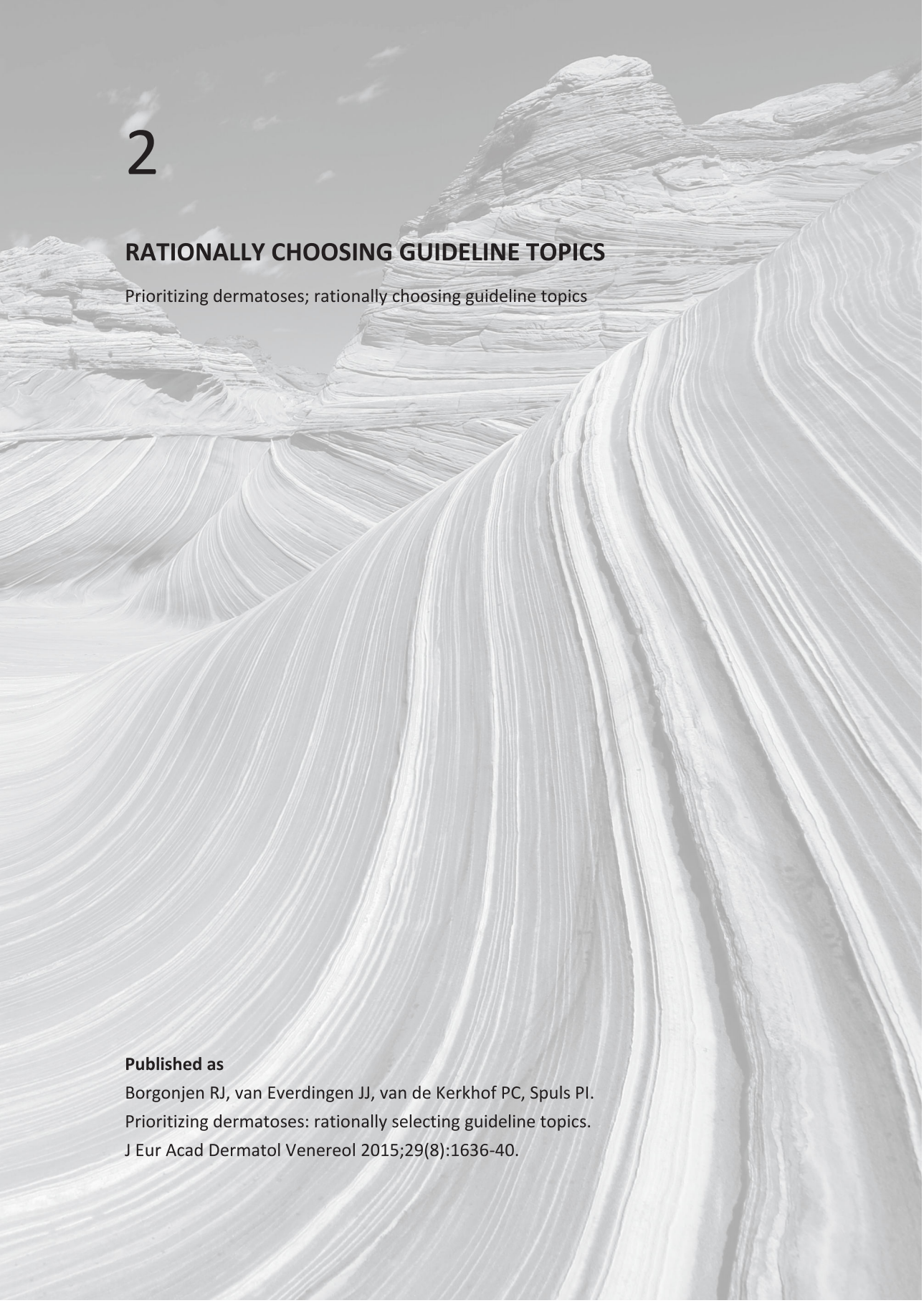
35 Fervers B, Burgers JS, Voellinger R, Brouwers M, Browman GP, Graham ID, et al.; ADAPTE Collaboration. Guideline adaptation: an approach to enhance efficiency in guideline development and improve utilisation. BMJ Qual Saf. 2011;20(3):228-36.

36 Spuls PI, Nast A. Evaluation of and perspectives on guidelines: what is important? J Invest Dermatol. 2010;130(10):2348-9.

37 Foy R, MacLennan G, Grimshaw J, Penney G, Campbell M, Grol R. Attributes of clinical recommendations that influence change in practice following audit and feedback. J Clin Epidemiol 2002;55(7):717-22.

38 Gartlehner G, West SL, Lohr KN, Kahwati L, Johnson JG, Harris RP et al. Assessing the need to update prevention guidelines: a comparison of two methods. Int J Qual Health Care 2004;16:399-406.
39 Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM et al. Clinical practice guidelines: how quickly do guidelines become outdated? JAMA 2001;286:1461-1467.
40 Pagliari C, Grimshaw J. Impact of group structure and multidisciplinary evidence-based guideline development: observed study. J Eval Clin Pract 2002;8:145–53.
41 Raine R, Sanderson C, Hutchings A, Carter S, Larkin K, Black N. An experimental study of determinants of group judgments in clinical guideline development. Lancet 2004;364:429–37.
42 Adviescommissie Richtlijnen. In opdracht: Raad voor Wetenschap, Opleiding en Kwaliteit van de Orde van Medisch Specialisten. Medisch specialistische richtlijnen 2.0. Utrecht, 2011.
43 Ollenschläger G, Marshall C, Qureshi S, Rosenbrand K, Burgers J, Mäkelä M et al.; Guidelines International Network (G-I-N). Improving the quality of health care: using international collaboration to inform guideline programmes by founding the Guidelines International Network (G-I-N). Qual Saf Health Care 2004;13(6):455-60.
44 Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. Lancet 2000;355:103-106.
45 Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA 1999;281:1900-1905.
46 Fervers B, Burgers JS, Haugh MC, Latreille J, Mlika-Cabanne N, Paquet L et al. Adaptation of clinical guidelines: literature review and proposition for a framework and procedure. Int J Qual Health Care. 2006;18(3):167-76.





2

RATIONALLY CHOOSING GUIDELINE TOPICS

Prioritizing dermatoses; rationally choosing guideline topics

Published as

Borgonjen RJ, van Everdingen JJ, van de Kerkhof PC, Spuls PI.
Prioritizing dermatoses: rationally selecting guideline topics.
J Eur Acad Dermatol Venereol 2015;29(8):1636-40.

SUMMARY

Background Clinical practice guideline (CPG) development starts with selecting appropriate topics, as resources to develop a guideline are limited. However, a standardized method for topic selection is commonly missing and the way different criteria are used to prioritize is not clear.
Objectives To select and prioritize dermatological topics for CPG development and elucidate criteria dermatologists find important in selecting guideline topics.
Methods All 410 dermatologists in the Netherlands were asked to create a top 20 of dermatological topics for which a guideline would be desirable, regardless of existing guidelines. They also rated, on a 5-point Likert scale, 10 determinative criteria derived from a combined search in literature and across (inter)national guideline developers. Top 20 topics received scores ranging from 0.01 to 0.2 and combined scores yielded a total score.
Results The 118 surveys (response 29%) identified 157 different topics. Melanoma, squamous cell carcinoma, basal cell carcinoma, psoriasis and atopic dermatitis are top priority guideline topics. Venous leg ulcer, vasculitis, varicose veins, urticaria, acne, Lyme borreliosis, cutaneous lupus erythematosus, pruritus, syphilis, lymphoedema, decubitus ulcer, hidradenitis suppurativa, androgenic alopecia and bullous pemphigoïd complete the top 20. A further 15 topics have overlapping confidence intervals. Mortality and healthcare costs are regarded as less important criteria in topic selection ($P < 0.04$), than other criteria like the potential to reduce unwanted variation in practice.
Conclusion Dermatological professional organizations worldwide succeeded in developing guidelines for all top 20 topics. Respondents mostly agree with (inter)national guideline programmes and literature concerning the criteria important to selecting guideline topics.

INTRODUCTION

2

Since the last century clinical practice guidelines (CPG) have been developed according to a method based on systematically collected scientific literature. Recommendations for daily clinical practice are formulated using this literature and everyday experience, hence making decisions and options more transparent.[1] Huge steps in improving the quality of guidelines were made with the introduction of the AGREE (Appraisal of Guidelines for Research and Evaluation) instruments and the GRADE (Grading of Recommendations Assessment, Development and Evaluation) methodology.[2,3] Nowadays, there are dozens of (inter)national guidelines, all seeking to assist the dermatologist and patient in making appropriate decisions and improve quality of care.[4] However, developing and updating guidelines costs time and money. To ensure that limited resources are used optimally, a balance must be struck between updating existing guidelines and introducing new topics.[5-7] Thus it is imperative to prioritize guideline topics.[8] Many criteria are mentioned in literature that are important to guideline topic selection, including prevalence or large unexplained variation in procedures.[7] Another method for prioritization is by consensus.[6-9] Almost every guideline development programme has different methods and criteria for setting priorities. To ensure acceptance, appreciation and implementation in daily practice, practitioners should be involved in the guideline development process, including topic selection and prioritization.[10-13] In our opinion, a more standardized practitioner-driven method to choose a guideline topic should exist.[7,8,14] This research describes the selection and prioritization of 20 dermatological guideline topics and delineates which criteria dermatologists find important in this selection process.

METHODS

In June 2010, all 410 Dutch dermatologists received a survey asking them to select and prioritize 20 dermatological topics for which a guideline would be desirable, regardless of current existing guidelines. All topics received a score varying from 0.01 to 0.2, with 0.2 for their number one and 0.01 for the 20th topic. The scores were pooled and an overall list of prioritized dermatological topics was made. Synonyms were added as a total score.

Furthermore dermatologists were asked to rate criteria on a 5-point Likert scale. The score varied from “not important at all” up to “very important” in selecting and prioritizing topics. The rated criteria are those frequently found by a combined literature search including reference checking (Appendix 1) and a search across 21 recognized (inter)national organizations that produce or collect (dermatological) CPGs

(Appendix 2). Several methodological papers on CPG prioritization and systematic reviews of selecting criteria were identified.[5-9,14-16] Criteria were categorized and double criteria deleted, eventually yielding 10 criteria (Table 1). Dermatologists were asked to mention additional criteria that they used in making their selection and priority. Conflict-of-interest statements were included in the survey and the returned surveys were handled anonymously. The availability of a dermatological guideline for each topic in the top 20 was checked using the list of 21 (inter)national organisations (Appendix 2). SPSS (Statistical Package for the Social Sciences) was used for all statistical analyses (Appendix 3).

Rank	Criterion	Mean score (range 0-5) and 95% CI
1	Relevancy for the stakeholders n=111	4.24 (4.09 – 4.40)
2	Unwanted variation in practice n=111	4.02 (3.86 – 4.18)
3	The potential to improve quality of life/ lessen burden of illness n=112	4.01 (3.82 – 4.19)
4	The applicability/implementability in practice n=111	3.98 (3.82 – 4.14)
5	The availability of scientific evidence to underpin recommendations n=111	3.95 (3.76 – 4.14)
6	The need/demand of a guideline across stakeholders n=110	3.87 (3.71 – 4.04)
7	The prevalence/incidence/morbidity of a disease n=112	3.65 (3.45 – 3.85)
8	The potential to improve the quality of public healthcare n=111	3.61 (3.44 – 3.78)
9	The potential to reduce costs on the macro-economic level n=110	3.15 (2.99 – 3.32) p = 0.04*
10	The mortality of a disease n=110	3.10 (2.83 – 3.37)

Table 1: Criteria for the selection and prioritization of guideline topics. * p-value Mann-Whitney test criterion 8 vs. 9; $\alpha = 0.05$. CI = confidence interval.

RESULTS

In total 118 dermatologists (29%), returned the survey pointing out 157 different dermatological topics for which a guideline would be desirable. Six surveys were excluded from analysis because respondents failed to provide a listing of any kind. Additionally, eighteen topics in nine surveys were excluded due to significantly different scores between negative and positive (n=5) or because of blank (n=4)

conflict-of-interest statements. The top 20 of most desired dermatological guideline topics is shown in Table 2. There is a statistically significant difference (p<0.000) in rank between the top 6 (Kruskall-Wallis 1-3 vs. 4-6 and 4-6 vs. 7-9) apart from the rank between psoriasis and atopic dermatitis (p-value 0.127) and that between actinic keratosis and venous leg ulcer (p-value 0.764). Other top 20 topic scores yielded no statistically significant differences when tested against their neighbouring ranking places and showed overlapping 95% confidence intervals with another 15 topics (Table 3).

Rank	Topic	In top 20 (n)	Sum and 95% CI
1	melanoma	97	17.8 (16.4-19.3)
2	basal cell carcinoma	94	16.6 (15.1-17.9)
3	squamous cell carcinoma	90	15.3 (13.6-17.0)
4	psoriasis	85	13.1 (11.6-14.7)
5	atopic dermatitis	72	11.3 (9.6-12.9)
6	actinic keratosis	55	8.0 (6.4-9.7)
7	venous leg ulcer	69	7.8 (6.5-9.2)
8	vasculitis	71	7.5 (6.2-8.8)
9	varicose veins	68	7.5 (6.1-8.8)
10	urticaria	62	6.2 (4.9-7.5)
11	acne	51	5.8 (4.4-7.1)
12	Lyme borreliosis	53	4.9 (3.7-6.0)
13	lupus erythematosus	50	4.5 (3.3-5.7)
14	pruritus	44	3.9 (2.9-4.8)
15	syphilis	36	3.5 (2.3-4.8)
16	lymphoedema	36	3.5 (2.3-4.6)
17	decubitus ulcer	37	3.4 (2.4-4.5)
18	hidradenitis suppurativa	37	3.4 (2.4-4.4)
19	androgenic alopecia	33	3.3 (2.2-4.5)
20	bullous pemphigoïd	36	3.2 (2.2-4.3)

Table 2: Top 20 dermatological topics. CI = confidence interval.

Criteria

Dermatologists rated 8 of the 10 criteria as important (Table 1). The potential to reduce costs on the macro-economic level and the mortality of a disease were considered moderately important and differed significantly in rank from the 8 other criteria (p-value 0.04 Mann-Whitney test criterion 8 vs. 9). Other criteria to select a guideline topic suggested by respondents were the political relevance of a topic, the role of the dermatologist in a multidisciplinary topic, and guidance in topics with many off-label or unregistered treatments.

Topic	In top 20 (n)	Sum and 95% CI
hand eczema	31	3.2 (2.1-4.2)
mastocytosis	38	3.1 (2.1-4.2)
toxic epidermal necrolysis	29	2.9 (1.8-4.0)
vitiligo	36	2.8 (1.9-3.8)
chronic venous insufficiency	23	2.6 (1.5-3.6)
chlamydia	24	2.3 (1.5-3.2)
gonorrhoea	26	2.3 (1.5-3.1)
rosacea	27	2.3 (1.4-3.2)
erysipelas	23	2.3 (1.3-3.3)
alopecia areata	23	2.1 (1.2-2.9)
pemphigus vulgaris	22	2.0 (1.1-2.9)
lichen sclerosus	28	2.0 (1.2-2.8)
lentigo maligna	18	1.9 (0.9-2.8)
hirsutism	18	1.6 (0.8-2.4)
contact dermatitis	15	1.6 (0.7-2.4)

Table 3: Additional topics with overlapping top 20 confidence intervals. CI = confidence interval.

DISCUSSION

With the appearance of guidelines on chronic pruritus, bullous pemphigoïd and androgenic alopecia, the worldwide dermatological professional bodies have succeeded in developing guidelines for the entire top 20 topics.[17-19] As the confidence intervals showed, there is an overlap between topics 12 to 20 and a further set of 15 topics (Table 3). Developing guidelines for those 15 topics should be considered in case they do not exist. However, many guideline programmes rely on experts that do most of the work voluntarily and have reached a critical point regarding the number of guidelines that they can keep up-to-date.[6] International collaboration in developing guidelines and adapting existing ones are ways to minimize (double) efforts and to maximize an efficient use of resources, particularly concerning lower priority topics. In our opinion, the statistically significant top 5 topics should be covered by (inter)national guidelines with a cyclic (modular) guideline maintenance programme, thus ensuring a 'living' guideline.

Table 3 showed that there are similarities in the criteria used to prioritize. Guidelines are often considered for common, costly topics which have large effects on morbidity and for which there is good evidence that appropriate health care can make a difference in outcomes. A wide variation in current care, probably due to professional uncertainty about how to care for the condition, is another important criterion.[7] The 10 criteria used in this study are not an absolute list and applying these criteria

together with other considerations requires an in-depth understanding.[5,13] Dermatologists or other practitioners can deliver that knowledge.
However, other stakeholders have to be involved too because the priorities set by practitioners may be conflicting with theirs.[2,6,7] For instance, health care policymakers would consider costs a top priority in decision making, whereas patients would most likely prefer guidance on issues related to quality of life.[8,9] An international survey across CPG developers revealed that decisions about the relative importance of prevalence, costs, mortality, etcetera, are made rather implicitly, given that only 30% of the respondents reported using explicit methods or criteria for setting priorities.[7] A 52% majority of the guideline development organizations in this study (Appendix 2) explicitly stated their criteria on a website, although none used a formal scoring system to quantify their criteria, as far as we know. Since the prioritization process across stakeholders is complicated by both quantitative and qualitative components, further prioritizing by the inclusion of objective quantitative data is warranted.[14] As an example, Oortwijn et al. gave points for the criteria 'burden of disease', 'benefit to the individual patient', 'number of patients', 'costs of the method per patient' and 'impact on health care costs'. Using those criteria, topics ended up in different categories of relevance.[14] Dutch governmental organizations, meanwhile, tried a quantification based on objective criteria like prevalence, mortality and costs.[20] The downside of applying a quantitative weighting is the general lack of data. Reveiz et al. reported that in only 13% of cases criteria were supported by good-quality literature. Defining suitable criteria and cut-off points is also a challenge and is strongly dependent on the actors involved in the prioritization process.[14] Subsequently, until more accurate data is available, further quantification for dermatological topics seems premature.

Limitations

The identified prioritized topics for guideline development are, in principle, only applicable to the Netherlands. Topics and priorities could vary in other countries, because of differences in health care systems, prevalence or dermatologists' attitudes. As this study deals only with topics considered important by a subsection of Dutch dermatologists (29%), there is potential response bias. Nevertheless, a survey has proved itself as a method for assessing preferences and repeating the survey in the future across Europe could bring momentum to a prioritization update.[9] The decision to specify precisely 20 topics may have been somewhat arbitrary, but was nevertheless informed by Dutch healthcare financial and professional resources. One point of consideration is the manner in which topics are grouped together, since a broader term has a higher priority score. In this study, specific topics were used

because CPGs are designed as tools to answer specific clinical questions. Still, the 20 topics are rather broad-banded and need translation into explicit clinical questions with graded outcomes when developing an actual CPG.

CONCLUSION

Skin cancer, psoriasis and atopic eczema are statistically significant top 5 topics and a further 30 topics with overlapping confidence intervals apply for the top 20. Dutch dermatologists mostly agree with (inter)national guideline programmes and literature concerning the criteria important in selecting and prioritizing a guideline topic. In an era of evidence-based guideline development, a standardized approach that uses explicit criteria and limits the influence of personal biases is needed, to ensure that the appropriate topics are selected. The shortlist of most wanted topics can serve as a basis for quantitative prioritization by other stakeholders, further pan-European research and translation into specific clinical questions.

APPENDIX 1

(((((priorit*[Title/Abstract]) OR selecti*[Title/Abstract])) AND ((Topic[Title/Abstract]) OR subject[Title/Abstract])) AND guideline[Title/Abstract]) OR (Practice guidelines as topic [Mesh] AND Program development'[Mesh] AND Program evaluation [Mesh])

APPENDIX 2

International

- Agency for Healthcare Research and Quality (AHRQ)
  http://www.ahrq.gov/
- Ärztliches Zentrum für Qualität in der Medizin (AQuMed/ÄZQ)
  http://www.aezq.de/
- Guidelines International Networks (GIN)
  http://www.g-i-n.net/
- Institute of Medicine
  http://www.nap.edu/
- National Institute for Clinical Excellence (NICE)
  http://www.nice.org.uk/
- New Zealand Guidelines Group (NZGG)
  http://www.nzgg.org.nz/
- Scottish Intercollegiate Guidelines Network (SIGN)
  http://www.sign.ac.uk/
- WHO. Handbook for guideline development.
  http://www.searo.who.int/

National

- Koninklijk Nederlands Genootschap voor Fysiotherapie (KNGF)
  https://www.cebp.nl/
- Landelijk Expertisecentrum Verpleging&Verzorging (LEVV)
  http://www.levv.nl/
- Nederlandse Huisartsen Genootschap (NHG)
  http://www.nhg.artsennet.nl/
- Orde van Medisch Specialisten.
  http://www.orde.artsennet.nl/
- Rijksinstituut voor Volksgezondheid en Milieu (RIVM)
  http://www.rivm.nl/
- Trimbos-instituut
  http://www.trimbos.nl/

- Vereniging Integrale Kankercentra (VIKC)
  http://www.oncoline.nl/
- Vereniging voor Arbeids- en Bedrijfsgeneeskunde (NVAB)
  http://nvab.artsennet.nl/
- Verpleegkundigen & Verzorgenden Nederland (V&VN)
  http://www.venvn.nl/

Dermatology

- American Academy of Dermatology
  http://www.aad.org/
- European Dermatology Forum (EDF)
  http://www.euroderm.org/
- Dutch Society of Dermatology and Venerology (NVDV)
  http://www.huidarts.info
- The British Association of Dermatologists (BAD)
  http://www.bad.org.uk/

APPENDIX 3

The Kruskall-Wallis test (alpha 0.05) was used to explore statistically significant differences in the ranking of dermatological topics and criteria (in groups of three). Mann-Whitney U tests were used in case of $p < 0.05$ to determine for which dermatological condition the ranking was statistically different. The null hypothesis is an interchangeable place in the ranking, meaning that the tested rankings are not statistically different.

REFERENCES

1 Grimshaw J, Russell I. Achieving health gain through clinical guidelines. I: Developing scientifically valid guidelines. Qual Health Care. 1993;2(4):243-8.

2 Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42.

3 Guyatt GH, Oxman AD, Schunemann HJ, et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clinical Epidemiol. 2011;64(4):380-2.

4 Field MJ, Lohr KN. Clinical Practice Guidelines: Directions for a New Program. National Academy Press. Washington DC, 1990.

5 Ketola E, Toropainen E, Kaila M, et al. Prioritizing guideline topics: development and evaluation of a practical tool. J Eval Clin Pract. 2007;13(4):627-31.

6 Reveiz L, Tellez DR, Castillo JS, et al. Prioritization strategies in clinical practice guidelines development: a pilot study. Health Res Policy Syst. 2010;8:7.

7 Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 2.Priority setting. Health Res Policy Syst. 2006;4:14.

8 Battista RN, Hodge MJ. Setting priorities and selecting topics for clinical practice guidelines. CMAJ. 1995;153(9):1233-7.

9 van der Sanden WJ, Mettes DG, Grol RP, et al. Development of clinical practice guidelines for dentists: methods for topic selection. Community Dent Oral Epidemiol. 2002;30(4):313-9.

10 Grol R, Thomas S, Roberts R. Development and implementation of guidelines for family practice: lessons from The Netherlands. J Fam Pract. 1995;40(5):435-9.

11 Lomas J. Making clinical policy explicit. Legislative policy making and lessons for developing practice guidelines. Int J Technol Assess Health Care. 1993;9(1):11-25.

12 Hayward RS, Guyatt GH, Moore KA, et al. Canadian physicians' attitudes about and preferences regarding clinical practice guidelines. CMAJ. 1997;156(12):1715-23.

13 Eccles MP, Grimshaw JM, Shekelle P, et al. Developing clinical practice guidelines: target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest. Implement Sci. 2012;7:60.

14 Oortwijn WJ, Vondeling H, van Barneveld T, et al. Priority setting for health technology assessment in The Netherlands: principles and practice. Health policy. 2002;62(3):227-42.

15 Brouwers MC, Chambers A, Perry J, et al. Can surveying practitioners about their practices help identify priority clinical practice guideline topics? BMC Health Serv Res. 2003;3(1):23.

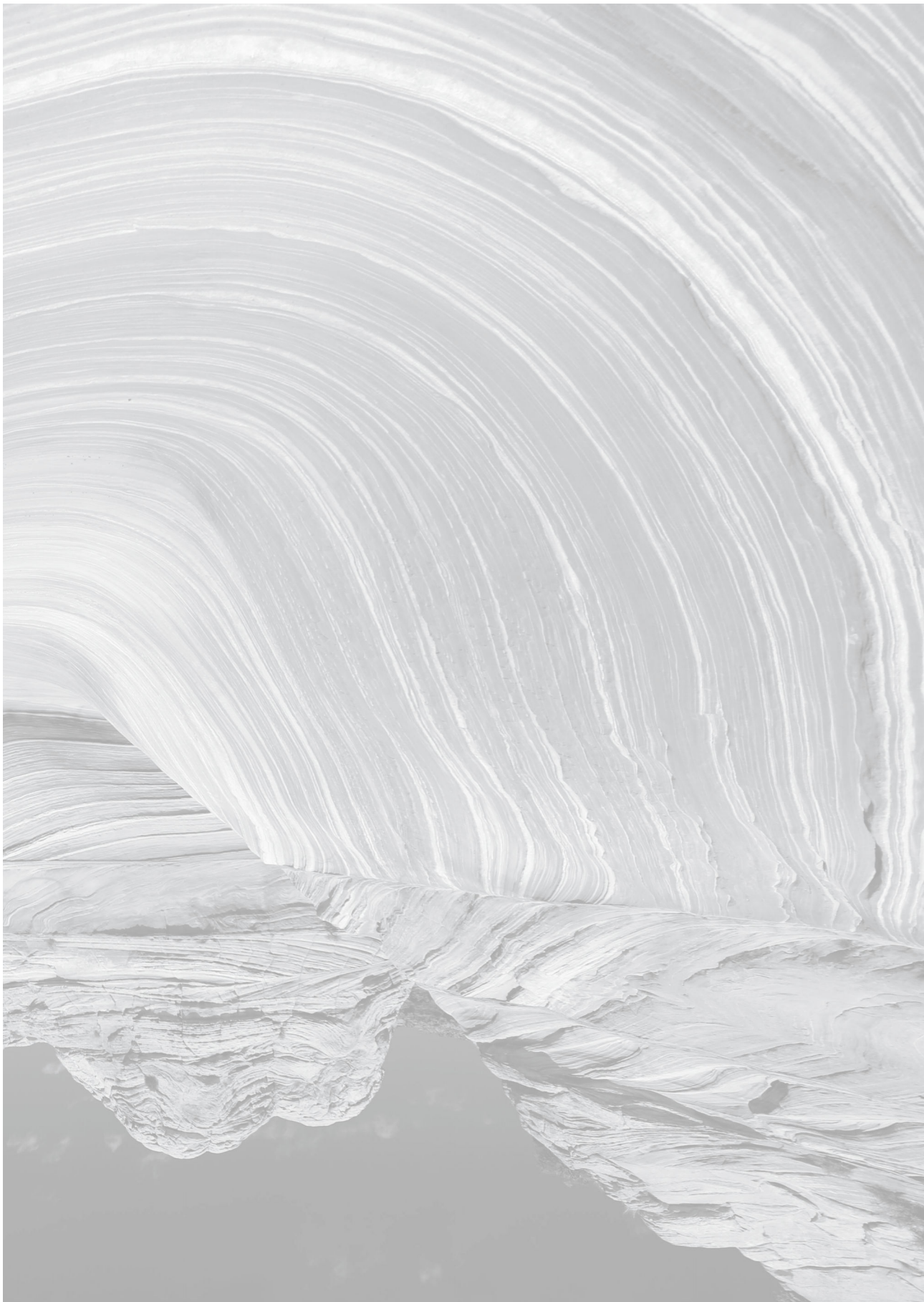
16 Burgers JS, Grol R, Klazinga NS, et al. Towards evidence-based clinical practice: an international survey of 18 clinical guideline programs. Int J Qual Health Care. 2003;15(1):31-45.

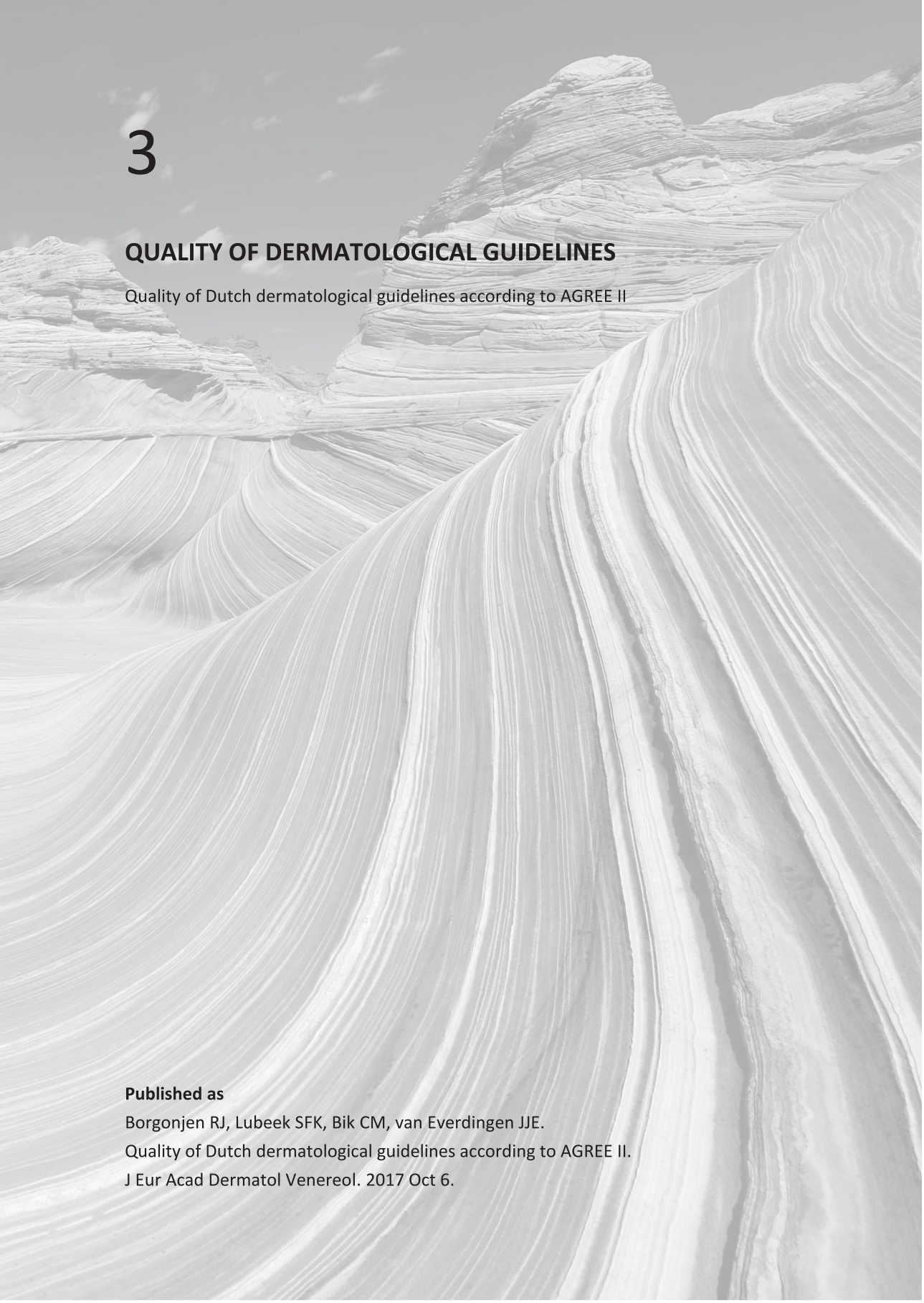
17 Weisshaar E, Szepietowski JC, Darsow U, et al. European guideline on chronic pruritus. Acta Derm Venereol. 2012;92(5):563-81.

18 Venning VA, Taghipour K, Mohd Mustapa MF,et al. British Association of Dermatologists' guidelines for the management of bullous pemphigoid 2012. Br J Dermatol. 2012;167(6):1200-14.

19 Blumeyer A, Tosti A, Messenger A, Reygagne P, Del Marmol V, Spuls PI, et al. Evidence-based (S3) guideline for the treatment of androgenetic alopecia in women and in men. J Dtsch Dermatol Ges. 2011;9 Suppl 6:S1-57.

20 Ploegmakers MMJ, de Vries Moeselaar A, Wiersma T, van Barneveld TA. (2012) Prioriteren onderwerpen voor richtlijnontwikkeling in Nederland 2010-2012. http://www.regieraad.nl/fileadmin/www.regieraad.nl/publiek/Actueel/rapporten/PRIORITERING_RICHTLIJNONDERWERPEN.pdf [accessed on 23 August 2013].





3

QUALITY OF DERMATOLOGICAL GUIDELINES

Quality of Dutch dermatological guidelines according to AGREE II

Published as

Borgonjen RJ, Lubeek SFK, Bik CM, van Everdingen JJE.
Quality of Dutch dermatological guidelines according to AGREE II.
J Eur Acad Dermatol Venereol. 2017 Oct 6.

SUMMARY

Objective To determine to what extent Dutch national dermatological guidelines meet internationally accepted quality criteria.
Methods Twenty Dutch dermatological guidelines, published between 2010 and 2016, were assessed with the AGREE II Instrument. This instrument consists of twenty-three items, grouped into six domains; 'scope and purpose', 'stakeholder involvement', 'rigour of development' and 'clarity of presentation'. Standardized domain scores were calculated according to the AGREE II manual.
Results The quality of the Dutch dermatological guidelines is on average good. The guidelines scored good to very good in the domains 'scope and purpose', 'stakeholder involvement', 'rigour of development' and 'clarity of presentation'. However the domains 'applicability' and 'editorial independence' received lower scores.
Conclusion On general Dutch dermatological guidelines meet the current international quality criteria. In particular certain aspects of applicability and editorial independence can be bettered. Possible improvements are discussed.

INTRODUCTION

Guidelines are documents containing recommendations to support healthcare professionals and healthcare users, aiming to improve the quality of healthcare. Nowadays a guideline is built on scientific research complemented with expertise and experiences from healthcare professionals and healthcare users.[1] The number of (dermatological) clinical practice guidelines is increasing. However, the quality of those guidelines is variable.[2-5] Besides adequate quality of medical content, a healthcare professional and/or user should be able to rely on a methodological sound guideline that is applicable in daily practice as well. Even more, because guidelines have become core instruments in quality policy of professional groups and are increasingly embraced by other stakeholders such as healthcare insurers and governmental institutions.[6] To objectively assess the methodological quality of guidelines, guideline experts developed the Appraisal of Guidelines and Research and Evaluation (AGREE) instrument. The original AGREE instrument is revised. The validated second version (AGREE II) evaluates the process of guideline development across six domains; 'scope and purpose', 'stakeholder involvement', 'methodology', 'clarity and presentation', 'applicability' and 'editorial independence'.[7] In 2007 eight Dutch dermatological guidelines were already assessed with the original AGREE instrument. Afterwards recommendations were formulated.[8] Also international evaluation of dermatological guidelines yielded several suggestions for improvement.[9-11] Due to expansion and revision of Dutch dermatological guidelines it is useful to see if the recommended improvements were applied. Also the availability of the revised AGREE instrument in which extra quality requirements and insights have been incorporated, makes renewed evaluation of the quality of Dutch dermatological guideline development appropriate. The aim of the present study is to determine to which extent current Dutch dermatological guidelines meet AGREE II quality criteria.

METHODS

All Dutch dermatological guidelines until January 2017 were selected and accessed using the Dutch Society of Dermatology and Venereology website (www.huidarts.info). This yielded twenty guidelines containing one or more dermatological topics (Table 1). Seven of those twenty guidelines were a revision of a guideline already assessed in 2007 using the original AGREE instrument. All twenty guidelines including available appendices were independently evaluated by at least two appraisers using the AGREE II instrument (RB, SL, CB). The AGREE II instrument consists of twenty-three items, divided across six domains (Table 1). Items are scored

individually on a seven-point Likert scale ranging from 'strongly agree' to 'strongly disagree'. Each AGREE II item contains additional information to assist in scoring. Subsequently standardised domain scores were calculated from the individual domain scores using the AGREE II manual.[12] Eventually domain scores reflect a percentage from the maximum possible score (100%) in that particular domain. An average domain score of 60% or higher was considered as good.[9-11,13]

Analysis

Every domain was analysed with the average score, standard deviation and 95% confidence interval. To asses inter-rater reliability the Intraclass Correlation Coefficient (ICC) was determined. Analyses were performed using SPSS (Statistical Package for the Social Sciences, Chicago, IL, USA) version 22.

	Scope and purpose	Stakeholder involvement	Methodology	Clarity and presentation	Applicability	Editorial independence
Acneiform eruptions (2010)	89	86	70	81	33	21
Small vessel vasculitis (2010)	83	86	75	83	23	38
Actinic keratosis (2010)	58	50	74	75	27	21
Squamous cell carcinoma (2011)	86	72	67	75	27	25
Psoriasis (2011)	81	81	73	86	42	29
Pressure ulcers (2011)	89	92	57	83	73	79
Lichen sclerosus and lichen planus (2012)	61	81	75	72	38	29
Sexually transmitted diseases (2013)	72	92	45	69	42	29
Melanoma (2012)	89	89	73	75	60	42
Lyme borreliosis (2013)	86	72	68	89	50	63
Contact dermatitis (2013)	94	61	69	89	38	42

Cellulitis and erysipelas (2013)	100	72	81	100	67	58
Lipoedema (2013)	83	72	44	94	54	17
Lymphoedema (2014)	89	89	63	94	67	17
Atopic dermatitis (2014)	83	83	60	89	50	33
Venous disease (2014)	94	89	79	94	63	50
Chronic urticaria (2015)	94	83	77	100	63	25
Basal cell carcinoma (2015)	83	67	66	92	54	67
Vitiligo (2015)	94	78	85	100	67	25
Hyperhidrosis (2016)	94	78	85	100	67	25
Average domain score	**85**	**78**	**69**	**87**	**49**	**37**
95% Confidence Interval	**81-90**	**74-83**	**64-73**	**83-91**	**43-56**	**29-44**

Table 1: Standardised domain scores (%) per domain and guideline, average domain score including 95% confidence interval; all calculated according to AGREE II manual. The guidelines are shown in chronological order. 80-100%, 60-79%, 40-59%, 20-39%, 0-19%.

RESULTS

The AGREE II scores of the twenty dermatological guidelines are presented in table 1. There are differences in domain scores. The domains 'scope and purpose', 'stakeholder involvement', 'methodology' and 'clarity and presentation' scored good to very good. However, the domains 'applicability' and 'editorial independence' lagged. The Intraclass Correlation Coefficient (two way mixed and absolute agreement) as a measure for the agreement between reviewers was 0.84 (95% confidence interval 0.72-0.96). A number above 0.7 is seen as sufficient.

Scope and purpose

This domain considers whether the aim of the guideline, the questions that are being discussed and the population the guideline covers are described explicitly. The average domain score of 85% (95%-confidence interval 81-90) was very good. Still, the guidelines 'Actinic keratosis' (58%), and 'Lichen sclerosus and lichen planus' (61%) could have been formulated more explicitly.

Stakeholder involvement

In this domain the AGREE II instrument considers whether all the relevant professionals are part of the guideline development group and whether the view and preferences of the target population have been taken into account. It is important to assemble a carefully balanced working group in which all relevant factions are represented, in order to create as much support for the guideline as possible in advance. This domain achieved a very good score; 78% (95% confidence interval 74-83). Most of the selected dermatological guidelines are guidelines in which numerous relevant disciplines are involved. The target users of the guidelines are explicitly mentioned. Nevertheless, the majority of the guidelines lost some points because the roles of the working group members were not explicitly mentioned. For example, it was not clear who worked on a specific question or who made a particular (evidence) table or summary. The guidelines 'Actinic keratosis', 'Basal cell carcinoma', 'Contact dermatitis' and 'Cellulitis and erysipelas' achieved a lower score in the item 'the views/preferences of the target population have been sought'. The "Psoriasis" guideline on the other hand included an extensive literature search on patient views, a focus group discussion with patients, a subsequent questionnaire among patients and patient representatives were part of the guideline working group. Though, it might just be easier to find patients willing to cooperate in a guideline about a chronic condition with an active patient association.

Methodology

The methodology domain is the most extensive domain, wherein it is tested whether systematic methods have been used to search for evidence. Besides that the selection criteria, the strengths and limitations of the body of evidence and the methods used to formulate recommendations must be described explicitly and clearly. In addition, in order to achieve a high domain score, the health benefits, side effects and risks must have been considered in formulating the recommendations. Also an explicit link between the recommendations and the underlying evidence should be present. Furthermore, ideally the guideline should have been reviewed by external experts before publication and a procedure to update the guideline ought to be provided. The average domain score was good; 69% (95% confidence interval 64-73). All guidelines described a systematic literature search, but not in all cases the search strategy and in- and exclusion criteria were attached. The methods used to formulate recommendations were sometimes less clear in the guidelines "Sexually transmitted diseases" and "Lipoedema", with a corresponding domain score of respectively 45 and 44%. Achieving a higher domain score was possible in some guidelines by mentioning

how feedback from expert review had been incorporated. Also the procedure of updating a guideline was not mentioned in all guidelines and was in other cases often not entirely clear. A minority of the guidelines used GRADE (Grading of Recommendations, Assessment, Development and Evaluation tool) as method to formulate several of its recommendations.[14] Grading recommendations can help to harmonise (the interpretation of) recommendations and clearly depict the link with the body of evidence.

Clarity and presentation

AGREE II scores in this domain are high as soon as the following conditions are met: the recommendations are specific and unambiguous, the different options for management of the condition or health issue are clearly presented and the key recommendations are easily identifiable. This was the case in almost all guidelines, partly because key recommendations were placed in separate frameworks or summaries. This resulted in a very good domain score of 87% (95% confidence interval 83-91). Recommendations can be specified by substituting words and phrasings such as 'usually', 'large number', 'when sufficient… are present' with more explicit ones.

Applicability

The domain applicability shows whether the guideline provides advice and/or tools for applying the recommendations in daily practice and whether facilitators and barriers to its application are described. This domain also checks if resource implications of applying the recommendations have been considered. Furthermore, it is important that monitoring and/or auditing criteria are presented. The domain score of 49% (95% confidence interval 43-56) can be regarded as moderate. Especially the more recent guidelines from 2014 onwards and the guidelines 'Melanoma' and 'Pressure ulcers' gave attention to implementation barriers and resource implications. Almost all guidelines do have a summary and separate information for healthcare users.

Editorial independence

This domain looks if the views of the funding bodies have influenced the content of the guideline and if competing interests of guideline development group members have been recorded and addressed or not. The average domain score was 37% (95% confidence interval 29-44). The funding body is mentioned in all guidelines, but in many cases a lack of information allows no further judgment about influence on the content or its developers. The guideline 'Pressure ulcers' scored the highest (79%) due to an extensive report about competing (financial) interests, a screening that can be

requested as well as declarations from the individuals members of the guideline development group.

DISCUSSION

The analysed Dutch dermatological guidelines all have a systematic search in literature and recommendations formulated with practical considerations in mind. Therefore they comply with the principles of evidence based guideline development. The AGREE II instrument is developed according to these principles to systematically appraise the quality of guidelines. A limitation in the use of the AGREE instrument is that it does not asses the quality of the content of a guideline recommendation. Just reporting an item is often enough to reach a high score. Furthermore, it is possible that the actual procedure deviates from the one that is reported. Nevertheless, the AGREE II instrument offers guideline developers support which items must be incorporated into a high quality guideline. It reassures the clinician that the recommendations and (scientific) background information in a guideline are realised in a valid way.
The individual items of the AGREE II instrument served as a basis for the Dutch 'Guideline for guidelines' written in command of the former Directing council Quality of care (Dutch: 'Regieraad Kwaliteit van Zorg') and the Medical Specialists Guidelines 2.0 in command of the Dutch Federation of Medical Specialists.[1] When the Directing council Quality of Care was abolished in 2012 and the Dutch Health Institute took over, the group responsible for the Guideline for Guidelines was installed as the Advice- and expert group on QUAlity standards (AQUA). This Advice- and expert group started to publish its yearly updated guidance on quality standards in 2014.

All these documents emphasise the need to give ongoing attention to implementation and resource implications in guidelines. The comparatively lower scores in the domains applicability and editorial independence are also demonstrated in earlier publications using the AGREE instrument to analyse guidelines.[13] These domains also received the lowest scores of respectively 44% and 38% in the review of Dutch dermatological guidelines in 2007 by Hukkelhoven et al.[8] However, the original version of the AGREE instrument used in 2007 scored item on a 4 point Likert scale. Therefore, comparing individual scores to the current 7 point Likert scale is difficult.
Internationally, there are a couple of papers that evaluate dermatological guidelines using the AGREE II instrument. In a review of seven evidence based European Dermatology Forum (EDF) guidelines average domain scores of 51 to 84% were reached. Older and less rigorously developed EDF guidelines scored substantially

lower and reached average domain scores of 39 to 61%.[15] Other international reviews appraised actinic keratosis and acne guidelines.[9,16]
It is noteworthy that the Dutch acne and actinic keratosis guidelines are missing in those reviews. Omitting Dutch guidelines in international publications is described previously by Spuls et al. in the case of psoriasis.[17] Consequently, it is vital that every Dutch guideline should have at least an English abstract and is actively added to international databases. This can save costs and effort for other guideline developers.[9,17] In this paper standardised domain scores were calculated according to the AGREE II manual, but analysing any differences in item scores and subsequently reaching consensus can certainly be meaningful as well. The agreement between reviewers also seems to depend on both the familiarity with the assessed guidelines and the AGREE II manual.[15,17]

Applicability
A lot of resources, expertise and above all human effort are needed to develop a guideline that reaches a high score in every item and domain of the AGREE instrument. The domain applicability was one of the two domains that had a lower average score. Yet, the application of a guideline in daily practice is vital to achieve the quality improvement guidelines strive for. It is a prerequisite to asses and document implementation barriers from the start. Composing an implementation plan in advance can be useful. This plan should contain strategies that enhance support and evaluation like guideline based audit, feedback, decision aids and continuing medical education. However, designing process indicators only because it is part of the procedure is useless.

Editorial independence
Cause of the lower domain score was mostly the lack of information about potential competing interests. This seems to be a broader issue, since a review of German dermatological guidelines found a declaration about competing interests in only 18% of the guidelines.[18] When a declaration about competing interests is absent, a majority of guideline authors nonetheless has some business with pharmaceutical companies. 7% of these guideline authors stated that their business with the pharmaceutical industry influenced the recommendations they formulated, although 19% of the authors thought that their co-authors were influenced.[19] Though most guideline authors are undoubtedly trustworthy, it is crucial to remove any doubts about conscious or unconscious conflict of interests.

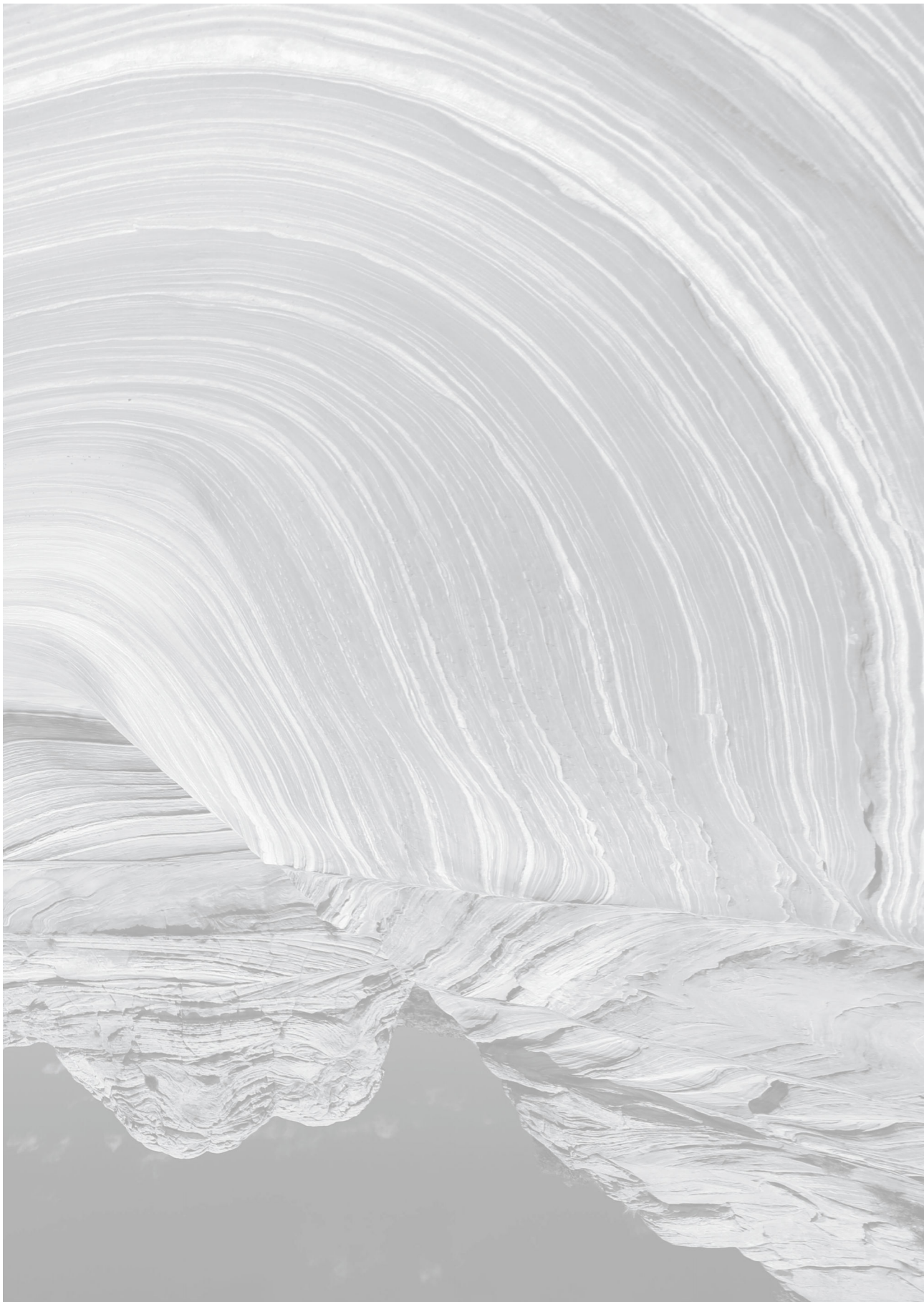
A strategy consisting of documenting how competing interests are handled, depicting any competing interests and non-participation in the case of sensitive group discussions can be useful.[18]

CONCLUSION

In general Dutch dermatological guidelines are methodological sound and meet the current international quality criteria described in the AGREE II instrument. Awareness of certain aspects of the domains 'applicability and 'editorial independence' is necessary to accomplish the improvement of healthcare guidelines strive for. Guideline development remains a means to an end.

REFERENCES

1 Werkgroep Richtlijn voor Richtlijnen; in opdracht: Regieraad Kwaliteit van Zorg. Richtlijn voor Richtlijnen. Den Haag, 2010.
2 Grilli R, Magrini N, Penna A et al. Practice guidelines developed by specialty societies: the need for a critical appraisal. Lancet 2000;355:103-106.
3 Shaneyfelt M, Mayo-Smith MF, Rothwangle J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA 1999;281:1900-1905.
4 Kung J, Miller RR, Mackowiak PA. Failure of clinical practice guidelines to meet institute of medicine standards: two more decades of little, if any progress. Arch Intern Med 2012;172:1628-1633.
5 Burgers JS, Fervers B, Haugh M et al. International Assessment of the Quality of Clinical Practice Guidelines in Oncology Using the Appraisal of Guidelines and Research and Evaluation Instrument. J Clin Oncol 2004;22:2000-7.
6 Buiting AMJ, van Barneveld T, Bartels PCM. Richtlijnen als basis voor kwaliteitsontwikkeling. Ned Tijdschr Klin Chem 2003;28:192-196.
7 Brouwers M, Kho ME, Browman GP, et al.; AGREE Next Steps Consortium. AGREE II: Advancing guideline development, reporting and evaluation in healthcare. Can Med Assoc J 2010;182:E839-842.
8 Hukkelhoven CWPM, Burgers JS, van Everdingen JJE et al. De kwaliteit van dermatologische CBO-richtlijnen beoordeeld met het AGREE Instrument. Ned Tijdschr Dermatol 2007;17:22-7.
9 Sanclemente G, Acosta JL, Tamayo ME et al. Clinical practice guidelines for treatment of acne vulgaris: a critical appraisal using the AGREE II instrument. Arch Dermatol Res 2014;306(3):269-77.
10 Nast A, Spuls Ph, Ormerod AD et al. A critical appraisal of evidence-based guidelines for the treatment of psoriasis vulgaris: 'AGREE-ing' on a common base for European evidence-based psoriasis treatment guidelines. JEADV 2009;23:782–787.
11 de Haas ERM, de Vijlder HC, Siewertz van Reesema W et al. Quality of clinical practice guidelines in dermatological oncology. JEADV 2007;21:1193–1198.
12 AGREE Next Steps Consortium. AGREE II. Instrument voor de beoordeling van richtlijnen. 2009. [English electronic version available: http://www.agreetrust.org].
13 Alonso-Coello P, Irfan A, Solà I et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. Qual Saf Health Care 2010;19(6):e58.
14 Guyatt GH, Oxman AD, Vist GE et al.; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;26;336(7650):924-6.
15 Werner RN, Marinović B, Rosumeck S et al. The quality of European dermatological guidelines: critical appraisal of the quality of EDF guidelines using the AGREE II instrument. J Eur Acad Dermatol Venereol 2016;30(3):395-403.
16 Kirby JS, Scharnitz T, Seiverling EV et al. Actinic Keratosis Clinical Practice Guidelines: An Appraisal of Quality. Dermatol Res Pract 2015;2015:456071.
17 Spuls Ph, Nast A. Evaluation of and Perspectives on Guidelines: What Is Important? Journal of Investigative Dermatology 2010;130:2348-2349.
18 Rosumeck S, Sporbeck B, Rzany B, Nast A. Disclosure of potential conflicts of interest in dermatological guidelines in Germany - an analysis - status quo andquo vadis. J Dtsch Dermatol Ges 2011;9(4):297-304.
19 Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. JAMA 2002;287:612–17.



4

TO A GUIDELINE FOR OFF-LABEL CYCLOSPORIN USE; A GRADE APPROACH

Off-label use of cyclosporin in dermatology: a GRADE review

To be submitted as

Bik CMJM*, Borgonjen RJ*,van Leeuwen H, Schram ME, van der Schroeff JG, van Everdingen JJE, Spuls PI.

Off-label Use of Cyclosporin in Dermatology: a GRADE review.

*Both authors contributed equally to this work

SUMMARY

Objective To summarize evidence regarding effectiveness, efficacy and safety of off-label cyclosporin use in dermatology.

Methods A literature search in Medline, EMBASE and CENTRAL was performed using cyclosporin and its synonyms as main search terms. Articles about the use of cyclosporin in an off-label dermatologic setting were independently assessed for eligibility by 2 authors. The initial search retrieved 9894 articles, and 297 articles were selected for detailed review. 79 articles matching the inclusion and exclusion criteria were included to evaluate methodologic quality and extract outcome data using GRADE (Grading of Recommendations, Assessment, Development and Evaluation).

Results The quality of evidence was mostly very low or low across a broad variety of outcome measures. Regarding chronic urticaria quality of evidence was moderate for some outcome measures. Subsequently, a weak recommendation could be formulated for treating alopecia areata, morbus Behcet, chronic actinic dermatitis, chronic contact dermatitis, chronic dermatitis of the hand, chronic urticaria, dermatomyositis, eosinophilic pustular folliculitis, Langerhans cell histiocytosis, leprosy type 1, lichen planus, lichen sclerosus, livedoid vasculitis, pemphigus vulgaris, palmoplantar pustulosis, prurigo, pyoderma gangrenosum, sarcoidosis, systemic sclerosis and toxic epidermal necrolysis with cyclosporin. A weak recommendation against was formulated for cutaneous T-cell lymphoma and hidradenitis suppurativa.

Conclusion

Literature based evidence for off-label cyclosporin use exists in 22 different dermatological conditions. Overall, the quality of the articles was low; downgrading due to sparse data or indirectness was common. Therefore, only weak recommendations could be given. Prospective long-term comparative registries and prospective studies should add useful evidence and provide legal support for off-label drug use in dermatology. Harmonising outcomes by defining core outcome sets is necessary.

INTRODUCTION

Cyclosporin (also known as cyclosporin A) (CsA) is an immunomodulating agent widely used in medicine today. It was the first immunosuppressive drug that allowed selective immunoregulation of T cells without excessive toxicity. Cyclosporin was isolated from the fungus *Tolypocladium inflatum* and was originally intended as an anti-fungal agent. Borel discovered its immunosuppressive activity in 1976.[1] Therapeutic effects of cyclosporin include a steroid-sparing effect, thereby reducing the toxicity associated with high dosage and prolonged usage of corticosteroids.[2] In the United States, cyclosporin is registered for the following indications; kidney-, heart- and liver transplantation, rheumatoid arthritis and psoriasis.[3] In most European countries cyclosporin is licensed for the treatment of atopic dermatitis and psoriasis.[4] It is extensively used off-label.[5] In an off-label setting, patients are treated with drugs that are not registered for that specific indication, population or dosage.[6,7] Only pharmaceutical companies can apply for registration of drugs for specific indications. However, the interests of pharmaceutical companies and those of physicians and patients often differ regarding the registration of drugs in difficult-to-treat and refractory dermatologic conditions and/or in rare conditions with subsequently scarce therapeutic options. As a result, in daily practice, many drugs are prescribed off-label, including 17-73% of dermatological prescibtions.[8] International and national regulatory agencies are applying rules and regulations concerning off-label medication use (http://www.fda.gov, http://www.mhra.gov.uk, and http://www.rivm.nl).[6]
To ensure evidence-based healthcare decision making, guidelines should be developed for drugs that are often prescribed off-label. Because cyclosporin is one such drug, we aimed to summarize the available evidence in the literature regarding the efficacy, effectiveness, and safety of off-label treatment with cyclosporin in dermatologic patients. Based on this evidence and taking clinical experience into account, we formulated clinical recommendations for the application of cyclosporin in specific dermatological conditions.

METHODS

Inclusion and exclusion criteria

Included were outcome data from randomized controlled trials (RCTs), cohort studies, and case series with more than 5 patients, in which dermatologic patients were treated with off-label cyclosporin, in order to determine the efficacy, effectiveness, and/or safety of cyclosporin with or without concomitant therapy. Excluded were reviews, case series with fewer than 5 patients, animal and in vitro studies, abstracts,

unpublished data, articles concerning topical treatment, articles concerning diseases primarily treated by other specialists, and duplicate publications. No restrictions were imposed regarding age, sex, skin type or date of publication. Only articles in English, French, German or Dutch were included.

Predefined review protocol

Literature Search

On June 28th, 2016, a literature search in the MEDLINE, EMBASE, and CENTRAL databases was performed. The keyword *cyclosporin* and its synonyms were used in combination with skin diseases (Table 1). Reference lists of included articles and reviews regarding cyclosporin were hand-searched to find additional eligible articles.

No.	Term(s)
1	(cyclosporin* or ciclosporin* or neoral* or sandimmun*).ti.
2	(cyclosporin* or ciclosporin* or neoral* or sandimmun*).ab.
3	exp skin disease/
4	*cyclosporin/
5	1 or 2 or 4
6	3 and 5
7	derm*.jn.
8	5 and 7
9	6 or 8
10	limit 9 to (human and (Dutch or English or French or German))

Table 1: Search Strategy for MEDLINE

Study Selection and Data Extraction

The titles and/or abstracts of the search results were screened for off-label cyclosporin treatment in dermatologic patients and selected if outcome data were present. To determine eligibility, the full text of the selected articles was assessed according to the predefined inclusion and exclusion criteria. Data concerning methodologic quality, study characteristics including study population (number, age range, sex of included patients, dermatologic disease concerned, previous treatments, and inclusion criteria for participation), design of the study (study type, prospective-retrospective, duration of treatment, dosage of cyclosporin, concomitant medication and duration of follow-up period), publication date, onset of action and duration of remission and outcome data regarding efficacy/effectiveness and safety were extracted. Preferably, direct outcomes graded by the authors as important or critical

for a specific skin disease were reported. If this was not possible indirect outcomes and other outcome measures were presented. All stages of study selection and data extraction were performed by 2 reviewers independently. Disagreements regarding study selection and data extraction were solved by discussion. Outcome data is depicted in GRADE (Grading of Recommendations, Assessment, Development and Evaluation) Summary of Finding tables (Addendum 1).

Data analysis

We assessed RCTs according to the criterion grading system described in the Cochrane Handbook for Systematic Reviews of Interventions and the GRADE handbook.[9,10] To assess the risk of bias, the following items were considered: random sequence generation, concealment of allocation, blinding (of participants, researchers and outcome assessment), incomplete outcome data and selective reporting. Each item was scored as adequate, inadequate, or unclear (Addendum 2). In turn the risk of bias table served as input for possible downgrading according to the GRADE risk of bias (severe or very severe). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist for observational studies was used to identify methodologic errors in cohort studies.[11] Case reports were only assessed for methodologic quality using the eighth GRADE criteria for assessing the quality of evidence (risk of bias, directness of evidence, consistency and precision of results, risk of publication bias, magnitude of the effect, dose-response gradient, and influence of residual plausible confounding). The quality of evidence was finally determined using GRADE. Because it is inappropriate to formulate safety conclusions based on information derived from studies with relatively small case numbers and short follow-up periods, we compared the adverse events (AEs) and serious adverse events (SAEs) reported in the included studies with the safety data from the Summary of Product Characteristics dated 10th March 2016.[12] Attention was given to the identification of AEs that seemed to be specific for an indication.

Clinical recommendations

Recommendations are characterized as strong or weak according to the quality of the supporting evidence and the balance between desirable and undesirable consequences of the alternative management options.[13] Considerations such as patient preferences, availability of other treatment options, organizational aspects, social consequences, and costs were assessed independently by two authors to generate balanced clinical recommendations.

An important consideration regarding off-label use is its safety. Known adverse events mentioned in the summary of product characteristics (SPC) are taken into account, as far as they were not already distilled from literature.

RESULTS

Study selection

An initial wide search retrieved 9894 articles. Checking for additional articles in references did not generate any new items. After screening the title and/or abstract for eligibility, 297 articles were selected (Figure 1). The main reason for exclusion was lack of relevance (e.g. use of cyclosporin in registered indications). After screening the full texts of the articles, 79 articles with potentially relevant outcome data matched our inclusion and exclusion criteria, of which 7 were RCTs, 14 were cohort studies, and 58 were case series. In total, 1992 patients with 22 different dermatologic diseases were described.

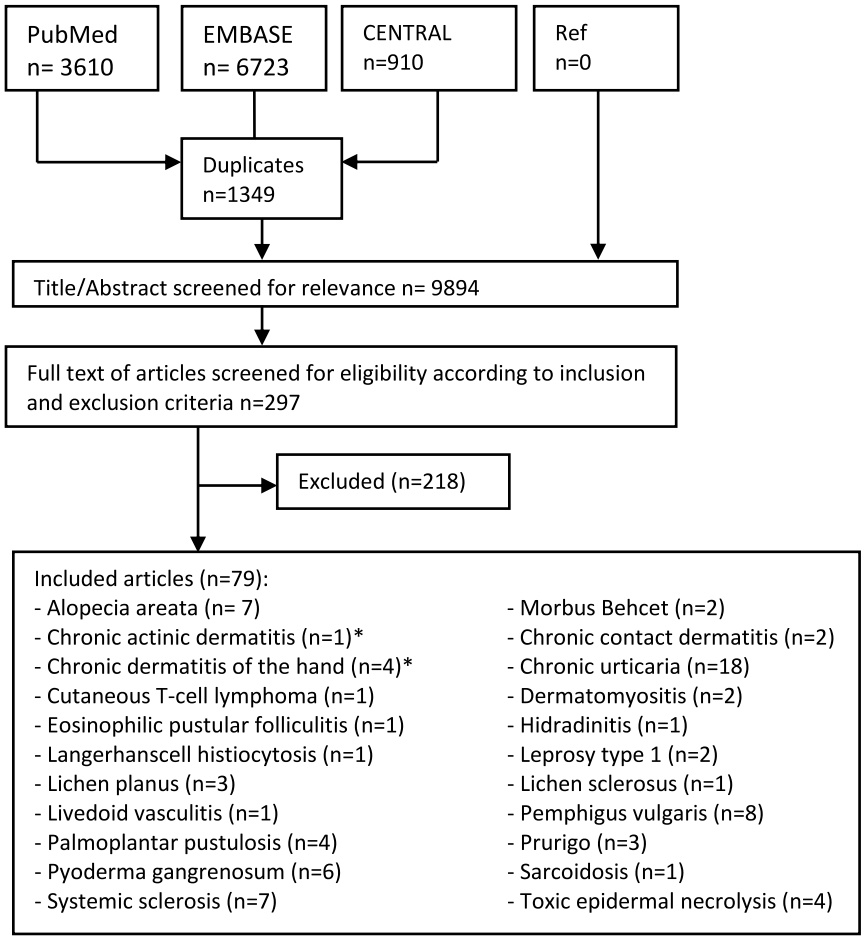
Outcome data

Diagnostic criteria and disease definitions were often not specified. The outcome measurements used to asses efficacy and effectiveness were heterogeneous. Outcomes measures included in this systematic review were; urticaria activity scores, autologous serum skin tests (ASST), a clinical severity score, patient visual analogue scores (VAS) rating the overall severity of urticaria, clustered physician's global assessment (PGA) scores, a severity score index, unspecified QoL (quality of life) scores, an itch intensity scale, immunohistochemic analysis, skin scores (in all studies in which this score was used the skin score was scored differently, i.e. symptom scoring scale, elastometry, plicometry, patients opinion and/or physician's general assessment), time to arrest, time to reepithelialisation, severity-of-illness score for toxic epidermal necrolysis (SCORTEN) and the standardized mortality ratio (SMR). In most of the studies however, only descriptive means were given to indicate changes in disease severity. Originally, important and critical outcomes should be established as GRADE evaluates the body of evidence per outcome and not per study. Eventually all outcomes and outcome measures were considered due to the lack of evidence on important outcome(s) (measures). Consequently, these surrogate outcomes were frequently downgraded because of indirectness. Onset of action varied from 2 to 11 days after the initial dosage. Duration of remission during and after treatment varied from 0 to 72 months. Concomitant treatment was allowed in most studies and not reported in 21 studies. Because cyclosporin is often regarded and used as a

corticosteroid-sparing agent, oral corticosteroids were most frequently used as concomitant medication (n=20). Oral corticosteroids were also used to support patients between initiation and the onset of action/therapeutic effect of cyclosporin, after which the corticosteroid dosage was tapered. A definition of onset of action was mostly absent. Topical corticosteroids were allowed in 1 study.
Results are summarized in Summary of Finding tables per disease and per outcome(measure) (Addendum 1).

Figure 1: Study flowchart

*Article reporting cyclosporin efficacy in patients with 2 different dermatologic diseases; discussed separately for each disease.



Summary of findings and recommendations

The Summary of Findings are discussed per disease in alphabetical order. After every textual Summary of Findings a box shows the eventual recommendation. The final formulated recommendation is the result of the quality of evidence in combination with other considerations and can be formulated as a weak or strong recommendation in favour of a certain therapy or as a weak or strong recommendation against a certain therapy (see box below). The aim of this procedure is to enhance the transparency of the guideline. To provide further insight general considerations that are valid for all skin conditions treated with off-label cyclosporin are depicted below. Disease specific considerations are highlighted after the recommendation.

Recommendation
Strong = if clinicians, based on the available evidence, are very certain that benefits do, or do not, outweigh risks or burdens, they will make a strong recommendation. Weak = if clinicians, based on the available evidence, believe that benefits and risks or burdens are finely balanced, or if considerable uncertainty exists about the magnitude of benefits and risks, they must make a weak recommendation.

General considerations
Risks associated with therapy Off-label use. Potential disease and population specific (serious) adverse events not known from SPC.
Burdens of Therapy Follow-up visits and regular laboratory and blood pressure monitoring. Daily oral tablet. Potential burden of side effects.
Costs Between € 169,64 and 169,92 for 1 month 2dd 100 mg. Not included are the costs for delivery, laboratory monitoring and visits to the clinic. (www.medicijnkosten.nl). Costs may vary with the number of follow-up visits and dosage of CsA.

Alopecia areata

Seven case series, concerning 124 alopecia patients in total, were included.[14-20] Four series used CsA with concomitant prednisolone.[17-20] The dose of CsA employed ranged from 2.5-7.5 mg/kg/day in six studies and Kim *et al.* administered 200 mg/day. In total 11 patients were lost to follow up. Duration of treatment varied from 2 to 12 months and the follow-up period could be as long as 20 months. Regrowth of vellus or terminal hair was used as outcome to indicate changes in disease severity. 25 to 75% of the patients had a favourable response to cyclosporin. In Constantopoulos *et al.* all 5 patients began to lose their hair 2 months after stopping CsA treatment. 7 of the 124 patients stopped treatment due to adverse events. Three serious adverse events were reported.

There is a weak clinical recommendation for the use of cyclosporin in alopecia areata.

Alternative treatment for progressive alopecia areata in patients experiencing a great impact on quality of life. *Considerations* Very low quality of evidence Very uncertain estimate for a moderate effect Only a few other treatment options available Alopecia areata often has a benign prognosis

Behçet disease

Two case series (Avci et al. and Diaz-Llopis et al.), concerning 38 patients in total, were found in which patients with Behçet disease (BD) were treated with CsA.21,22 The dose employed was 5-7 mg/kg/day. In both series appearance of ulcerations was used as outcome. Diaz-Llopis et al. showed a decrease of buccal ulcerations and chronic cutaneous vasculitis lesions.[22] Avci *et al.* found a 75% response in patients with oral ulcerations.[21] Acneiform lesions did not occur in 17 out of 18 patients during treatment and 17 out of 20 patients did not show any erythema nodosum-like lesions at the time of treatment. In total, 11 adverse events, of which none were serious, were reported in these case series. Nausea and fatigue were reported most. After 12 months of treatment one patient dropped out due to nephrotoxicity.

There is a weak clinical recommendation for the use of cyclosporin in morbus Behcet.

If conventional treatment options are contra-indicated or have failed in chronic Behçet.
Considerations
Very low quality of evidence
Very uncertain estimate for a low to moderate treatment effect

Chronic actinic dermatitis
One case series was included with 6 chronic actinic dermatitis patients.[23] The dose employed ranged between 1.5 and 4.0 mg/kg/day. Outcome measurements were descriptive. The overall response is good in 4 patients, moderate in 1 and poor in another 1. No adverse events were mentioned.

There is a weak clinical recommendation for the use of cyclosporin in chronic actinic dermatitis.

Only when conventional therapy fails and a quick response is needed in severe exacerbations.
Considerations
Very low quality of evidence
Very uncertain estimate for a moderate effect
More research is needed to give any further advise
Only a few other treatment options available

Chronic contact dermatitis
Two case series, concerning 21 patients, were found in which patients with chronic contact dermatitis were treated with CsA.[24,25] The dose employed was 5.0 mg/kg/day in both studies. After 4 weeks Flori *et al.* lowered the dose to 3.0 mg/kg/day.[25] Outcome measures were the percentage of patients with clinical improvement or a reduction in a score of clinical manifestations. Manifestations on which could be scored (0-3) were: itch, burning, erythema, oedema, vesicles, crusts, scaling and

infiltration. Both studies found clinical improvement in all patients. Higgins *et al.* reported 1 mild adverse event; distal paraesthesia.

There is a weak clinical recommendation for the use of cyclosporin in chronic contact dermatitis.

Alternative treatment when non-systemic treatments and avoidance of exacerbating stimuli have failed. *Considerations* Very low to low quality of evidence Very uncertain estimate for a good effect

Chronic dermatitis of the hand

One RCT (Granlund et al. 1996), 1 follow-up study (Granlund et al. 1998) and 2 case series (Reitamo et al. and Christoffers et al.) were found in which 150 patients with chronic hand dermatitis and chronic hand eczema were treated with CsA.[23,26-28] CsA dosage was 1.25-3.0 mg/kg/day. Outcome was measured by descriptive means or using the clustered physician's global assessment (PGA) score. Two treatment arms were compared in Granlund et al. 1996. The first arm consisted of CsA with plain cream, the other of BDP cream (betamethasone-17,21-dipropionate) with placebo capsules. After the first phase there was crossover between the treatment arms if therapy failed. The results show 50% treatment success in the CSA group and 32% treatment success in the BDP group in the first phase. Although there is a difference, it is not statistically significant ($p = 0.233$). After crossover, the second phase, there was 67% treatment success in the CSA group and 62% treatment success in the BDP group. This difference was not significant. During follow-up 50% of both groups relapsed. In total 7 patients withdrew, 4 from the CSA group and 3 from the BDP group. In Christoffers et al. 63% of the patients improved at least 50% in their hand eczema (PGA). In the other case series at least (86%) of the patients responded to therapy. Two patients stopped treatment after 3 weeks due to adverse events; headache, hypertension and paresis of the abducens nerve.

There is a weak clinical recommendation for the use of cyclosporin in chronic dermatitis of the hand.

If conventional treatment options are contra-indicated or have failed.
Considerations
Very low to low quality of evidence
Uncertain estimate for moderate effect
Great heterogeneity between studies and definitions of hand eczema: subgroups such as acrovesicular or hyperkeratotic hand eczema could not be specified.

Chronic urticaria

In total, 18 studies were found in the literature in which patients with chronic urticaria were treated with CsA; 2 RCT's, 5 cohorts and 11 case series.[29-46] Outcome measurements used in the studies were very diverse. The main outcome, efficacy/effectiveness, used the percentage of patients with a complete or partial response. Other outcome measures were; response measured by urticaria activity scores, the autologous serum skin test (ASST), clinical severity score, and a 10-cm visual analogue score (VAS) indicating the overall severity of a patients urticaria. Also, the time to remission, duration of remission, time to relapse and adverse events were used to assess efficacy.

All 18 studies concerning cyclosporin in urticaria patients showed an overall decrease in disease severity and other measured outcomes after treatment with cyclosporin. Two included RCTs showed that cyclosporin treatment had results significantly more favourable than placebo treatment.[29,30]

Grattan et al. compared CsA with placebo.[29] Mean reduction in the outcome measure UAS (urticaria aggregate score) from baseline at 4 weeks was 12.7 (95% confidence interval (CI) 6.6-18.8) for patients randomized to active medication and 2.3 (95% CI 23.3 to 7.9) for placebo (p 0.005). The reduction in the outcome VAS at 4 weeks for active (mean 3.0, 95% CI 1.5-4.5) and placebo was also significant (p 0.026). Mean weal numbers and itch went down within 5 days of starting active and placebo capsules, and this improvement was sustained during active treatment. The number of chronic urticaria patients (as defined in this study) who would need treatment (NNT) with cyclosporin to achieve one response at 4 weeks was 2.4 (95% CI 1-6-5.0). Of the responders 75% relapsed after stopping CsA treatment. After 4 weeks 17 non-responding patients participated in an open label study. Eleven out of 17 (64.7%) showed remission. Overall there were 19 responders, of which 5 did not relapse during follow-up.

In Vena et al. CsA mono therapy was compared with placebo in three treatment arms.[30] The two arms, where CsA was given showed a statistically significant improvement. In total 38 patients withdrew; 8 from the first group, 13 from the second group and 17 from the placebo group. A meta-analysis of the RCTs could not be performed.

Five cohorts were included.31-35 In Di Leo et al. patients were divided into 4 groups, the groups received a different dosage of CsA; up to 1 to 3 mg/kg/day.[31] The outcome mean total symptom severity score decreased the most in the groups with the highest dosage; by 63% in 1 mg/kg/day, 76% in 2 mg/kg/day, and 85% in 3 mg/kg/day. During CsA treatment, in Di Gioacchino et al., the clinical severity score dropped significantly.[32] The ASST was negative in 13 patients and positive in 3 of 16 patients, with total remission of symptoms. Significant score reduction was observed also in patients experiencing relapses that resolved spontaneously ($p = 0.005$) or with antihistamines ($p = 0.03$). Two patients failed to complete the study.

In all the other cohorts and the 12 included case series, descriptive means were used to report changes in disease severity.[33-46] In those studies, at least 57% of the patients had a favourable response to cyclosporin, although no clear definition was given. A few mild adverse events were reported in the cohorts and case series. Nausea and vomiting (e.g. gastro-intestinal complaints) were reported most. In total twenty subjects dropped out due to high blood pressure. Only Vena et al. described 2 SAE; 1 with precordialgia and 1 with acute gastroenteritis. Both patients had to be admitted to the hospital and recovered completely. The events were not considered to be drug related.[30]

There is a weak clinical recommendation for the use of cyclosporin in chronic urticaria.

Only if antihistamines in high, off-label, dosages fail.
Only in severe recalcitrant chronic urticaria.
Considerations
Very low to moderate quality of evidence
Uncertain estimate for a good effect
Cyclosporin has a more unfavourable safety profile than antihistamines

Cutaneous T-cell lymphoma

One case series with 11 patients was included in which patients with cutaneous T-cell lymphomas (CTCL) were treated with CsA.[47] Cooper et al. treated CTCL patients with CsA 15 mg/kg/day. Outcome measure was the disappearance of all clinical evidence of active tumour for a minimum of 4 weeks. They found a complete response in 2 out of 11 patients. Both responding patients had recurrent disease 1 week after stopping CsA. They responded well on recontinuation with CsA, but both patients had to stop therapy due to renal toxicity. One serious adverse event was reported; a patient received additional glyburide, an oral hypoglycaemic agent, and died after 1 week of therapy due to complications attributed to hypoglycaemia.

There is a weak clinical recommendation against the use of cyclosporin in cutaneous T-cell lymphoma.

Considerations
Very low quality of evidence
Very uncertain estimate for a low effect
High off-label dosage with extra safety concerns
The course of the disease probably remains unchanged

Dermatomyositis

Two studies, one cohort and one case series (Grau et al. 1994, Danko et al. 1991), were found in which patients with dermatomyositis were treated with CsA.[48,49] The dosage CsA that was employed was 3-5 mg/kg/day. Complete remission, clinical improvement and therapeutic failure were used as outcome measures. In Grau et al. two groups were compared; CsA and prednisone plus azathioprine.[48] In the CsA group 1 patient failed to achieve complete remission. In the CsA group 1 serious adverse event was reported; non-oliguric renal failure. Danko et al. found remarkable improvement in 6 out of 10 patients and good improvement in the remaining 4 patients.[49] It was possible to decrease the prednisone maintenance dosage in all patients. Nausea and hypertrichosis was reported as adverse event.

There is a weak clinical recommendation for the use of cyclosporin in dermatomyositis.

If conventional treatment options are contra-indicated or have failed.
Considerations
Very low quality of evidence
Very uncertain estimate for a moderate effect
Possible prednison sparing effect
Multidisciplinary disease: a reumatologist should take part in treatment decisions before starting systemic therapy.

Eosinophilic pustular folliculitis

One case series of 6 eosinophilic pustular folliculitis (EPF) patients treated with CsA was part of the Fukamachi et al. study.[50] The dose employed was 100-150 mg/day. Outcome measure was a severity score index. Before and after treatment eruption elements (e.g. pustules), symptoms and number of involved sites were assessed. Five patients had partial remission, defined as a severity score lower than that of pre-treatment. One patient got into complete remission (severity score 0 at the end of treatment). No adverse events were reported.

There is a weak clinical recommendation for the use of cyclosporin in eosinophilic pustular folliculitis.

If more conventional treatment options are contra-indicated or have failed.
Considerations
Very low quality of evidence
Very uncertain estimate for a moderate effect

Hidradenitis suppurativa

One case series, concerning 18 patients, was found in the literature in which patients with hidradenitis were treated with CsA.[51] The dose employed was 2.0–3.5 mg/kg/day. Effectiveness was measured as no change, slight improvement and clinically significant improvement. Seven patients were scored as experiencing ‘‘slight improvement’’ and 9 patients were described as having experienced ‘‘no change’’. Two patients discontinued treatment due to adverse effects. One complained of intolerable gastrointestinal adverse effects, the other of unspecific complaints: ‘‘could not bear it’’.

There is a weak clinical recommendation against the use of cyclosporin in hidradenitis suppurativa.

Considerations
Very low quality of evidence
No clinically significant improvement
Other superior alternative therapies available

Langerhans cell histiocytosis
In one case series 8 patients with Langerhans cell histiocytosis (LCH) were treated with CsA.[52] The dose employed was 12 mg/kg/day. Outcome measures were the extent and pattern of disease manifestations. Two out of 8 had a complete response and 2 had a partial response. Hypertrichosis was seen in almost all patients and 2 patients had hypertension.

There is a weak clinical recommendation for the use of cyclosporin in Langerhans cell histiocytosis.

If conventional treatment options are contra-indicated or have failed.
Considerations
Very low quality of evidence
Very uncertain estimate for a low effect
High off-label dosage with extra safety concerns

Leprosy
Two studies; one RCT, Lambert et al. concerning 13 patients, and one case series, Marlowe et al. concerning 41 patients, were found in the literature in which patients with leprosy type I reaction (LT1) and leprosy type II erythema nodosum leprosum (ENL) were treated with CsA.[53,54] The dose employed was 5-7.5 mg/kg/day. Outcome was measured using sensory testing, voluntary muscle power was graded using the modified Medical Research Council scale and clinical improvement, defined as a reduction in clinical severity score at time of therapy. In Lambert et al. one group was treated with prednisone, the other group was treated with CsA.[53] Patients with chronic ENL receiving cyclosporin had almost twice as many severe flare-ups in the

intervention period (week 0–16) than patients receiving prednisolone only. This was statistically significant (p 0.017). The only significant difference in the amounts of prednisolone taken by participants with new ENL in the two treatment groups was in the overall total amount taken (p 0.028), suggesting that cyclosporin has a steroid sparing effect in the management of ENL. There was 1 serious adverse event; the patient developed osteomyelitis and needed an amputation of the left hallux. In Marlowe et al. found that 35 out of 41 (85%) of the patients improved.[54] In total, 7 adverse events, of which none were serious, were reported in this case series. Mostly gastro-intestinal complaints were reported.

There is a weak clinical recommendation for the use of cyclosporin in leprosy.

Only for a possible prednison sparing effect and if conventional treatment options are contra-indicated or have failed. *Considerations* Very low to low quality of evidence Incosistency between studies and effects Only a few other treatment options available Heterogenic disease

Lichen planus

In total, 3 case series, concerning 33 patients, were found in the literature in which patients with lichen planus were treated with CsA.[55-57] The dose employed was 3-5 mg/kg/day. Outcome measures were the percentage of patients with a complete or partial remission. Mozzanica et al. also used laboratory markers, like antibodies, to monitor disease severity.[55] Higgins et al. got 2 dropouts due to adverse events.[56] All patients showed improvement; 32 complete remission and 1 partial remission. In total, 6 adverse events (asthenia (3), nausea (1), increase of triglycerides (4), rise of creatinine level (1) and nonspecific malaise (1)), of which 1 was serious (myocardial infarction within 24 hours after the first dose of CsA), were reported.

There is a weak clinical recommendation for the use of cyclosporin in lichen planus.

Alternative systemic treatment if conventional and topical treatment options are contra-indicated or have failed.
Considerations
Very low quality of evidence
Very uncertain estimate for a good effect

Lichen sclerosus

One case series, concerning 5 patients, was found in the literature in which patients with lichen sclerosus were treated with CsA.[58] CsA was given in a dose of 3-4 mg/kg/day. Outcome measure was a change in mean total symptom score. Severity of the following symptoms were scored: itch, burning, pain and stinging. Bulbul Baskan et al. showed that the mean total symptom score regressed significantly. Before CsA treatment the mean score was 8.6, after treatment 0.8. Nausea, mild hypertrichosis and mucositis were reported in one patient.

There is a weak clinical recommendation for the use of cyclosporin in lichen sclerosus.

If more conventional and topical treatment options fail.
Considerations
Very low quality of evidence
Very uncertain estimate for a good effect

Livedoid vasculitis

One study, a case series by Leclerc et al. concerning 5 patients, was found in the literature in which patients with livedoid vasculitis were treated with CsA.[59] The CsA dose employed ranged from 1.7-3 mg/kg/day. Outcome was the percentage of patients with clinical improvements. Leclerc et al. found significant improvements in all 5 patients. A few, non-serious adverse events are reported; hirsutism (1), hypertension (3) and increase of creatinine level (1).

There is a weak clinical recommendation for the use of cyclosporin in livedoid vasculitis.

If conventional treatment options are contra-indicated or have failed.
Considerations
Very low quality of evidence
Very uncertain estimate for a moderate effect

Palmoplantar pustulosis

In total, 5 studies, 2 RCT's concerning 98 patients and 3 case series, were found in the literature in which patients with palmoplantar pustulosis were treated with CsA.[60-64] Dosage employed was 1 – 3.75 mg/kg/day. Outcome measures were the percentage of patients with treatment success (defined as: reduction of 50% or more in the number of pustules compared with baseline). Also, other efficacy parameters such as PASI, QoL, erythema, infiltration and scaling, were used to assess efficacy. Both RCT's employed an adequate randomization. Erkko et al. and Reitamo et al. were both very short in explaining the methodology.[60,61] Therefore it is unclear whether there was concealment of allocation, adequate blinding and selected reporting (Addendum 2). Both studies did not report incomplete data but were free of other bias. Erkko et al. had a treatment success of 48% after 1 month in the CsA group. At the same time 19% of the placebo group had a good response. This difference was statistically significant ($p < 0.02$). Erkko et al. also showed a statistically significant reduction in all efficacy parameters (erythema, infiltration and scaling) after 1 month.[60] Reitamo et al. showed treatment success in 89% of the patients in the CsA group after 1 month. In the placebo group this was 21%. This is a statistically significant difference ($p < 0.001$).[61] In the case series from Hayama et al. PASI gradually decreased.[62] Improvement was found in all enrolled patients. QoL scores before and after the administration of cyclosporin differed statistically significant. In the other 2 case series at least 91% of the patients experienced improvement.[63,64] Adverse events reported were: fatigue (3), nausea (3), hypertrichosis (1) and rise of creatinine level (1). There was no need to withdraw patients because of adverse events.

There is a weak clinical recommendation for the use of cyclosporin in palmoplantar pustulosis.

As an alternative systemic treatment after retinoids, topical treatments and local light therapy have failed or are contra-indicated.
Considerations
Very low to low quality of evidence
Uncertain estimate for a good effect

Pemphigus vulgaris

In total, 8 studies, 4 cohorts and 4 case series, were found in the literature in which patients with pemphigus vulgaris were treated with CsA.[65-72] Outcome measures were the percentage of patients with a complete or partial remission. The dosage of CsA was 2.5-8 mg/kg/day. All the cohort studies compared a combination therapy of CsA and prednisone with prednisone alone. Ioannides et al. found complete remission in 5 (29%) patients of the prednisone only group and in 4 (25%) patients of the combined group after 6 months of therapy.
Olszewska et al. found that the time to clinical remission was 7.2 ±13.1 months in the prednisone group and 8.1 ±11.8 months in the combination group. Relapse during therapy was almost similar in both groups.
Chrysomallis et al. 'time to remission' was similar in both groups. In the prednisone alone group this time was 28 days (range 20-40) and in the combined group 25 days (range 18-38). In the case series Mobini et al. was the only study in which CsA was used without concomitant prednisolone. Mobini et al. showed 100% complete remission.[69] In the other case series at least 75% of the patients with combination therapy responded. They also showed a reduction in prednisone dosage. The most reported adverse events were hypertrichosis, hypertension and cutaneous infections.

There is a weak clinical recommendation for the use of cyclosporin in pemphigus vulgaris.

If conventional treatment options are contra-indicated or have failed. When there is a need for a corticosteroid sparing effect and a reduction of the cumulative steroid dose, combination therapy may be considered.
Considerations
Very low quality of evidence
Possible steroid sparing effect, but no significant effect of cyclosporin in addition to oral corticosteroids
Very uncertain estimate for a moderate effect of combination treatment
Very uncertain estimate for a poor effect of cyclosporin mono therapy

Prurigo
In total, 3 case series concerning 43 patients, were found in the literature in which patients with essential prurigo or actinic prurigo were treated with CsA.[73-75] The dose used ranged from 2.5-5.0 mg/kg/day. Outcome measures were the percentage of patients with clinical improvement measured according to an itch intensity scale. At least 92% of the patients showed a response. The skin biopsies taken after treatment showed a significant reduction of all cell types and adhesion molecules (except for IgG and IgM). The mild adverse events were gastric upset (2), cold hands (1), hypertension (2), increase of creatinine level (2), angioedema (1), depression (1), gingival hyperplasia (2), muscle pain (1), hypercholesterinaemia (1), numbness of toes (1) and an increase in blood pressure (1). One serious adverse event was described by Siepmann et al.[75] After 1 month of CSA therapy an embolism of the popliteal artery occurred. This patient stopped CsA treatment.

There is a weak clinical recommendation for the use of cyclosporin in prurigo.

Only after topical therapy and other more conventional options have failed or are contra-indicated.
Considerations
Very low quality of evidence
Very uncertain estimate for a good effect
Great heterogeneity in the included group of patients: unspecified prurigo

Pyoderma gangrenosum

In total 1 RCT and 5 case series, concerning 165 patients, were found in the literature in which patients with pyoderma gangrenosum (PG) were treated with CsA.[76-81] The dose of CsA employed ranged from 4.0 to 10.0 mg/kg/day. Outcome measures were the percentage of patients with a complete or partial remission. In the study by Capella et al. time to reepithelialisation was also used.[77] In the RCT oral prednisone was compared to cyclosporin. Ormerod et al. showed that mean (SD) speed of healing at six weeks was -0.21 (1.00) cm2/day in the cyclosporin group compared with -0.14 (0.42) cm2/day in the prednisolone group; this was not significant.[76] Two serious adverse reactions occurred: a ruptured abdominal aortic aneurysm and acute kidney injury with increased serum creatinine levels. In the case series in Capella et al. found a mean re-epithelialization time of 68 cm2/week.[77] All ulcers healed. In the other case series at least 66% showed healing. Adverse events reported consisted of nausea (3), vomiting (1), headache (1), hypertension (5), tuberculosis (1), weight loss (1), anorexia (1), cellulitis (1), neuropathy (3), nephrotoxicity (2), hypertrichosis (1) and acne (1).

There is a weak clinical recommendation for the use of cyclosporin in pyoderma gangrenosum.

For extensive or progressive pyoderma gangrenosum or after topical therapy has failed in milder pyoderma.
Considerations
Low to moderate quality of evidence
Uncertain estimate for a good effect
No head-to-head comparison with potentially safer topical treatments

Sarcoidosis

One case series, concerning 6 patients, was found in the literature in which patients with dermatological symptoms of sarcoidosis were treated with CsA.[82] The dose employed was 5.0 mg/kg/day. Outcome measure was the percentage of patients with (total) remission of skin lesions. All skins lesions healed. None of the patients relapsed. Adverse events were hypertrichosis in two patients and nausea in one.

There is a weak clinical recommendation for the use of cyclosporin in sarcoidosis.

Alternative systemic treatment for the skin symptoms of sarcoidosis if topical treatment options have failed.
Considerations
Very low quality of evidence
Very uncertain estimate for a good effect

Systemic sclerosis

In total 7 studies, 2 cohort and 5 case series, concerning 78 patients, were found in which patients with systemic sclerosis were treated with CsA.[83-89] Dosage CsA employed was 1 -5 mg/kg/day. Outcome measures were the percentage of patients with a decrease in skin score. In all studies skin score was scored differently, i.e. symptom scoring scale, elastometry, plicometry, or patients opinion and physician's general assessment. In Clements et al., there is a statistically significant better improvement in skin score compared with the placebo group ($p<0.004$).[83] In Filaci et al. there was a better improvement in skin score compared with the Iloprost group.[84] In the case series Basso et al. 7 out of 9 patients show progressive decline in skin involvement.[85] Gisslinger et al. showed statistically significant improvement in skinscore.[86] Ippoliti et al. showed significant improvement of skin elastometry and marked regression of skin fibrosis with a reduction of cutaneous infiltrated in all 5 patients.[87] Morton et al. found marked improvement of symptoms in 5 patients (62.5%).[88] In the study by Worle et al. clinical improvement was found in 4 out of 5 patients.[89] In total, 16 adverse events, of which none were serious, were reported in these case series.

There is a weak clinical recommendation for the use of cyclosporin in systemic sclerosis.

Alternative treatment if conventional treatment options are contra-indicated or have failed.
Considerations
Very low quality of evidence
Very uncertain estimate for a good effect
Severe disease with limited treatment options

Toxic epidermal necrolysis and Steven-Johnson Syndrome

Two cohort studies, Kirchhof et al. and Arevalo et al. and 2 case series, Valeyrie et al. and Singh et al with 121 patients, were included in which patients with toxic epidermal necrolysis (TEN) and Steven-Johnson Syndrome (SJS) were treated with CsA.[90-93] The dosage of CsA used was 3.0 mg/kg/day. Outcome measures were the difference in time to arrest and time to reepithelialisation, the severity-of-illness score for toxic epidermal necrolysis (SCORTEN) and the standardized mortality ratio (SMR). In the cohort study Arevalo et al. found a significant reduction in time to arrest and reepithelialisation time in the CsA group.[90] All patients in the CsA group survived and 3 patients in the control group died due to complications of TEN. In Kirchhof et al. the calculated SMR suggests a survival benefit to cyclosporin use and an increased mortality associated with IVIg use.[91] One patient in the CsA group had a rise of creatinine levels. Valeyrie et al. concluded that cyclosporin had possibly contributed to the survival of all patients, even though the comparison with the expected death rate did not reach significance.[92] In Singh et al. cyclosporin had significantly reduced the time to the arrest of progression of SJS/TEN ($p = 0.043$), the total re-epithelization time ($p = 0.010$) and hospitalization stay ($p = 0.026$). There was no mortality in patients treated with cyclosporin.[93] Adverse events included moderately raised blood pressure, corneal ulceration and decreased renal function. In total two patients died under CsA treatment.

There is a weak clinical recommendation for the use of cyclosporin in toxic epidermal necrolysis and Steven-Johnson syndrome.

If conventional treatment options are contra-indicated or have failed. *Considerations* Very low to low quality of evidence Uncertain estimate for a moderate effect Extra concern for selection bias. Due to the possible mortality asociated with SJS/TEN and the additional risks of cyclosporin therapy, the patient population in the studies may not represent the average TEN population.

Safety

When comparing the adverse events from the Summary of Product Characteristics with the reported AEs and SAEs in the off-label studies, the reported AEs and SAEs are comparable to the known adverse effects or might be explained by the age of the patients and/or the severity of their disease. The studies in this review were mostly

short-term, with limited populations and often incomplete or missing reports of AEs, probably because of the emphasis on effectiveness/efficacy as outcome in the study designs. Therefore, further conclusions regarding (long-term) safety could not be drawn. Nonetheless, the reported AEs and SAEs did not reveal any particular new safety concern.

DISCUSSION

Methodologic quality

As noted before, a literature search can be a useful instrument to screen for promising off-label indications.[94] The main difficulty in assessing the off-label effectiveness/efficacy and safety of cyclosporin was again the poor quality of study design (overall risk of bias, heterogeneity of the data) and a lack in comprehensive reporting of RCTs and other studies in general. Major sources of heterogeneity were the use of different and unvalidated diagnostic criteria, co-medication and previous medication, patient population, treatments itself and the outcomes used. The numerous descriptive case series were prone to confounding, and the retrospective design raised concerns about selection bias. Consequently, only weak recommendations could be made on effectiveness and no recommendations could be made on population, dosage and safety.

Limitations

Case series with less than 5 patients were excluded in this review. The additional value of including numerous case series with less than 5 patients was tested for safety issues. No safety issues were missed. The inclusion of these small case series would lead to an unclear overview.

Clinical research implications

As shown in this and in an earlier review on azathioprine, the available evidence for off-label treatment is scarce in most indications, and promising case series are often not followed by well-designed RCTs.[94] Generating higher quality of evidence for (rare) diseases, for which we do not have (many) effective or safe treatments, ensures future treatment options. Thereby, long-term prospective cohorts/registries, such as the RUBRIC database, should be created to generate long term safety data for off-label drug use within dermatology.[95,96] It is vital that future studies are methodologically sound and reported well and this review highlights the need to standardize the way studies regarding off-label treatment are being conducted.

Besides well validated diagnostic criteria, the use of checklists for generating and reporting data, together with the incorporation of patient reported outcomes in studies is necessary. Most of the reporting checklists and standards, such as CONSORT for randomised trails and STROBE for observational studies, are found on the Enhancing the QUAlity and Transparency Of health Research network website (http://www.equator-network.org/). PROs (patient reported outcomes) are reports provided by patients about their own health, quality of life, or functional status associated with the healthcare or treatment they have received. Patient reported outcome measures are the tools and/or instruments that have been developed to ensure both a valid and reliable measurement of these patients-reported outcomes, such as quality of life measures. Luckily, there is evermore recognition of the importance of involving patients in clinical research and within the wider context of development and evaluation of health care service delivery and quality improvement. The inclusion of PROs through the use of PROMs with clinical outcomes in research and clinical practice provides a more complete understanding of the impact of a therapy on the patient.[97] Patient reported outcomes should be part of a core outcome set (COS). A COS is an agreed minimum set of outcomes or outcome measures. The existence or use of a COS does not imply that outcomes in a particular trial should be restricted to those in the relevant core outcome set. Rather, there is an expectation that the core outcomes will be collected and reported, making it easier for the results of trials to be compared, contrasted and combined as appropriate; while researchers continue to explore other outcomes as well (http://miror-ejd.eu/individual-research-projects/). Unfortunately, there were no COS available for the included diseases in this review. Dermatological COS are available for vitiligo and eczema and a couple of COS are being developed at this moment.[98-101] It is desirable that COS are formulated for more dermatologic diseases.

Implications for clinical practice

Results from this review can be used to update or formulate clinical practice guidelines for off-label cyclosporin prescription and thus ensure evidence-based healthcare decision making. Future studies should be performed to confirm the role of cyclosporin in diseases that currently have a weak recommendation because of (very) low quality of evidence but in which cyclosporin has shown promising results such as; chronic contact dermatitis, Langerhans cell histiocytosis, lichen planus, livedoid vasculitis, pemphigus vulgaris, pyoderma gangrenosum, sarcoidosis, systemic sclerosis and toxic epidermal necrolysis.

CONCLUSION

Evidence for off-label cyclosporin use exists in 22 different dermatological conditions. Overall, the quality of the included articles was low; downgrading due to sparse data or indirectness was common. Therefore, only weak recommendations could be given. Prospective long-term comparative registries and prospective studies should add useful evidence and provide legal support for off-label drug use in dermatology. Harmonising outcomes by defining core outcome sets is necessary.

ADDENDUM 1

Summary of Findings tables

Cyclosporin in observational studies for alopecia areata					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Hair growth assessed with: significant hair growth (opinion of clinical assessor) after 2.5–6 mg/kg/day cyclosporin for 4.41 months (mean) [a]	10 out of 25 (40%) patients showed significant hair growth.		25 (1 observational study)	⊕◯◯◯ VERY LOW [b]	Only alopecia areata patient > 50% alopecia were included. Only male patients included. 3 (12%) patients discontinued treatment due to adverse events.
Good response assessed with: opinion of clinical assessor after 5-7.5 mg/kg/day cyclosporin for 6 months [c]	5 out of 6 patients had a good response (83%).		6 (1 observational study)	⊕◯◯◯ VERY LOW [d,e]	One patient withdrew after 2 weeks.
Complete remission assessed with: complete regrowth of hair after 5 mg/kg/day cyclosporin for 6-12 months [f]	2 out of 15 patients (13%) had complete remission.		15 (1 observational study)	⊕◯◯◯ VERY LOW [d]	One patient withdrew after 1 month (severe hypertension). 5 patients had > 70% regrowth.
Complete remission assessed with: complete regrowth of hair after 200 mg cyclosporin for 2-3.5 months [g]	33 out of 46 patients had complete remission (77%).		46 (1 observational study)	⊕◯◯◯ VERY LOW [h]	Prednisolone was used as concomitant medication in all patients. 24% of responders relapsed during follow-up. 3 patients withdrew.

Adequate response assessed with: > 70% regrowth after 2.5 mg/kg cyclosporin for 5-8 months [i]	6 out of 18 patients had an adequate response (33%).		18 (1 observational study)	⊕◯◯◯ VERY LOW [d,h]	Also alopecia universalis patients. Prednisolone was used as concomitant medication in all patients.
Number of responders to therapy assessed with: > 75% regrowth after 5 mg/kg/day cyclosporin for 6 months [j]	2 out of 8 patients were responders (25%).		8 (1 observational study)	⊕◯◯◯ VERY LOW [d,h]	Both responders relapsed after treatment. 3 patients withdrew (oedema, hypertension, abnormal lab tests). Prednisolone was used as concomitant medication in all patients.
Favourable hair growth assessed with: opinion of clinician after 2.5 mg/kg/day cyclosporin for 5 months [k]	6 out of 6 patients had favourable hair growth (100%).		6 (1 observational study)	⊕◯◯◯ VERY LOW [b,e,h]	No relapses after 6 months follow-up. Prednisolone was used as concomitant medication in all patients.

Explanations

a. Acikgoz et al. 2014[14], b. Retrospective series, unclear outcome measure, varying duration of treatment, unclear assessment, sparse data, c. Constantopoulos et al. 1996[15], d. Retrospective small series with sparse data, e. Indirect and unspecified outcome measure, f. Ferrando et al. 1999[16], g. Kim et al. 2008[17], h. Uncertainty of effect due to natural course and concomitant medication, i. Shaheedi et al. 2008[18], j. Shapiro et al. 1997[19], k. Teshima et al. 1992[20]

Cyclosporin in observational studies for Behcet					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Oral symptoms free interval assessed with: weeks of no reported oral ulceration during 5 mg/kg/day cyclosporin for > 6 months [a]	Average oral symptom free interval lengthened from 2.7 to 4.9 weeks.		24 (1 observational study)	⨁◯◯◯ VERY LOW [b]	
Remission of oral ulceration assessed with: reported buccal ulceration during 2-7 mg/kg/day cyclosporin for 30-38 months [c]	11 out of 14 had complete remission of buccal ulcerations at 5 mg/kg/day and 7 out of 14 at 2 mg/kg/day.		14 (1 observational study)	⨁◯◯◯ VERY LOW [b,d]	In case of poor response with cyclosporin 7 mg/kg/day prednisolone was allowed. Starting dose was 5 mg/kg/day. Good response: tapering to 2 mg/kg/day.
Decrease in oral ulceration assessed with: number and size (mm) of reported oral ulcers during 5 mg/kg/day cyclosporin for > 6 months [a]	Number of ulcers diminished from 4.6 to an average of 1.7. The size of the oral ulcers decreased from an average of 5.7 mm to 3.6 mm.		24 (1 observational study)	⨁◯◯◯ VERY LOW [b]	
Remission of genital ulceration assessed with: reported genital ulcers during 5 mg/kg/day cyclosporin for > 6 months [a]	No genital ulcerations appeared in 17 out of 21 (80%) patients during treatment.		24 (1 observational study)	⨁◯◯◯ VERY LOW [b]	

Remission of acneiform lesions assessed with: reported acneiform lesions during 5 mg/kg/day cyclosporin for > 6 months [a]	Acneiform lesions did not occur in 17 out of 18 patients during treatment.		24 (1 observational study)	⨁◯◯◯ VERY LOW [b]	
Remission of erythema nodosum-like lesions assessed with: reported during 5 mg/kg/day cyclosporin for > 6 months [a]	17 out of 20 patients did not have any erythema nodosum-like lesions during treatment.		24 (1 observational study)	⨁◯◯◯ VERY LOW [b]	
Remission of chronic cutaneous vasculitis lesions assessed with: reported vasculitis lesions during 2-7 mg/kg/day cyclosporin for 30-38 months [c]	4 out of 5 (80%) had complete remission at 5 mg/kg/day and 2 out of 5 (40%) at 2 mg/kg/day.		5 (1 observational study)	⨁◯◯◯ VERY LOW [b,d]	In case of poor response with cyclosporin 7 mg/kg/day prednisolone was allowed. Starting dose was 5 mg/kg/day. Good response: tapering to 2 mg/kg/day.

Explanations

a. Avci et al. 1997[21], b. Case series with sparse data, c. Diaz-Llopis et al. 1990[22], d. Possible bias due to concomitant medication

Cyclosporin in observational studies for chronic actinic dermatitis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Remission of symptoms assessed with: good response (not otherwise specified) after 1.5-4 mg/kg/day cyclosporin for 4-26 weeks (mean 6 weeks) [a]	3 out of 6 patients (50%).		6 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	Initial dosage was 2.5 mg/kg/day. One patient got lost to follow-up. 2 patients relapsed during follow-up.

Explanations

a. Granlund et al. 1998[23], b. Very small series with sparse data, c. Undefined outcome measure

Cyclosporin compared to placebo or cyclosporin in observational studies for chronic contact dermatitis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Total reduction in severity score assessed with: score (0-3) in itch, burning, erythema, oedema, vesicles, crusts, scaling and infiltration after 5 mg/kg/day cyclosporin and after 4 weeks 3 mg/kg/day for 3 months [a]	100% reduction of score in all 15 patients after 4 weeks.		15 (1 observational study)	⨁⨁◯◯ LOW [b]	Two relapses after 6 months follow-up.
Clinical improvement assessed with: % of patients after 2 weeks 5 mg/kg/day cyclosporin [c]	100% (6 out of 6) showed clinical improvement. 4 out of 6 healed completely.		6 (1 observational study)	⨁◯◯◯ VERY LOW [d,e]	No follow-up mentioned.

Explanations

a. Flori et al. 1993[25], b. Unclear assessment of outcome measure, c. Higgins et al. 1991[24], d. Small series with sparse data, e. Unclear outcome measure

Cyclosporin 3mg/kg/day with plain cream compared to betamethasonedipropionate cream with placebo capsules for chronic dermatitis of the hand[28]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with betamethasonedipropionate cream with placebo capsules	**Risk with cyclosporin 3mg/kg/day with plain cream**				
% of patients with treatment success assessed with: decrease in disease activity score to ≤50% of patient's own baseline [a]	286 per 1.000	**499 per 1.000** (.. to ..) [b]	**RR 1.7482** (-- to --)	41 (1 RCT)	⨁⨁◯◯ LOW [b,c]	During follow-up 50% of both groups relapsed in the first 2 weeks. 7 patients withdrew (4 cyclosporin group and 3 betamethasone group) 6 weeks treatment, non-responders did crossover.

Explanations

a. Disease activity score = disease activity from 0 (none) to 3 (severe) plus extent of the disease, use of emollients, itch and sleep disturbances and overall assessment of efficacy (1/very good to 5/no effect), b. No confidence intervals and unclear intention to treat analysis, c. Unclear concealment of allocation and researcher blindation

Cyclosporin compared to placebo or cyclosporin in observational studies for chronic dermatitis of the hand					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Effectiveness according to physician's global assessment (PGA) assessed with: PGA score after 3 months cyclosporin (either start ≤3.5 mg/kg/day and gradually increased to max. 5 mg/kg or start with >3.5– 5 mg/kg and slowly tapered) [a,b]	63% of the patients improved at least 50% in their hand eczema (PGA1)		121 (1 observational study)	⨁◯◯◯ VERY LOW [c,d]	During treatment with cyclosporin, patients were allowed to use topical emollients and/or topical corticosteroids. 65% had recurrent vesicular hand eczema and 42% atopic eczema.
Duration of drug survival assessed with: number of days which an individual continued the first treatment episode [a]	The overall drug survival rates after 6 months, 1, 2 and 3 years were 61.7%, 45.2%, 18.6% and 13.9% respectively. The median survival duration for cyclosporin in hand eczema was 0.86 year (10.3 months).		121 (1 observational study)	⨁◯◯◯ VERY LOW [c,e]	48% discontinued due to adverse events (hypertension (21%), gastrointestinal complaints (8%), headache (6%), fatigue (6%) and a serum creatinine increase of >30% (6%).
Effectiveness assessed with: percentage of patients with a complete, good or moderate response after 2-16 weeks cyclosporin 1.25-2.5 mg/kg/day [f]	1 out of 7 patients had a complete response. 3 out of 7 had a good response, of which 2 relapsed after 3 weeks and 8 months. Moderate response was seen in 2 out of 7 patients.		7 (1 observational study)	⨁◯◯◯ VERY LOW [d,g]	3 patients a hyperkeratotic type and 4 a vesicular type. Duration of treatment varied from 2 to 16 weeks. 1 patient lost to follow-up.

Patients in remission assessed with: Percentage of patient in remission at 1-year follow-up after 6 weeks 3 mg/kg/day cyclosporin [h]	21 out of 27 (78%) patients in remission after 1 year of follow-up.		27 (1 observational study)	⊕◯◯◯ VERY LOW [d]	Same patients as in Granlund et al. 1996 RCT

Explanations

a. Christoffers et al. 2016[26], b. PGA score: PGA 1: good effect; more than 50% improvement, PGA 2: moderate effect; less than 50% improvement, PGA 3: failure of treatment; no improvement or worse., c. Retrospective series with sparse data about variable dosing, d. Imprecise outcome measure, e. Indirect outcome measure, f. Reitamo et al. 1994[27], g. Small case series with sparse data and variable dosing and follow-up, h. Granlund et al. 1998[23]

Cyclosporin 5 mg/kg/day compared to prednisone 20 mg/day for chronic urticaria[34]					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Complete response assessed with: % of patients reaching complete reduction in symptoms after 2 months treatment	100% (10/10) complete response in both groups		20 (1 observational study)	⊕◯◯◯ VERY LOW [a]	No concomitant drugs were used.

Explanations

a. No adequate reporting of follow-up, time to response and duration of remission

Cyclosporin 4 mg/kg/day and cetirizine 20 mg/day compared to cetirizine 20 mg/day for chronic urticaria[29]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with cetirizine 20 mg/day	**Risk with cyclosporin 4 mg/kg/day and cetirizine 20 mg/day**				
Clinical response assessed with: Mean reduction in UAS Scale from: 0 to 42 follow up: 4 weeks[a]	The mean clinical response was **2.3**	The mean clinical response in the intervention group was 10,4 more (4,3 more to 16,5 more)	-	30 (1 RCT)	⊕⊕⊕◯ MODERATE [b]	1 withdrawal in cyclosporin group not included in further analysis
Clinical response assessed with: Reduction in VAS score by patients follow up: 4 weeks[c]	The mean clinical response was **0.7**	The mean clinical response in the intervention group was 2,3 more (0,8 more to 3,8 more)	-	30 (1 RCT)	⊕⊕⊕◯ MODERATE [b]	1 withdrawal in cyclosporin group not included in further analysis
Clinical response assessed with: NNT to achieve <25% baseline UAS score follow up: 4 weeks[d]	NNT with cyclosporin was 2.4 (95% CI 1.6-5.0)			30 (1 RCT)	⊕⊕⊕◯ MODERATE [b]	
Clinical response assessed with: number of patients with reduction < 25% of baseline UAS follow up: 4 weeks	11 patients responded (65%)			17 (1 observational study)	⊕◯◯◯ VERY LOW [e]	Open label cyclosporin for 4 weeks for non-responders (from cyclosporin group and placebo group)

Explanations

a. UAS = Urticaria Activity Score. A weekly aggregate of weal count (3 categories) and severity of itch (3 categories)., b. Small sample size and unclear blinding, c. VAS = Visual Analogue Score, d. NNT = Number Needed to Treat, e. Small sample size und no blinding.

Cyclosporin 1-5 mg/kg for chronic urticaria						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with cyclosporin**				
Complete response assessed with: % of patients with complete remission of symptoms after 1-month cyclosporin 2.5 mg/kg [a]	70% (19 out 27 patients) had a complete response			27 (1 observational study)	⨁◯◯◯ VERY LOW [b]	
Complete response assessed with: % of patients with complete remission of symptoms after 1-week cyclosporin 3mg/kg, 1-week 2mg/kg and 1 week 1 mg/kg [c]	The mean complete response was **0** %	The mean complete response in the intervention group was 68 % higher (0 to 0)	-	35 (1 observational study)	⨁◯◯◯ VERY LOW [d]	6 drop-outs in cyclosporin group; not analysed in % complete response. Unclear whether and how much concomitant drugs were used.
Complete or partial response assessed with: % of patients with complete or partial remission of symptoms after 1 month cyclosporin 2.5-3.5 mg/kg [e]	75% (9 out of 12 patients) had a partial or complete response			12 (1 observational study)	⨁◯◯◯ VERY LOW [f,g]	All subjects used antihistaminic drugs; unclear how much. 3 patients used prednisone, unclear how much.
Complete response assessed with: % of patients with complete remission of symptoms after 1 month cyclosporin 3 mg/kg [h]	100% (7 out of 7) had a complete response			7 (1 observational study)	⨁◯◯◯ VERY LOW [i]	No information about comedication or follow-up. Relapses (4/7) were successfully treated with cyclosporin.

Cyclosporin 1-5 mg/kg for chronic urticaria						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with cyclosporin**				
Complete response assessed with: % of patients with complete remission of symptoms after 3 weeks of tapering doses of cyclosporin [j,k]	13% (2 out of 15) had a complete response			15 (1 observational study)	⨁◯◯◯ VERY LOW [f]	After 1 week 100% complete response but relapses with tapering.
Clinical response assessed with: % reduction after 5 months of a total score consisting of a 4-point scale for pruritus severity, no. of wheals, diameter of wheals and pruritus duration [l]	88% reduction in total score			30 (1 observational study)	⨁◯◯◯ VERY LOW [m,n]	Varying dosage (1.5 mg/kg-2.5 mg/kg). 13.3% drop-out due to high blood pressure.
Clinical response assessed with: % mean total score reduction (pruritus severity, no. of wheals and diameter of wheals) after 6 months 1, 2, or 3 mg/kg cyclosporin [o]	Mean total score reduction of 63% (1 mg/kg), 76% (2 mg/kg) and 86% (3 mg/kg).			110 (1 observational study)	⨁◯◯◯ VERY LOW [n,p]	Oral steroids and antihistamines as concomitant medication.

Cyclosporin 1-5 mg/kg for chronic urticaria						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with cyclosporin**				
Complete response assessed with: % complete remission after 6 months 1, 2, or 3 mg/kg cyclosporin [o]	Complete response 39% (25 patients with 1 mg/kg), 28% (32 patients with 2 mg/kg) and 45% (53 patients with 3 mg/kg)			110 (1 observational study)	⨁◯◯◯ VERY LOW [p]	Oral steroids and antihistamines as concomitant medication.
Complete response assessed with: % complete remission of symptoms after 2-24 months <3 mg or 3 mg/kg cyclosporin [q]	Complete remission 100% (7 out of 7 patients)			7 (1 observational study)	⨁◯◯◯ VERY LOW [f,i]	Varying dosage and treatment duration. Oral steroids and antihistamines as concomitant medication. Study in children.
Complete response assessed with: Complete remission oedematous erythema and pruritus after 11.3 months (mean) 3 mg/kg/day cyclosporin [r]	Complete remission 100% (15 out of 15)			15 (1 observational study)	⨁◯◯◯ VERY LOW [f,i]	

Cyclosporin 1-5 mg/kg for chronic urticaria						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with cyclosporin**				
Complete response assessed with: Complete remission after 3 months 3 mg/kg cyclosporin of a total score consisting of a 4-point scale for pruritus severity, no. of wheals, diameter of wheals and pruritus duration [s]	Complete remission 25% (30 out of 120)			120 (1 observational study)	⨁◯◯◯ VERY LOW [f]	Concomitant oral steroids 10-240 mg/day. Partial remission (score 1-2) in 43%. Large drop-out due to severe adverse events (16.5%).
Complete response assessed with: % complete absence of urticaria with mean dosage of 2.1 mg/kg cyclosporin (unknown duration) [t]	39% complete response (25 out of 64 patients)			64 (1 observational study)	⨁◯◯◯ VERY LOW [i]	Unclear duration of treatment.
Complete response assessed with: Complete remission of urticaria after 3 mg/kg/day cyclosporin and gradual tapering (median duration 5.5 months) [u]	100% (16 out of 16 patients)			16 (1 observational study)	⨁◯◯◯ VERY LOW [f,p]	Study in children. 31% relapse after discontinuation cyclosporin.

Cyclosporin 1-5 mg/kg for chronic urticaria						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with cyclosporin**				
Clinical response assessed with: Patient reported alleviation of symptoms with 2.5-5 mg/kg/day cyclosporin for a mean of 14 weeks [v]	18% (2 out of 11)			11 (1 observational study)	⨁◯◯◯ VERY LOW [i,n,p]	Chronic solar urticaria patients. 1 drop-out due to adverse event.
Complete response assessed with: Absence of pruritus and wheals with min. 3 months 3-5mg/kg/day cyclosporin[w]	59% (17 out of 29 patients)			29 (1 observational study)	⨁◯◯◯ VERY LOW [f]	6 patients had a relapse after successful treatment. Antihistamines were allowed in excellent response category.
Clinical severity score assessed with: after 5 mg/kg/day cyclosporin for 8 weeks and then 4 mg/kg/day for 8 weeks [x]	Significant drop in clinical severity score (p = 0.002).			40 (1 observational study)	⨁◯◯◯ VERY LOW [y]	Originally designed as RCT, however randomization abandoned after 2 weeks and all patients received cyclosporin monotherapy. Data only available for open observational part. Antihistamines allowed.

Cyclosporin 1-5 mg/kg for chronic urticaria						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with cyclosporin**				
Negative ASST in originally positive ASST patients assessed with: after 5 mg/kg/day cyclosporin for 8 weeks and then 4 mg/kg/day for 8 weeks [x,z]	The ASST was negative in 13 patients and positive in 3 of 16 patients, with total remission of symptoms.			40 (1 observational study)	⨁◯◯◯ VERY LOW [aa,y]	Originally designed as RCT, however randomization abandoned after 2 weeks and all patients received cyclosporin monotherapy. Data only available for open observational part. Antihistamines allowed.

Explanations

a. Serhat et al. 2008[33], b. Small sample size and an indirect non-comparable study design, c. Toubi et al. 1997[35], d. Small sample size, no intention to treat analysis, large drop-out, possible bias due to concomitant drugs, e. Barlow et al. 1993[37], f. Small sample size, possible bias due to concomitant medication, g. Unclear definition of outcome measure, h. Daher et al. 2010[38], i. Very small sample size and sparse data, j. Week 1: 5 mg/kg/day, week 2: 3.5 mg/kg/day and week 3: 2.5 mg/kg/day, k. Ilter et al. 1999[36], l. Boubouka et al. 2011[39], m. Large drop-out and varying dosage, n. Indirect outcome measure, o. Di Leo et al. 2011[31], p. Possible bias due to non-specified concomitant medication, q. Doshi et al. 2009[41], r. Ohtsuka et al. 2010[40], s. Kessel et al. 2010[42], t. Lim et al. 2013[43], u. Neverman et al. 2014[45], v. Harubielle et al. 2015[44], w. Asero 2015[46], x. Di Gioacchino 2003[32], y. Sparse data; changing of original study design, z. ASST = autologous serum skin test, aa. ASST is indirect outcome measure

Cyclosporin 3-5mg/kg and cetirizine 10 mg/day compared to cetirizine 10 mg/day for chronic urticaria[30]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with cetirizine 10 mg/day	**Risk with cyclosporin 3-5mg/kg and cetirizine 10 mg/day**				
Improvement in severity assessed with: Mean improvement in severity score as reported by Breneman et al. (%) follow up: 8 weeks	The mean improvement in severity was **23.3** %	The mean improvement in severity in the intervention group was 39,2 % higher (0 to 0)	-	66 (1 RCT)	⨁◯◯◯ VERY LOW [b,c,d]	Data from the 16 weeks cyclosporin arm. Drop out 48.6% in cetirizine group, 22.6% in the 16 week cyclosporin group. Primary study outcome.
Improvement in severity assessed with: Mean improvement in severity score as reported by Breneman et al. (%) follow up: 8 weeks	The mean improvement in severity was **23.3** %	The mean improvement in severity in the intervention group was 38,8 % higher (0 to 0)	-	68 (1 RCT)	⨁◯◯◯ VERY LOW [b,c,d]	Data from the 8 weeks cyclosporin arm. Drop out 48.6% in cetirizine group, 36.4% in the 8 week cyclosporin group. Primary study outcome.
Improvement in severity assessed with: Mean improvement in severity score as reported by Breneman et al. (%) follow up: 16 weeks	The mean improvement in severity was **25.0** %	The mean improvement in severity in the intervention group was 27,9 % higher (0 to 0)	-	66 (1 RCT)	⨁◯◯◯ VERY LOW [b,c,d]	Data from the 16 weeks cyclosporin arm. Secondary study outcome.

Cyclosporin 3-5mg/kg and cetirizine 10 mg/day compared to cetirizine 10 mg/day for chronic urticaria[30]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with cetirizine 10 mg/day	**Risk with cyclosporin 3-5mg/kg and cetirizine 10 mg/day**				
Improvement in severity assessed with: Mean improvement in severity score as reported by Breneman et al. (%) follow up: 16 weeks	The mean improvement in severity was **25.0** %	The mean improvement in severity in the intervention group was 20,6 % higher (0 to 0)	-	68 (1 RCT)	⨁◯◯◯ VERY LOW [b,c,d]	Data from the 8 weeks cyclosporin arm. Secondary study outcome.
Improvement in severity assessed with: Mean improvement in severity score as reported by Breneman et al. (%) follow up: 24 weeks	The mean improvement in severity was **30.2** %	The mean improvement in severity in the intervention group was 11,5 % higher (0 to 0)	-	66 (1 RCT)	⨁◯◯◯ VERY LOW [b,c,d]	Data from the 16 weeks cyclosporin arm. Secondary study outcome.
Improvement in severity assessed with: Mean improvement in severity score as reported by Breneman et al. (%) follow up: 24 weeks	The mean improvement in severity was **30.2** %	The mean improvement in severity in the intervention group was 16,7 % higher (0 to 0)	-	68 (1 RCT)	⨁◯◯◯ VERY LOW [b,c,d]	Data from the 8 weeks cyclosporin arm. Secondary study outcome.
Improvement of quality of life assessed with: Mean improvement in DLQI (%) Scale from: 0 to 30 follow up: 24 weeks [e]	The mean improvement of quality of life was **34.6** %	The mean improvement of quality of life in the intervention group was 27,6 % higher (0 to 0)	-	66 (1 RCT)	⨁⨁◯◯ LOW [b,d]	Data from the 16 weeks cyclosporin arm. Secondary study outcome.

Cyclosporin 3-5mg/kg and cetirizine 10 mg/day compared to cetirizine 10 mg/day for chronic urticaria[30]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with cetirizine 10 mg/day	**Risk with cyclosporin 3-5mg/kg and cetirizine 10 mg/day**				
Improvement of quality of life assessed with: Mean improvement in DLQI (%) Scale from: 0 to 30 follow up: 24 weeks [e]	The mean improvement of quality of life was **34.6** %	The mean improvement of quality of life in the intervention group was 29 % higher (0 to 0)	-	68 (1 RCT)	⨁⨁◯◯ LOW [b,d]	Data from the 8 weeks cyclosporin arm. Secondary study outcome.
Rescue therapy assessed with: Number of people requiring rescue medication	19.4% needed rescue therapy in the 16 week cyclosporin group versus 42.4% in the 8 week cyclosporin group and 31.4% in the cetirizine only group.			99 (1 RCT)	⨁⨁◯◯ LOW [b,d]	During the 24 weeks post randomisation period
Relief of symptoms assessed with: Odds of SGA of relief of symptoms follow up: 24 weeks [f]	1.7 odds of relief in the 16 week cyclosporin group versus 0.9 in the 8 week cyclosporin group and 1.7 in the cetirizine only group.			99 (1 RCT)	⨁◯◯◯ VERY LOW [b,d,g]	
Serious adverse events assessed with: Events that needed hospitalization follow up: 24 weeks	1 patient had a serious adverse event in the 16 weeks cyclosporin group versus 1 in the 8 weeks cyclosporin group and 0 in the cetirizine only group.			99 (1 RCT)	⨁⨁◯◯ LOW [b,d]	Events were not drug related according to the authors of the study

Explanations

a. No confidence interval reported, b. Large drop-out and unclear randomization, c. Unclear if score is a real proxy for disease severity, d. Small sample size and no confidence interval reported, e. DLQI = Dermatology Life Quality Index, f. SGA = subject's global assessment, g. Unclear how odds were calculated

Cyclosporin compared to placebo or cyclosporin in observational studies for cutaneous T-cell lymphoma					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Complete response assessed with: remission of disease activity after 15 mg/kg/day cyclosporin for 1 week up to 13 months [a]	2 out of 11 (18%) had a complete response.		11 (1 observational study)	⊕◯◯◯ VERY LOW [b,c]	Unknown if concomitant therapy was used. Follow-up time and time to response are not mentioned. Renal toxicity in 6 patients.

Explanations
a. Cooper et al. 1993[47], b. Small series with sparse data, c. Unclear outcome measure and assessment

Cyclosporin compared to placebo or in observational studies for dermatomyositis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Effectiveness assessed with: percentage of patients with clinical improvement (no definition) after 3-12 months 5 mg/kg/day cyclosporin [a]	Remarkable improvement in 6 out of 10 (60%) patients.		10 (1 observational study)	⊕◯◯◯ VERY LOW [b,c]	Prednisone was used as concomitant medication. Unknown whether the patients had skin symptoms or not.

Explanations
a. Danko et al. 1991[49], b. Small series with sparse data, c. Unclear outcome measure, dosing and concomitant medication

Cyclosporin 5 mg/kg/day compared to prednisone (and azathioprine) for dermatomyositis						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with prednisone (and azathioprine)	**Risk with cyclosporin 5 mg/kg/day**				
Time to complete remission assessed with: in days with either prednisone 20 mg and after 2 months azathioprine (100 mg/day or 2 mg/kg) or cyclosporin 5 mg/kg/day [a]	The mean time to complete remission was **23** days	The mean time to complete remission in the intervention group was 14,4 days lower	-	55 (1 observational study)	⨁◯◯◯ VERY LOW [b]	Large standard deviations.
Therapy failure assessed with: Number of patients without improvement with cyclosporin 5 mg/kg/day or prednisone 20 mg/day and after 2 months azathioprine [a]	4 therapy failures (9%) in prednisone and azathioprine group. 1 failure (10%) in cyclosporin group.			55 (1 observational study)	⨁◯◯◯ VERY LOW [b]	
Rash clearance - not reported	-		-	-	-	

Explanations a. Grau et al. 1994[48], b. Historical control group and variable dosage

Cyclosporin compared to placebo or cyclosporin in observational studies for eosinophilic pustular folliculitis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Complete reduction in severity score index assessed with: after 2-12 weeks 100-150 mg/day cyclosporin [a]	1 out of 6 achieved complete reduction (18%).		6 (1 observational study)	⨁◯◯◯ VERY LOW [b]	Only one patient used cyclosporin as monotherapy.

Explanations a. Fukamachi et al. 2009[50], b. Small series with sparse data, unclear outcome assessment and variable dosing and co-medication

Cyclosporin compared to placebo or cyclosporin in observational studies for hidradenitis suppurativa					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Effectiveness graded as ‘‘worsening’’, ‘‘no change’’, ‘‘slight improvement’’ or ‘‘clinically significant improvement’’ assessed with: patients and physicians subjective score after 2.0–3.5 mg/kg/day cyclosporin (varying duration) [a]	2 out 18 patients had clinically significant improvement. 7 out of 18 patients had slight improvement.		18 (1 observational study)	⨁◯◯◯ VERY LOW [b]	Retrospective series. One significant improvement only occurred after co-medication with clavulanic acid and amoxicillin.

Explanations a. Anderson et al. 2016[51], b. Small retrospective series with unclear outcome measures and variable dosage

Cyclosporin compared to placebo or cyclosporin in observational studies for Langerhans cell histiocytosis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Complete response assessed with: after 2 to 12 months cyclosporin 12 mg/kg/day [a]	2 out of 8 patients (25%) had a complete response.		8 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	4 patients were still on steroid therapy during cyclosporin treatment. The dosage was tapered whenever possible.

Explanations a. Arico et al. 1995[52], b. Small series with sparse data, c. Unclear outcome assessment

Cyclosporin compared to placebo or cyclosporin in observational studies for leprosy type I reactions					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Reduction in clinical severity score assessed with: % of patients after 3 months cyclosporin 5.0 mg/kg/day [a]	35 out of 41 patients (85%) showed improvement. 57% kept the improvement after 12 weeks follow-up.		41 (1 observational study)	⨁◯◯◯ VERY LOW [b]	2 patients lost to follow-up. 13 patients relapsed. 15 patients got a higher dosage of 7.5 mg/kg/day.

Explanations

a. Marlowe et al. 2007[54], b. Unclear outcome measure and assessment

Cyclosporin 7.5 mg/kg/day (tapered after 12 weeks) and prednisone 40 mg tapered in 4 weeks compared to prednisone 60 mg/day (tapered with 5 mg/week) for acute leprosy type II reactions[53]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with prednisone 60 mg/day (tapered with 5 mg/week)	**Risk with cyclosporin 7.5 mg/kg/day (tapered after 12 weeks) and prednisone 40 mg tapered in 4 weeks**				
Number of ENL recurrence episodes per patient after 16 weeks assessed with: occurrence of ENL requiring the institution or change of treatment [a]	The mean number of ENL recurrence episodes per patient after 16 weeks was **2.4**	The mean number of ENL recurrence episodes per patient after 16 weeks in the intervention group was 1,11 fewer (0 to 0)	-	13 (1 RCT)	⨁⨁◯◯ LOW [b,c]	

Time to ENL recurrence after initial control assessed with: weeks to event	The mean time to ENL recurrence after initial control was **23** weeks	The mean time to ENL recurrence after initial control in the intervention group was 13,8 weeks fewer (0 to 0)	-	13 (1 RCT)	⊕◯◯◯ VERY LOW [b,c]	
Severity of ENL at recurrence assessed with: rated by two physicians as none, mild, moderate or severe	Both had 2 severe ENL recurrences after 16 weeks.			13 (1 RCT)	⊕◯◯◯ VERY LOW [b,c]	No statistically significant differences.
Amount of additional prednisolone required by patients after 32 weeks assessed with: mg	The mean amount of additional prednisolone required by patients after 32 weeks was **2800** mg	The mean amount of additional prednisolone required by patients after 32 weeks in the intervention group was 700 mg lower (0 to 0)	-	13 (1 RCT)	⊕◯◯◯ VERY LOW [b,c]	Data was read from graph.
Frequency of adverse events assessed with: any event leading to admission or prolonged admission, study unblinding or death	45 adverse events in prednisone only group (33 patients) and 8 in intervention group (17 patients).			50 (1 RCT)	⊕⊕◯◯ LOW [b]	Data from all participant, both acute and chronic ENL. [a]
Difference in score in Quality of Life assessment between start and end	No statistically significant difference in score changes for any of the SF-36 scales between study arms, except in the mental health score the cyclosporin group (+23 vs. -3) had a better improvement.			13 (1 RCT)	⊕◯◯◯ VERY LOW [b,c]	Data was read from graph.

Explanations

a. ENL = erythema nodosum leprosum, b. Small RCT with inadequate blinding, c. No confidence intervals given; no statistical significance

Cyclosporin 7.5 mg/kg/day (tapered after 12 weeks) and prednisone 40 mg tapered in 4 weeks compared to prednisone 60 mg/day (tapered with 5 mg/week) for chronic leprosy type II reactions[53]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with prednisone 60 mg/day (tapered with 5 mg/week)	**Risk with cyclosporin 7.5 mg/kg/day (tapered after 12 weeks) and prednisone 40 mg tapered in 4 weeks**				
Number of ENL recurrence episodes per patient after 16 weeks assessed with: occurrence of ENL requiring the institution or change of treatment [a]	The mean number of ENL recurrence episodes per patient after 16 weeks was **2.0**	The mean number of ENL recurrence episodes per patient after 16 weeks in the intervention group was 0,3 more (0 to 0)	-	13 (1 RCT)	⊕⊕◯◯ LOW [b,c]	
Time to ENL recurrence after initial control assessed with: weeks to event	The mean time to ENL recurrence after initial control was **11.25** weeks	The mean time to ENL recurrence after initial control in the intervention group was 4,15 weeks fewer (0 to 0)	-	13 (1 RCT)	⊕◯◯◯ VERY LOW [b,c]	
Severity of ENL at recurrence assessed with: rated by two physicians as none, mild, moderate or severe	14 severe flare-ups in cyclosporin group after week 16 and 6 in patients receiving prednisolone only.			13 (1 RCT)	⊕◯◯◯ VERY LOW [b,c]	Statistically significant difference. Data was read from graph.

Amount of additional prednisolone required by patients after 32 weeks assessed with: mg	The mean amount of additional prednisolone required by patients after 32 weeks was **2200** mg	The mean amount of additional prednisolone required by patients after 32 weeks in the intervention group was 1.900 mg higher (0 to 0)	-	13 (1 RCT)	⊕◯◯◯ VERY LOW [b,c]	Data was read from graph.
Frequency of adverse events assessed with: any event leading to admission or prolonged admission, study un-blinding or death	45 adverse events in prednisone only group (33 patients) and 8 in intervention group (17 patients).			50 (1 RCT)	⊕⊕◯◯ LOW [b]	Data from all participant, both acute and chronic ENL. [a]
Difference in score in Quality of Life assessment between start and end	No statistically significant difference in score changes for any of the SF-36 scales between study arms.			13 (1 RCT)	⊕◯◯◯ VERY LOW [b,c]	Data was read from graph.

Explanations

a. ENL = erythema nodosum leprosum, b. Small RCT with inadequate blinding, c. No confidence intervals given; no statistical significance

Cyclosporin compared to placebo or cyclosporin in observational studies for lichen planus					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Complete remission assessed with: reported itch and rash after 5 mg/kg/day cyclosporin (unknown duration) [a]	Complete response in 100% (6 out of 6 patients).		8 (1 observational study)	⊕◯◯◯ VERY LOW [b,c,d]	2 patients dropped out due to adverse events (not analysed). Itch was relieved after 7.5 days; rash after 6 weeks.
Complete remission assessed with: symptoms after 4-5 mg/kg/day cyclosporin for 1 month [e]	100% remission (20 out of 20 patients).		20 (1 observational study)	⊕◯◯◯ VERY LOW [b,c,d]	2 partial relapses; in remission with topical treatment.
Improvement assessed with: reduction in clinical symptoms after up to 1 month 3 mg/kg/day cyclosporin [f]	80% (4 out of 5) had an excellent improvement.		5 (1 observational study)	⊕◯◯◯ VERY LOW [b,c,d]	Significant decrease of inflammatory cells.

Explanations
a. Higgins et al. 1989[56], b. Small series with sparse data, c. Unclear outcome assessment and measure, d. All small case series show a large effect, e. Leoni et al. 1997[57], f. Mozzanica et al. 1991[55]

Cyclosporin compared to placebo or cyclosporin in observational studies for lichen sclerosus					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Change in mean total symptom score assessed with: Severity of itch, burning, pain and stinging after 3-4 mg/kg/day cyclosporin for 3 months. [a]	Reduction in mean total symptom score: from 8.6 to 0.8.		5 (1 observational study)	⊕◯◯◯ VERY LOW [b]	No recurrences during 12-month follow-up.

Explanations
a. Bulbul Baskan et al. 2007[58], b. Small series with sparse data and indirect outcome measure

Cyclosporin compared to placebo or cyclosporin in observational studies for livedoid vasculitis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Complete remission assessed with: clinical symptoms after 1.7-3 mg/kg/day cyclosporin [a]	100% complete remission (5 out of 5 patients).		5 (1 observational study)	⊕◯◯◯ VERY LOW [b]	Different duration of treatment in each patient. 2 recurrences after 5-6 months, healed after increase in dosage.

Explanations
a. Leclerc et al. 2000[59], b. Small series with sparse data and possible bias in outcome assessment

Cyclosporin compared to placebo or cyclosporin in observational studies for palmoplantar pustulosis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Effectiveness assessed with: % of patients with significant reduction in all efficacy parameters after 1-month cyclosporin 1 mg/kg or placebo [a,b]	In the placebo group 19% and in the cyclosporin group 48%.		58 (1 RCT)	⊕⊕◯◯ LOW [c,d]	
Effectiveness assessed with: % of patients with reduction of 50% or more in the number of pustules compared with baseline after placebo or 1 month 2.5 mg/kg/day cyclosporin [e]	In the placebo group 21% and in the cyclosporin group 89%.		40 (1 RCT)	⊕⊕◯◯ LOW [c,d]	
Complete response assessed with: after 12 weeks 2.5 mg/kg cyclosporin [f]	3 out of 7 patients.		7 (1 observational study)	⊕◯◯◯ VERY LOW [g,h]	Rapid relapse after withdrawal of cyclosporin.
Complete response assessed with: after cyclosporin 2.5 mg/kg for 1 to 3 months [i]	Complete clearing of skin symptoms: 6 out of 6 patients (100%)		6 (1 observational study)	⊕◯◯◯ VERY LOW [h,j]	Patients were allowed to use topical steroids for treatment of psoriasis elsewhere.
Pain score assessed with: score before and after 8 weeks cyclosporin 2-3 mg/kg/day [k,l]	Reduction of pain score from 2.2 to 0.5.		7 (1 observational study)	⊕◯◯◯ VERY LOW [m]	All patients also had pustulotic arthro-osteitis. Unclear concomitant medication. Data was read from graph.

Cyclosporin compared to placebo or cyclosporin in observational studies for palmoplantar pustulosis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Modified PASI assessed with: score before and after 8 weeks cyclosporin 2-3mg/kg/day [k]	Reduction of modified PASI from 12 to 3.		7 (1 observational study)	⊕◯◯◯ VERY LOW [m]	All patients also had pustulotic arthro-osteitis. Unclear concomitant medication. Data was read from graph.
Quality of life assessed with: SF-8 questionnaire scores before and after 8 weeks cyclosporin 2-3 mg/kg/day [k]	Reduction of SF-8 score from 24 to 12.		7 (1 observational study)	⊕◯◯◯ VERY LOW [m]	All patients also had pustulotic arthro-osteitis. Unclear concomitant medication. Data was read from graph.

Explanations

a. Erkko et al. 1998[61], b. reduction of 50% or more in the number of pustules compared with baseline and reduction of erythema, infiltration and scaling, c. Unclear concealment of allocation, adequate blinding and selected reporting, d. Unclear confidence intervals, e. Reitamo et al. 1993[60], f. Meinardi et al. 1990[64], g. Small series with sparse data, h. No explanation was provided, i. Reitamo et al 1989[63], j. Small series with sparse data and varying treatment duration, k. Hayama et al. 2010[62], l. 0- no pain, 1- mild pain not interfering with daily activities, 2- moderate pain interfering but not preventing daily activities, and 3- severe pain preventing daily activities, m. Small series with sparse data, large standard deviations and indirect outcome measures

Cyclosporin and prednisolone compared to prednisolone as mono therapy for pemphigus vulgaris						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Complete remission assessed with: after 6 months 5 mg/kg/day cyclosporin and prednisolone or prednisolone as mono therapy [a]	Complete remission in 5 out of 17 (29%) patients of the prednisone only group and in 4 out of 16 (25%) patients in the cyclosporin and prednisolone group.			33 (1 observational study)	⨁◯◯◯ VERY LOW [b]	
Complete remission assessed with: after 12 months cyclosporin 5 mg/kg and prednisolone or with prednisolone as mono therapy [c]	Complete remission in all patients (16 out of 16) of the combination group. No data about control group.			31 (1 observational study)	⨁◯◯◯ VERY LOW [d]	Groups were not comparable in terms of prednisone dosage. Retrospective control group.
Time to remission (months) assessed with: with 2.5-3.0 mg/kg cyclosporin and prednisolone or prednisolone as mono therapy [e]	The mean time to remission (months) was **7.2**	The mean time to remission (months) in the intervention group was 0,9 more (0 to 0)	-	34 (1 observational study)	⨁◯◯◯ VERY LOW [f,g,h]	Groups were not comparable in terms of prednisone dosage. Retrospective study. 7 patients lost to follow-up (not analysed). Relapse during therapy was similar in both groups.

Time to remission (days) assessed with: with 5 mg/kg cyclosporin and prednisolone or prednisolone as mono therapy [i]	The mean time to remission (days) was **28** days	The mean time to remission (days) in the intervention group was 3 days fewer (0 to 0)	-	18 (1 observational study)	⊕◯◯◯ VERY LOW [d,g,h]	No demographic information. In both groups 2 patients relapsed.

Explanations

a. Ioannides et al. 2000[66], b. Unclear outcome assessment, c. Lapidoth et al. 1994[68], d. Sparse data and concerns about comparability, e. Olszewska et al. 2007[65], f. Concerns about comparability (dosage and drop outs), g. Inconsistent results regarding the time to remission compared with other studies, h. Very large standard deviations; confidence interval for difference not given, i. Chrysomallis et al. 1994[67]

Cyclosporin compared to placebo or cyclosporin in observational studies for pemphigus vulgaris					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Effectiveness assessed with: global assessment (treatment success) after 6 months 6-8 mg/kg cyclosporin [a]	1 out of 4 patients achieved treatment success with cyclosporin as monotherapy. 2/3 failures in remission after starting combination therapy. 4 out of 4 in remission with combination therapy (prednisolone).		8 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	Concomitant prednisolone.
Effectiveness assessed with: global assessment (treatment success) after 2 to 15 months 6-8 mg/kg cyclosporin [d]	1 out of 4 patients achieved treatment success with cyclosporin as monotherapy. 2/3 failures in remission after starting combination therapy. 5 out of 5 in remission with combination therapy (prednisolone).		9 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	Concomitant prednisolone. 4 out of 5 with combination therapy relapsed; relapses treated with prednisolone.
Complete remission assessed with: after 13 to 20 months 1-3 mg/kg cyclosporin. [e]	6 out of 6 showed (100%) complete remission.		6 (1 observational study)	⨁◯◯◯ VERY LOW [c,f]	Time to remission was 8-10 weeks. No recurrences during follow-up.
Complete remission assessed with: after 12 to 43 months 3-5 mg/kg cyclosporin [g]	5 out of 8 patients (62.5%).		8 (1 observational study)	⨁◯◯◯ VERY LOW [f]	Concomitant prednisolone. Reduction of prednisone dosage by 33%.

Explanations

a. Barthelemy et al. 1986[71], b. Small series with unclear outcome assessment and dosage, c. No explanation was provided, d. Barthelemy et al. 1988[70], e. Mobini et al. 1997[69], f. Small series with varying dosage, g. Veraldi et al. 1997[72]

Cyclosporin compared to placebo or cyclosporin in observational studies for prurigo					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Effectiveness assessed with: physician global assessment (very good, good, minimal, no response) after cyclosporin 3-5 mg/kg/day [a]	71.4% (10 out of 14 patients) had a very good response.		14 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	Heterogenic patient group with several underlying factors contributing to prurigo. The best results were seen in atopic patients.
Itch intensity score assessed with: before and after 2 months 5 mg/kg cyclosporin [d]	Significant decrease of itch intensity in 100% of patients.		10 (1 observational study)	⨁◯◯◯ VERY LOW [c,e]	One relapse after 1 month follow-up.
Clinical improvement assessed with: assessed by physician after 6-8 months 2.5 mg/kg/day cyclosporin [f]	18/19 (94.7%) progressive clinical improvement.		19 (1 observational study)	⨁◯◯◯ VERY LOW [c,e]	Skin biopsy: significant reduction of all cell types and adhesion molecules, except for IgG and IgM. No relapses during follow-up.

Explanations

a. Siepmann et al. 2008[75], b. Sparse data, varying dosage, unknown treatment duration and follow-up, c. Imprecise and indirect outcome measure, d. Teofoli et al. 1998[74], e. Unclear outcome assessment, f. Umana et al. 2002[73]

Cyclosporin 4 mg/kg/day compared to oral prednisolone 0.75 mg/kg/day for pyoderma gangrenosum[76]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with oral prednisolone 0.75 mg/kg/day	**Risk with cyclosporin 4 mg/kg/day**				
Speed of healing assessed with: cm2 per day healing of ulcers on digital images after 6 weeks		The mean speed of healing in the intervention group was 0,003 lower (0,2 lower to 0,21 higher)	-	108 (1 RCT)	⊕⊕⊕◯ MODERATE [a]	Primary outcome. 20% unblinded outcome data.
Healed ulcus assessed with: % of patients with healed ulcus after 6 months	47% in cyclosporin group versus also 47% in prednisolone group			112 (1 RCT)	⊕⊕⊕◯ MODERATE [a]	Most important secondary outcome. Hazard ratio was 0.94 (95% CI 0.55 to 1.63).
Global treatment response assessed with: 7 point Likert scale (completely clear - worse) by clinicians and participants and from digital images	Clinicians: 59% completely or almost clear with cyclosporin versus 51% in prednisolone group. Patients: 49% completely or almost clear with cyclosporin versus 53% in prednisolone group. From images: 54% completely or almost clear with cyclosporin versus 44% in prednisolone group.			112 (1 RCT)	⊕⊕◯◯ LOW [a,b]	Data read from graph. Percentage can differ slightly from original data.
Resolution of inflammation assessed with: a scale by Foss et al. recorded by clinicians and participants after 6 weeks	5 (9%) in cyclosporin group versus 6 (12%) in prednisolone group.			107 (1 RCT)	⊕⊕◯◯ LOW [a,b]	Odds ratio 1.03 (0.27 to 3.97).

Cyclosporin 4 mg/kg/day compared to oral prednisolone 0.75 mg/kg/day for pyoderma gangrenosum[76]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with oral prednisolone 0.75 mg/kg/day	**Risk with cyclosporin 4 mg/kg/day**				
Self-reported pain assessed with: mean severity of pain using a score from 0 to 4 (none, mild, moderate, severe, or extreme) at week 6	1.10 (SD 1.0) mean pain score in cyclosporin group versus 1.49 in prednisolone group (SD 1.3)			77 (1 RCT)	⊕⊕⊕◯ MODERATE [a]	
Quality of life assessed with: the DLQI16, European quality of life-5 dimensions, three levels (EQ-5D-3L and EQ VAS) at final visit [c]	DLQI (range 0-30, high score=worse): 4.8 (SD 6.8) cyclosporin versus 6.3 (SD 7.6) prednisolone EQ-5D-3L (range-0.594-1.000, (low scores=worse): 0.76 (SD 0.30) cyclosporin versus 0.63 (SD 0.41) prednisolone EQ VAS (range 0-100, low scores=worse): 73.2 (SD 20.5) cyclosporin versus 70.6 (SD 22.3) prednisolone			112 (1 RCT)	⊕⊕◯◯ LOW [a,d]	
Number of treatment failures assessed with: % patients who withdrew from randomised treatment (intolerance, worsening, target lesion unhealed after 6 months)	Treatment failure was documented in 29/59 (49%) participants in the cyclosporin group and 26/53 (49%) in the prednisolone group (p=0.88).			112 (1 RCT)	⊕⊕⊕◯ MODERATE [a]	

Cyclosporin 4 mg/kg/day compared to oral prednisolone 0.75 mg/kg/day for pyoderma gangrenosum[76]						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with oral prednisolone 0.75 mg/kg/day	**Risk with cyclosporin 4 mg/kg/day**				
Adverse reactions assessed with: adverse events that were possibly, probably, or definitely related to the study drug	40 (68%) participants in the cyclosporin group and 35 (66%) in the prednisolone group experienced at least one adverse reaction			112 (1 RCT)	⊕⊕⊕◯ MODERATE [a]	9 serious adverse reactions occurred: 2 in the cyclosporin group and 7 in the prednisolone group.
Time to recurrence assessed with: interval between the target lesion healing and a further episode of pyoderma gangrenosum (at any site).	The mean time to recurrence was **612** days	The mean time to recurrence in the intervention group was 30 days fewer (0 to 0)	-	52 (1 RCT)	⊕⊕◯◯ LOW [a]	The period of follow- up available varied depending on the time at which participants were randomised into the trial. Hazard ratio 1.43 (95%CI 0.5-4.07).

Explanations

a. No blinding of clinicians and participants, large confidence intervals, doubts about intention to treat analyses, b. Indirect outcome measure, c. DLQI = dermatology life quality index, EQ-5D-3L=European quality of life-5 dimensions, three levels; EQ VAS=European quality of life visual analogue scale, d. Large standard deviation, variation in number of patients with data, e. CI for days not reported. Hazard ratio with CI is reported.

Cyclosporin (observational studies) for pyoderma gangrenosum					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Healing of pyoderma ulcer assessed with: Healing during treatment with cyclosporin at 5–7 mg/kg/day (2-10 months)[a]	7 out of 8 patients (87.5%) showed complete healing after 2-10 months therapy (mean 4.3 months).		8 (1 observational study)	⊕◯◯◯ VERY LOW [b]	Systemic corticosteroids 1–2 mg/kg/d with progressive reduction were allowed. One patient died during treatment.
Time to healing assessed with: Time to re-epithelialization after 4-5mg/kg/day cyclosporin[c]	Mean time to healing 68 cm2/week.		5 (1 observational study)	⊕◯◯◯ VERY LOW [b]	Concomitant medication with prednisolone was allowed. No relapses during follow-up. All ulcers healed after 8 weeks.
Complete healing assessed with: Healing after 8-10 mg/kg/day cyclosporin for 3-7 months[d]	4 out of 7 patients (57%) healed completely.		7 (1 observational study)	⊕◯◯◯ VERY LOW [b]	One patient required continuous cyclosporin treatment.
Complete healing assessed with: complete healing after cyclosporin 5-6mg/kg/day[e]	10 out of 11 patients (91%) healed completely.		11 (1 observational study)	⊕◯◯◯ VERY LOW [b]	Unknown duration of treatment. One patient was treated with 10 mg/kg/day. Two relapses. One patient lost to follow-up. 3 patients with concomitant prednisolone.
Complete healing assessed with: Healing after 3-6 mg/kg/day cyclosporin for a mean 3.5 months[f]	49 out of 51 disease episodes (96%) showed complete healing.		22 (1 observational study)	⊕◯◯◯ VERY LOW [b]	Relapses in 66%. 9 patients required continuous cyclosporin treatment or concomitant medication.

Explanations a. Perreira et al. 2013[78], b. Small number of patients, retrospective, sparse data, c. Capella et al. 1999[77], d. Elgart et al. 1991[81], e. Matis et al. 1992[80], f. Vidal et al. 2004[79]

Cyclosporin compared to placebo or cyclosporin in observational studies for sarcoidosis					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Percentage of patients with total remission of skin lesions assessed with: after 5 mg/kg/day cyclosporin (unknown duration) [a]	All (6 out of 6; 100%) skin lesions were healed at 6 months.		6 (1 observational study)	⨁◯◯◯ VERY LOW [b]	Flucortolone and methotrexate as concomitant medication. No relapses during 6 months follow-up.

Explanations
a. Pia et al. 1996[82], b. Unknown treatment duration, small series.

Cyclosporin and iloprost compared to iloprost as monotherapy for systemic slerosis						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with iloprost as monotherapy	**Risk with cyclosporin and iloprost**				
Change in Skin score assessed with: before and after 12 months iloprost or iloprost and cyclosporin 2.5 mg/kg/day [a]	The mean change in Skin score was **-2.1**	The mean change in Skin score in the intervention group was 2 lower (0 to 0)	-	20 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	

Explanations
a. Filaci et al. 1999[84], b. Unclear outcome assessment, c. No confidence interval for difference

Cyclosporin compared to placebo or cyclosporin in observational studies for systemic slerosis						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with cyclosporin**				
Change in Skin score assessed with: after 48 weeks cyclosporin 1 mg/kg/day (increased monthly to a maximum of 5 mg/kg) or placebo [a]	The mean change in Skin score was **-0.8**	The mean change in Skin score in the intervention group was 1,5 lower (0 to 0)	-	23 (1 observational study)	⨁◯◯◯ VERY LOW [b,c,d]	Placebo comparison group from a different study.
Effectiveness assessed with: global assessment after 3-6 years 2.5 mg/kg/day cyclosporin [e]	After 3 years 7 out of 9 patients had a progressive decline in skin involvement.			9 (1 observational study)	⨁◯◯◯ VERY LOW [c,f]	
Change in Skin score from baseline assessed with: after 6 months 5 mg/kg/day cyclosporin [g]	Baseline: 39.6 ±5.8 After 6 months: 31.0 ±5.6 (p<0.05)			8 (1 observational study)	⨁◯◯◯ VERY LOW [c,d,f]	Concomitant medication was used in some patients. Two dropouts due to death at 7 and 9 months.
Skin elastometry assessed with: global improvement after 4 months 3-5 mg/kg/day cyclosporin [h]	Significant improvement of skin elastometry.			5 (1 observational study)	⨁◯◯◯ VERY LOW [c,f]	Marked regression of skin fibrosis with a reduction of cutaneous infiltrates in all 5 patients.

Cyclosporin compared to placebo or cyclosporin in observational studies for systemic slerosis						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with placebo	**Risk with cyclosporin**				
Effectiveness assessed with: global assessment after 2.4-6.9 mg/kg/day cyclosporin[i]	5 out of 8 patients had marked improvements of symptoms.			8 (1 observational study)	⨁◯◯◯ VERY LOW [c,j]	Concomitant medication was used in some patients.
Effectiveness assessed with: global assessment after 3-26 months 2.2-5.6 mg/kg cyclosporin [k]	Clinical improvement in 4 out of 5 patients. Progression of disease in 1.			5 (1 observational study)	⨁◯◯◯ VERY LOW [c,j]	

Explanations

a. Clements et al. 1993[83], b. Placebo group from other study, c. No explanation was provided, d. Very large standard deviations; unknown confidence interval for difference, e. Basso et al. 2001[85], f. Small series with unclear outcome assessment, g. Gisslinger et al. 1991[86], h. Ippoliti et al. 1994[87], i. Morton et al. 2000[88], j. Small series with sparse data and varying treatment duration, k. Worle et al. 1990[89]

Cyclosporin 3 mg/kg/day and tapering to 2mg/kg/day after 7 days compared to systemic corticosteroids for toxic epidermal necrolysis/Stevens Johnson syndrome						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with systemic corticosteroids	**Risk with cyclosporin 3 mg/kg/day and tapering to 2mg/kg/day after 7 days**				
Complete healing of the skin without any erosion (re-epithelization) assessed with: after cyclosporin or dexamethasone injection followed by oral prednisolone in the dosage of ≥1 mg/kg/day [a]	The mean complete healing of the skin without any erosion (re-epithelization) was **23** days	The mean complete healing of the skin without any erosion (re-epithelization) in the intervention group was 8,5 days fewer (0 to 0)	-	20 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	5 out of 11 patients received 7 days cyclosporin (instead of 14) due to good response. Two dead in corticosteroid group (not analysed).
Duration of hospitalization (days)	The mean duration of hospitalization (days) was **26** days	The mean duration of hospitalization (days) in the intervention group was 7,3 days fewer (0 to 0)	-	20 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	
Onset of stabilization (days) assessed with: when new lesions ceased to appear	The mean onset of stabilization (days) was **4.75** days	The mean onset of stabilization (days) in the intervention group was 1,6 days fewer (0 to 0)	-	20 (1 observational study)	⨁◯◯◯ VERY LOW [b,c]	

Explanations

a. Singh et al. 2013[93], b. Retrospective control cohort. No intention to treat analysis. Wide standard deviations. No confidence interval for difference., c. No explanation was provided

Cyclosporin 3-5 mg/kg/day for an average of 7 days compared to intravenous immunoglobulin for toxic epidermal necrolysis/Stevens Johnson syndrome						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with intravenous immunoglobulin	**Risk with cyclosporin 3-5 mg/kg/day for an average of 7 days**				
In-hospital mortality (SMR) assessed with: standardized mortality ratio (sum of observed deaths/sum of expected deaths) after cyclosporin or IVIg 1 g/kg/d for 3 days [a]	The mean in-hospital mortality (SMR) was **1.43**	The mean in-hospital mortality (SMR) in the intervention group was 1,01 lower (0 to 0)	-	50 (1 observational study)	⨁⨁◯◯ LOW [b]	No difference regarding corticosteroids beforehand. Retrospective cohorts. The observed morality was 29.7% (N = 11) for IVIg- and 5.9% (N = 1) for cyclosporin patients.

Explanations

a. Kirchhof et al. 2014[91], b. Retrospective cohorts with possible selection bias

Cyclosporin 3.0 mg/kg/day for 2 weeks and then tapering compared to prednisone > 1mg/kg and cyclophosphamide 300 mg/kg for toxic epidermal necrolysis/Stevens Johnson syndrome						
Outcomes	**Anticipated absolute effects*** (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
	Risk with prednisone > 1mg/kg and cyclophosphamide 300 mg/kg	**Risk with cyclosporin 3.0 mg/kg/day for 2 weeks and then tapering**				
Time to arrest assessed with: Days until no new erosions of blisters [a]	The mean time to arrest was **3.6** days	The mean time to arrest in the intervention group was 2,2 days fewer (0 to 0)	-	17 (1 observational study)	⊕◯◯◯ VERY LOW [b]	Three patients died in control group. Not due to adverse events.
Time to re-epithelization assessed with: Days until no erosions where present	The mean time to re-epithelization was **17.6** days	The mean time to re-epithelization in the intervention group was 5,6 days fewer (0 to 0)	-	17 (1 observational study)	⊕◯◯◯ VERY LOW [c]	

Explanations

a. Arevalo et al. 2000[90], b. Indirect outcome measure, c. Large standard deviations

Cyclosporin compared to placebo or cyclosporin in observational studies for toxic epidermal necrolysis/Stevens Johnson syndrome					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Mortality after 1 month assessed with: after 3 mg/kg/day cyclosporin for 10 days and tapering in a month. [a]	SCORTEN predicted deaths: 2.75. Observed: 0.		29 (1 observational study)	⊕⊕◯◯ LOW	Patients with reduced creatinine clearance were not included in the study. In three patients cyclosporin was stopped because of side-effects.

Cyclosporin compared to placebo or cyclosporin in observational studies for toxic epidermal necrolysis/Stevens Johnson syndrome					
Outcomes	**Anticipated absolute effects*** (95% CI)	Relative effect (95% CI)	№ of participants (studies)	Quality of the evidence (GRADE)	Comments
Duration of hospitalization (days) assessed with: after 3 mg/kg/day cyclosporin for 10 days and tapering in a month. [a]	16 days (SD 9.1 days).		29 (1 observational study)	⊕◯◯◯ VERY LOW [b]	
Stabilization of BSA involvement assessed with: between day 0 and day 3 after 3 mg/kg/day cyclosporin. [a]	18 out of 29 patients (62%).		29 (1 observational study)	⊕◯◯◯ VERY LOW [b]	
Time to re-epithelisation assessed with: complete healing of the skin after 3 mg/kg/day cyclosporin for 10 days and tapering in a month. [a]	12.4 days (SD 7.7 days).		29 (1 observational study)	⊕◯◯◯ VERY LOW [b]	

Explanations
a. Valeyrie-Allanore et al. 2010[92]. b. No explanation was provided

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

***The risk in the intervention group** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval. SD: Standard deviation.

APPENDIX 2 Risk of bias table of included randomized controlled trials

	Adequate randomisation?	Adequate concealment of allocation?	Adequate blinding?	Incomplete data reported?	Free of selected reporting?	Free of other bias?
Granlund *et al.* 1996[28]	NO	NO	Participants YES Researchers UNCLEAR Outcome assessment UNCLEAR	NO	YES	YES
Grattan et al. 2000[29]	YES	UNCLEAR	Participants YES Researchers UNCLEAR Outcome assessment UNCLEAR	NO	UNCLEAR	YES
Vena et al. 2006[30]	UNCLEAR	UNCLEAR	Participants UNCLEAR Researchers UNCLEAR Outcome assessment UNCLEAR	NO	UNCLEAR	YES
Erkko et al. 1998[61]	YES	UNCLEAR	Participants YES Researchers UNCLEAR Outcome assessment UNCLEAR	NO	UNCLEAR	YES
Reitamo et al. 1993[60]	YES	UNCLEAR	Participants UNCLEAR Researchers UNCLEAR Outcome assessment UNCLEAR	NO	UNCLEAR	YES
Ormerod et al. 2015[76]	YES	YES	Participants NO Researchers NO Outcome assessment YES	NO	YES	YES
Lambert et al. 2016[53]	UNCLEAR	UNCLEAR	Participants YES Researchers UNCLEAR Outcome assessment YES	NO	UNCLEAR	YES

REFERENCES

1 Borel JF. Ciclosporin and its future. Progress in Allergy. 1986;38:9-18.
2 Faulds D, Goa KL, Benfield P. Cyclosporin. A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic use in immunoregulatory disorders. Drugs. 1993;45(6):953-1040. Erratum in: Drugs. 1993;46(3):377.
3 https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm (accessed 10.10.2017)
4 http://www.eudrapharm.eu (accessed 10.10.2017)
5 Dehesa L, Abuchar A, Nuno-Gonzalez A, Vitiello M, Kerdel FA. The use of cyclosporine in dermatology. J Drugs Dermatol. 2012;11(8):979-87.
6 Stafford RS. Regulating off-label drug use--rethinking the role of the FDA. N Engl J Med. 2008;3;358(14):1427-9.
7 Brockmeyer NH, Brucklacher U, Potthoff A, Reich-Schupke S. Off-label use in dermatology in Germany: what has changed since 2004. J Dtsch Dermatol Ges. 2009;7(11):938-45.
8 Sugarman JH, Fleischer AB Jr, Feldman SR. Off-label prescribing in the treatment of dermatologic disease. J Am Acad Dermatol. 2002;47(2):217-23.
9 Higgings JPT, Altman DG. Assessing risk of bias in included studies. In: Higgings JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of interventions Version 5.1.0. The Cochrane Collaboration; 2011.
10 https://gdt.gradepro.org/app/handbook/handbook.html (accessed 10.10.2017)
11 von EE, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. [The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies]. Rev Esp Salud Publica 2008;82:251-259.
12 http://www.medicines.org.uk/emc/medicine/1307 (accessed 10.10.2017)
13 Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383-94.
14 Açıkgöz G, Calışkan E, Tunca M, Yeniay Y, Akar A. The effect of oral cyclosporine in the treatment of severe alopecia areata. Cutan Ocul Toxicol. 2014;33(3):247-52.
15 Constantopoulos, A., Tsoumacas, C. and Tsivitanidou, T. Cyclosporine in severe alopecia areata in children. J Eur Acad Dermatol Venereol. 1996;7:190-192.
16 Ferrando J, Grimalt R. Partial response of severe alopecia areata to cyclosporine A. Dermatology. 1999;199(1):67-9.
17 Kim BJ, Min SU, Park KY, Choi JW, Park SW, Youn SW et al. Combination therapy of cyclosporine and methylprednisolone on severe alopecia areata. J Dermatolog Treat. 2008;19(4):216-20.
18 Shaheedi-Dadras M, Karami A, Mollaei F, Moravvej H, Malekzad F. The effect of methylprednisolone pulse-therapy plus oral cyclosporine in the treatment of alopecia totalis and universalis. Arch Iran Med. 2008;11(1):90-3. Erratum in: Arch Iran Med. 2008;11(2):241.
19 Shapiro J, Lui H, Tron V, Ho V. Systemic cyclosporine and low-dose prednisone in the treatment of chronic severe alopecia areata: a clinical and immunopathologic evaluation. J Am Acad Dermatol. 1997;36(1):114-7.
20 Teshima H, Urabe A, Irie M, Nakagawa T, Nakayama J, Hori Y. Alopecia universalis treated with oral cyclosporine A and prednisolone: immunologic studies. Int J Dermatol. 1992;31(7):513-6.
21 Avci O, Gürler N, Güneş AT. Efficacy of cyclosporine on mucocutaneous manifestations of Behçet's disease. J Am Acad Dermatol. 1997;36(5 Pt1):796-7.
22 Díaz-Llopis M, Cervera M, Menezo JL. Cyclosporin A treatment of Behçet's disease: a long-term study. Curr Eye Res. 1990;9 Suppl:17-23.

23 Granlund H, Erkko P, Reitamo S. Long-term follow-up of eczema patients treated with cyclosporine. Acta Derm Venereol. 1998;78(1):40-3.

24 Higgins EM, McLelland J, Friedmann PS, Matthews JN, Shuster S. Oral cyclosporin inhibits the expression of contact hypersensitivity in man. J Dermatol Sci. 1991;2(2):79-83.

25 Flori L, Perotti R, Mazzatenta C, Andreassi L. Cyclosporin A in the treatment of severe allergic contact dermatitis. J Eur Acad Dermatol Venereol. 1993;2:200-206.

26 Christoffers WA, Politiek K, Coenraads PJ, van der Schaft J, de Bruin-Weller MS, Schuttelaar ML. Drug survival of cyclosporine in the treatment of hand eczema: a multicentre, daily use study. J Eur Acad Dermatol Venereol. 2016;30(1):63-6.

27 Reitamo S, Granlund H. Cyclosporin A in the treatment of chronic dermatitis of the hands. Br J Dermatol. 1994;130(1):75-8.

28 Granlund H, Erkko P, Eriksson E, Reitamo S. Comparison of cyclosporine and topical betamethasone-17,21-dipropionate in the treatment of severe chronic hand eczema. Acta Derm Venereol. 1996;76(5):371-6.

29 Grattan CE, O'Donnell BF, Francis DM, Niimi N, Barlow RJ, Seed PT et al. Randomized double-blind study of cyclosporin in chronic 'idiopathic' urticaria. Br. J. Dermatol. 2000;143(2):365-372.

30 Vena GA, Cassano N, Colombo D, Peruzzi E, and Pigatto P. Cyclosporine in chronic idiopathic urticaria: a double-blind, randomized, placebo-controlled trial. J Am Acad Dermatol. 2006;55(4):705-709.

31 Di Leo E, Nettis E, Aloia AM, Moschetta M, Carbonara M, Dammacco F et al. Cyclosporin-A efficacy in chronic idiopathic urticaria. Int J Immunopathol Pharmacol. 2011;24(1):195-200.

32 Di Gioacchino M, Di Stefano F, Cavallucci E, Verna N, Ramondo S, Paolini F et al. Treatment of chronic idiopathic urticaria and positive autologous serum skin test with cyclosporine: clinical and immunological evaluation. Allergy Asthma Proc. 2003;24(4):285-90.

33 Serhat Inaloz H, Ozturk S, Akcali C, Kirtak N, Tarakcioglu M. Low-dose and short-term cyclosporine treatment in patients with chronic idiopathic urticaria: a clinical and immunological evaluation. J Dermatol. 2008;35(5):276-82.

34 Loria MP, Dambra PP, D'Oronzio L, Nettis E, Pannofino A, Cavallo E et al. Cyclosporin A in patients affected by chronic idiopathic urticaria: a therapeutic alternative. Immunopharmacol Immunotoxicol. 2001;23(2):205-13.

35 Toubi E, Blant A, Kessel A, Golan TD. Low-dose cyclosporin A in the treatment of severe chronic idiopathic urticaria. Allergy. 1997;52(3):312-6.

36 Ilter N, Gürer MA, Akkoca MA. Short-term oral cyclosporine for chronic idiopathic urticaria. J Eur Acad Dermatol Venereol. 1999;12(1):67-9.

37 Barlow RJ, Kobza BA, Greaves MW. Treatment of severe, chronic urticaria with cyclosporine A. Eur J Dermatol. 1993;3:273-5.

38 Daher N, Lieberman P. Immunosuppression medication as an approach to treat chronic idiopathic urticaria in children. Curr Allergy Asthma Rep. 2010;10(1):1-2.

39 Boubouka CD, Charissi C, Kouimintzis D, Kalogeromitros D, Stavropoulos PG, Katsarou A. Treatment of autoimmune urticaria with low-dose cyclosporin A: A one-year follow-up. Acta Derm Venereol. 2011;91(1):50-4.

40 Ohtsuka T. Response to oral cyclosporine therapy and high sensitivity-CRP level in chronic idiopathic urticaria. Int J Dermatol. 2010;49(5):579-84.

41 Doshi DR, Weinberger MM. Experience with cyclosporine in children with chronic idiopathic urticaria. Pediatr Dermatol. 2009;26(4):409-13.

42 Kessel A, Toubi E. Cyclosporine-A in severe chronic urticaria: the option for long-term therapy. Allergy. 2010;65(11):1478-1482.

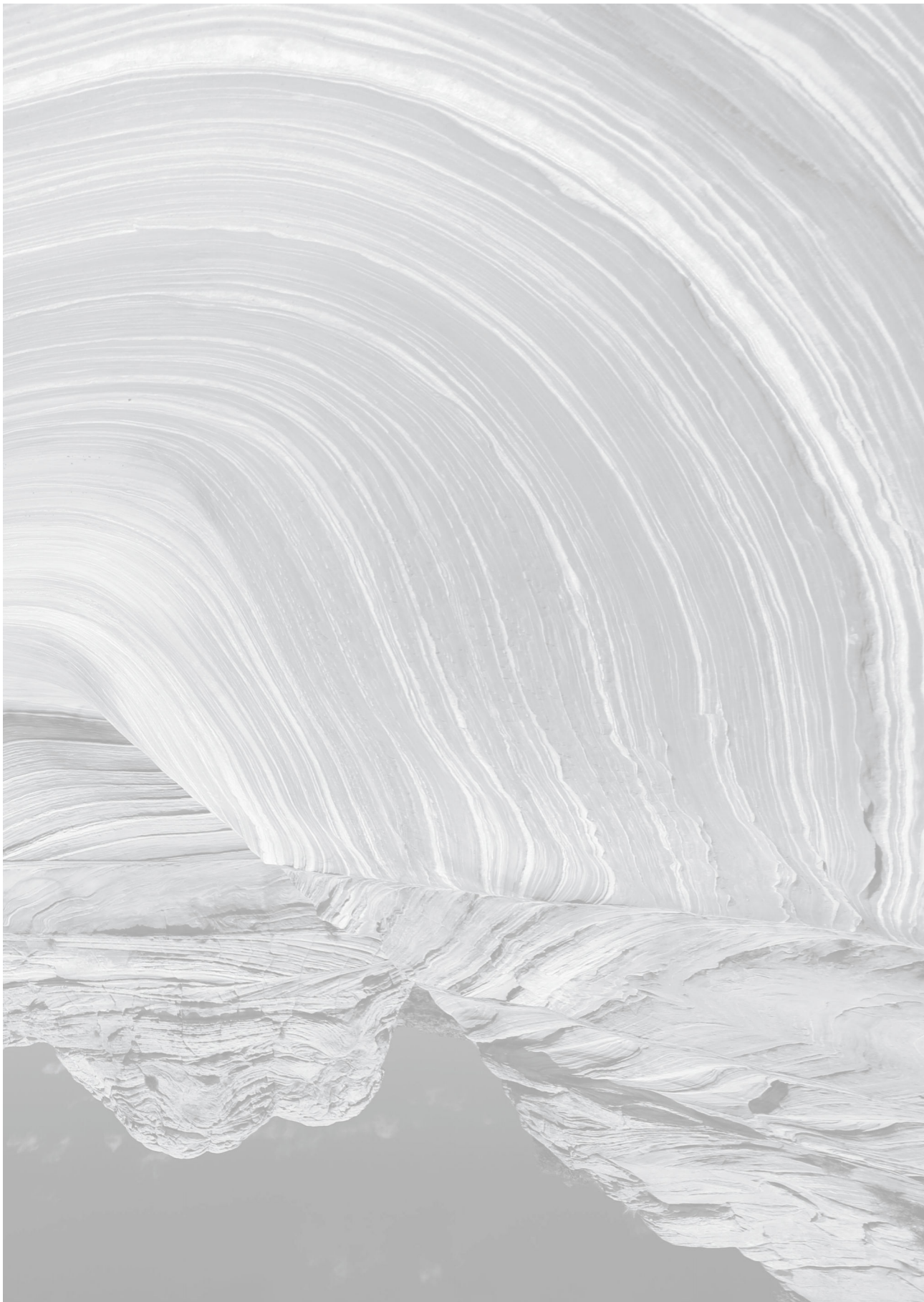
43 Lim KH, Lee SE, Kang MG, Park HK, Kim MY, Jo EJ et al. Cyclosporine in the management of chronic urticaria: the efficacy and safety data in Koreans. Allergy: European Journal of Allergy and Clinical Immunology. 2013;68(Suppl. 97):498-636.
44 Hurabielle C, Bedane C, Avenel-Audran M, Adamski H, Aubin F, Jeanmougin M et al. No major effect of cyclosporin A in patients with severe solar urticaria: a french retrospective case series. Acta Derm Venereol. 2015;95(8):1030-1.
45 Neverman L, Weinberger M. Treatment of chronic urticaria in children with antihistamines and cyclosporine. J Allergy Clin Immunol Pract. 2014;2(4):434-8.
46 Asero R. Plasma D-dimer levels and clinical response to ciclosporin in severe chronic spontaneous urticaria. J Allergy Clin Immunol. 2015;135(5):1401-3.
47 Cooper DL, Braverman IM, Sarris AH, Durivage HJ, Saidman BH, Davis CA et al. Cyclosporine treatment of refractory T-cell lymphomas. Cancer. 1993;71(7):2335-41.
48 Grau JM, Herrero C, Casademont J, Fernández-Solà J, Urbano-Márquez A. Cyclosporine A as first choice therapy for dermatomyositis. J Rheumatol. 1994;21(2):381-2.
49 Dankó K, Szegedi G. Cyclosporin A treatment of dermatomyositis. Arthritis Rheum. 1991;34(7):933-4.
50 Fukamachi S, Kabashima K, Sugita K, Kobayashi M, Tokura Y. Therapeutic effectiveness of various treatments for eosinophilic pustular folliculitis. Acta Derm Venereol. 2009;89(2):155-9.
51 Anderson MD, Zauli S, Bettoli V, Boer J, Jemec GB. Cyclosporine treatment of severe Hidradenitis suppurativa--A case series. J Dermatolog Treat. 2016;27(3):247-50.
52 Aricò M, Colella R, Conter V, Indolfi P, Pession A, Santoro N et al. Cyclosporine therapy for refractory Langerhans cell histiocytosis. Med Pediatr Oncol. 1995;25(1):12-6.
53 Lambert SM, Nigusse SD, Alembo DT, Walker SL, Nicholls PG, Idriss MH et al. Comparison of Efficacy and Safety of Ciclosporin to Prednisolone in the Treatment of Erythema Nodosum Leprosum: Two Randomised, Double Blind, Controlled Pilot Studies in Ethiopia. PLoS Negl Trop Dis. 2016;26;10(2):e0004149.
54 Marlowe SN, Leekassa R, Bizuneh E, Knuutilla J, Ale P, Bhattarai B, et al. Response to ciclosporin treatment in Ethiopian and Nepali patients with severe leprosy Type 1 reactions. Trans R Soc Trop Med Hyg. 2007;101(10):1004-12.
55 Mozzanica N, Cattaneo A, Legori A, Pigatto P, Finzi AF. Immunohistologic evaluation of the effect of cyclosporine treatment on the lichen planus immune infiltrate. J Am Acad Dermatol. 1991;24(4):550-4.
56 Higgins EM, Munro CS, Friedmann PS, Marks JM. Cyclosporin A in the treatment of lichen planus. Arch Dermatol. 1989;125(10):1436.
57 Leoni A, Chieregato C. Cyclosporine in the treatment of diffuse lichen ruber planus. 1997;183:30+33-34.
58 Bulbul Baskan E, Turan H, Tunali S, Toker SC, Saricaoglu H. Open-label trial of cyclosporine for vulvar lichen sclerosus. J Am Acad Dermatol. 2007;57(2):276-8.
59 Leclerc A, Braeken C, Marot L, Tennstedt D, Lachapelle M. La vasculite livédoïde. Traitement par ciclosporine: A propos de cinq cas. Lês Nouvelles Dermatologiques. 2000;19:356-60.
60 Reitamo S, Erkko P, Remitz A, Lauerma AI, Montonen O, Harjula K. Cyclosporine in the treatment of palmoplantar pustulosis. A randomized, double-blind, placebo-controlled study. Arch Dermatol. 1993;129(10):1273-9.
61 Erkko P, Granlund H, Remitz A, Rosen K, Mobacken H, Lindelöf B, et al. Double-blind placebo-controlled study of long-term low-dose cyclosporin in the treatment of palmoplantar pustulosis. Br J Dermatol. 1998;139(6):997-1004.

62 Hayama K, Inadomi T, Fujisawa D, Terui T. A pilot study of medium-dose cyclosporine for the treatment of palmoplantar pustulosis complicated with pustulotic arthro-osteitis. Eur J Dermatol. 2010;20(6):758-62.
63 Meinardi MM, de Rie MA, Bos JD. Oral cyclosporin A is effective in clearing persistent pustulosis palmaris et plantaris. Acta Derm Venereol. 1990;70(1):77-9.
64 Reitamo S, Puska P, Lassus A. Cyclosporin in the treatment of palmo-plantar pustulosis. Br J Dermatol. 1989;120(6):857.
65 Olszewska M, Kolacinska-Strasz Z, Sulej J, Labecka H, Cwikla J, Natorska U, et al. Efficacy and safety of cyclophosphamide, azathioprine, and cyclosporine (ciclosporin) as adjuvant drugs in pemphigus vulgaris. Am J Clin Dermatol. 2007;8(2):85-92.
66 Ioannides D, Chrysomallis F, Bystryn JC. Ineffectiveness of cyclosporine as an adjuvant to corticosteroids in the treatment of pemphigus. Arch Dermatol. 2000;136(7):868-72.
67 Chrysomallis F, Ioannides D, Teknetzis A, Panagiotidou D, Minas A. Treatment of oral pemphigus vulgaris. Int J Dermatol. 1994;33(11):803-7.
68 Lapidoth M, David M, Ben-Amitai D, Katzenelson V, Lustig S, Sandbank M. The efficacy of combined treatment with prednisone and cyclosporine in patients with pemphigus: preliminary study. J Am Acad Dermatol. 1994;30(5 Pt 1):752-7.
69 Mobini N, Padilla T Jr, Ahmed AR. Long-term remission in selected patients with pemphigus vulgaris treated with cyclosporine. J Am Acad Dermatol. 1997;36(2 Pt 1):264-6.
70 Barthelemy H, Frappaz A, Cambazard F, Mauduit G, Rouchouse B, Kanitakis J, et al. Treatment of nine cases of pemphigus vulgaris with cyclosporine. J Am Acad Dermatol. 1988;18(6):1262-6.
71 Barthelemy H, Biron F, Claudy AL, Souteyrand P, Thivolet J. Cyclosporin A: a new immunosuppressive agent in bullous pemphigoid and pemphigus. Transplant Proc 1986;18:913-4.
72 Veraldi S, Bocor M, Sarchi G, Caputo R. Treatment of pemphigus vulgaris with cyclosporine. Journal of Dermatological Treatment. 1997;8:2.
73 Umaña A, Gómez A, Durán MM, Porras L. Lymphocyte subtypes and adhesion molecules in actinic prurigo: observations with cyclosporin A. Int J Dermatol.2002;41(3):139-45.
74 Teofoli P, De Pita O, Frezzolini A, Lotti T. Antipruritic effect of oral cyclosporin A in essential senile pruritus. Acta Derm Venereol. 1998;78(3):232.
75 Siepmann D, Luger TA, Ständer S. Antipruritic effect of cyclosporine microemulsion in prurigo nodularis: results of a case series. J Dtsch Dermatol Ges. 2008;6(11):941-6.
76 Ormerod AD, Thomas KS, Craig FE, Mitchell E, Greenlaw N, Norrie J, et al.; UK Dermatology Clinical Trials Network's STOP GAP Team. Comparison of the two most commonly used treatments for pyoderma gangrenosum: results of the STOP GAP randomised controlled trial. BMJ. 2015;350:h2958.
77 Capella GL, Frigerio E, Fracchiolla C, Altomare G. The simultaneous treatment of inflammatory bowel diseases and associated pyoderma gangrenosum with oral cyclosporin A. Scand J Gastroenterol. 1999;34(2):220-1.
78 Pereira N, Brites MM, Gonçalo M, Tellechea O, Figueiredo A. Pyoderma gangrenosum--a review of 24 cases observed over 10 years. Int J Dermatol. 2013;52(8):938-45.
79 Vidal D, Puig L, Gilaberte M, Alomar A. Review of 26 cases of classical pyoderma gangrenosum: clinical and therapeutic features. J Dermatolog Treat. 2004;15(3):146-52.
80 Matis WL, Ellis CN, Griffiths CE, Lazarus GS. Treatment of pyoderma gangrenosum with cyclosporine. Arch Dermatol. 1992;128(8):1060-4.

81 Elgart G, Stover P, Larson K, Sutter C, Scheibner S, Davis B, et al. Treatment of pyoderma gangrenosum with cyclosporine: results in seven patients. J Am Acad Dermatol. 1991;24(1):83-6.
82 Pia G, Pascalis L, Aresu G, Rosetti L, Ledda MA. Evaluation of the efficacy and toxicity of the cyclosporine A-flucortolone-methotrexate combination in the treatment of sarcoidosis. Sarcoidosis Vasc Diffuse Lung Dis. 1996;13(2):146-52.
83 Clements PJ, Lachenbruch PA, Sterz M, Danovitch G, Hawkins R, Ippoliti A et al. Cyclosporine in systemic sclerosis. Results of a forty-eight-week open safety study in ten patients. Arthritis Rheum. 1993;36(1):75-83.
84 Filaci G, Cutolo M, Scudeletti M, Castagneto C, Derchi L, Gianrossi R, et al. Cyclosporin A and iloprost treatment of systemic sclerosis: clinical results and interleukin-6 serum changes after 12 months of therapy. Rheumatology (Oxford). 1999;38(10):992-6.
85 Basso M, Filaci G, Cutolo M, Murdaca G, Derchi L, Gianrossi R, et al. Long-term treatment of patients affected by systemic sclerosis with cyclosporin A. Ann Ital Med Int. 2001;16(4):233-9.
86 Gisslinger H, Burghuber OC, Stacher G, Schwarz W, Punzengruber C, Graninger W, et al. Efficacy of cyclosporin A in systemic sclerosis. Clin Exp Rheumatol. 1991;9(4):383-90.
87 Ippoliti G, Miori L, Negri M, Rovati B, Lorenzutti F, Zerbinati N, et al. Cyclosporine in treatment of progressive systemic sclerosis: clinical and immunologic findings. Transplant Proc. 1994;26(6):3117-8.
88 Morton SJ, Powell RJ. Cyclosporin and tacrolimus: their use in a routine clinical setting for scleroderma. Rheumatology. 2000;39(8):865-9.
89 Wörle B, Hein R, Krieg T, Meurer M. Cyclosporin in localized and systemic scleroderma--a clinical study. Dermatologica. 1990;181(3):215-20.
90 Arévalo JM, Lorente JA, González-Herrada C, Jiménez-Reyes J. Treatment of toxic epidermal necrolysis with cyclosporin A. J Trauma. 2000;48(3):473-8.
91 Kirchhof MG, Miliszewski MA, Sikora S, Papp A, Dutz JP. Retrospective review of Stevens-Johnson syndrome/toxic epidermal necrolysis treatment comparing intravenous immunoglobulin with cyclosporine. J Am Acad Dermatol. 2014;71(5):941-7.
92 Valeyrie-Allanore L, Wolkenstein P, Brochard L, Ortonne N, Maître B, Revuz J, et al. Open trial of ciclosporin treatment for Stevens-Johnson syndrome and toxic epidermal necrolysis. Br J Dermatol. 2010;163(4):847-53.
93 Singh GK, Chatterjee M, Verma R. Cyclosporine in Stevens Johnson syndrome and toxic epidermal necrolysis and retrospective comparison with systemic corticosteroid. Indian J Dermatol Venereol Leprol. 2013;79(5):686-92.
94 Schram ME, Borgonjen RJ, Bik CM, van der Schroeff JG, van Everdingen JJ, Spuls PI; Off-Label Working and Project Group of Dutch Society of Dermatology and Venereology. Off-label use of azathioprine in dermatology: a systematic review. Arch Dermatol. 2011;147(4):474-88.
95 Nijsten T, Spuls PI, Naldi L, Stern RS. The misperception that clinical trial data reflect long-term drug safety: lessons learned from Efalizumab's withdrawal. Arch Dermatol 2009;145:1037-1039.
96 Lecluse LL, Naldi L, Stern RS, Spuls PI. National registries of systemic treatment for psoriasis and the European 'Psonet' initiative. Dermatology 2009;218:347-356.
97 Weldring T, Smith SMS. Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). Health Services Insights. 2013;6:61-68.

98 Chalmers JR, Simpson E, Apfelbacher CJ, Thomas KS, von Kobyletzki L, Schmitt J et al. Report from the fourth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). Br J Dermatol. 2016;175(1):69-79.
99 Eleftheriadou V, Thomas K, van Geel N, Hamzavi I, Lim H, Suzuki T et al.; Vitiligo Global Issues Consensus Group. Developing core outcome set for vitiligo clinical trials: international e-Delphi consensus. Pigment Cell Melanoma Res. 2015;28(3):363-9.
100 Thorlacius L, Ingram JR, Garg A, Villumsen B, Esmann S, Kirby JS et al. Protocol for the development of a core domain set for hidradenitis suppurativa trial outcomes. BMJ Open. 2017;20;7(2):e014733.
101 Iyengar S, Williamson PR, Schmitt J, Johannsen L, Maher IA, Sobanko JF et al. Development of a core outcome set for clinical trials in rosacea: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. Trials. 2016;17(1):429.

4



5

COMPARISON OF THREE DIFFERENT METHODS OF GUIDELINE DEVELOPMENT

Prospective comparison of three different methods of guideline development for treatment of actinic keratosis

Published as

Borgonjen RJ, van Everdingen JJ, Bik CM, Tuut MK, Spuls PI, van de Kerkhof PC. Prospective comparison of three guideline development methods for treatment of actinic keratosis.

BMJ Qual Saf 2011;20(10):832-41.

SUMMARY

Objectives To compare three methods of guideline development, to see whether using alternative evidence-based methods resulted in variation of recommendations for treating actinic keratosis.

Methods Method 1 followed a standard multiple session evidence-based approach with a working group. In method 2 recommendations were formulated by a working group during a 2-day conference. Method 3 used one epidemiologist to summarise the evidence and one dermatologist to make clinical recommendations afterwards. Graded recommendations and levels of evidence were compared per therapy across three draft guidelines. The primary outcome was the extent of accordance or discordance. Secondary outcomes were total costs and time period necessary to make a draft guideline.

Results Therapeutic recommendations and levels of evidence differed in some occasions. However, intraclass correlations between levels of evidence were significant (method 1 versus 2: $p = 0.003$ and method 1 versus 3: $p < 0.001$). Regarding recommendation variation method 1 and method 2 correlated significant at 0.755 ($p = 0.001$). Method 1 versus 3 and method 2 versus 3 also showed significant, but lower, correlation coefficients (respectively, 0.493 ($p = 0.026$) and 0.673 ($p = 0.007$)). Method 3 was the cheapest and quickest (€ 24,770 and 4 months) and method 1 was the most expensive and slowest method (€ 48,100 and 14 months).

Conclusion The value of a guideline using alternative evidence-based methods seems to at least equal that of a guideline composed in multiple sessions, that is, for topics with a monodisciplinary character and a relatively small number of conducted trials. In addition, the presented alternatives were more time- and cost-efficient.

INTRODUCTION

Actinic keratosis is the most prevalent pre-malignant skin disease in people with skin type I-III and presents itself as solitary or multiple, variable erythematous and irregularly keratinized lesions in sun-exposed areas of the skin.[1,2] Consequently chronic exposure to ultraviolet radiation is an important etiological factor.[3] There are multiple treatment options, varying in (cosmetic) effectiveness and side-effects.[2] A Dutch clinical practice guideline (CPG) is available to guide dermatologists in treating actinic keratosis.[4] Over the last decades the method of developing guidelines has changed from an expert consensus-based approach to an evidence-based approach.[5] These days a search for the best available evidence in literature is obligatory. Although there are several methods used in daily practice, it is more or less standard to develop guidelines with a working group following a transparent evidence-based approach.[6] However, the traditional evidence-based approach in multiple sessions has several problems, such as the time and money required to develop a CPG. The alternative methods presented here have elements that have the potential to be quicker (fewer sessions), thus making it easier to keep abreast of developments as they occur, and to perhaps better sustain group motivation (a problem often encountered by guideline developers). Nevertheless, the question remains whether the quality is affected or not. Hence, we conducted a prospective study, comparing three methods of guideline development, to test whether using alternative evidence-based methods results in variation of guideline recommendations and to analyse the cost-effectiveness.

METHODS

The prospective three-arm comparative trial investigated the following guideline development methods.

Method 1

A working group followed an evidence-based approach with multiple sessions (traditionally 6-12) and was blinded regarding the existence of other working groups. Clinical participants extracted the data in evidence tables and wrote texts. The working group of method 1 held six meetings in which they discussed their individual input and formulated conclusions and recommendations in a draft guideline.

Method 2

A second working group was composed similar to that in method 1 to avoid confounding (Table 1). This group, although aware of being part of an experiment, formulated recommendations during a 2-day conference, without knowing the topic in advance. Outside consultation was prohibited and any additional information used was documented. In 2 days a draft guideline with recommendations for the treatment of actinic keratosis was made, based on self-made evidence tables distilled from received articles. Evidence tables from the evidence report in method 3 were available and used in cases of insurmountable time pressure.

Categories	Method 1	Method 2	Method 3*
Clinical participants	6 dermatologists 1 pathologist	6 dermatologists 1 pathologist	1 dermatologist
Ratio academic/non-academic dermatologists	3:2	3:2	NR
No. of participants with PhD in dermato-oncology	2	2	NR
Age (average)	35-49 yr (43,4)	36-51 yr (45)	NR
Ratio male/female	3:3	3:3	NR
No. of participants with former experience in guideline development	2	2	0
Conflicting interest reported	none	none	none

Table 1: Characteristics of the 3 different methods. * = an epidemiologist made an evidence report. NR = not relevant

Method 3

Method 3 used an epidemiologist who systematically summarised the evidence from the available literature in an evidence report with evidence tables. A single independent dermatologist (not part of any of the other methods) subsequently formulated recommendations for clinical practice, without knowledge of the results of the other groups. However, the dermatologist was aware of being part of a trial. Receiving the same set of articles based on the results of a systematic literature search, each method yielded a draft guideline ready for comparison. While composing

their guidelines, each working group benefited from the full methodological support (for example retrieving and printing requested articles, take down during meetings, answering epidemiological questions), without content interference, of experienced guideline developers with knowledge of research methodology.

Literature

In Medline, EMBASE (Excerpta Medica Database) and CENTRAL (Cochrane Central Register of Controlled Trials) databases (from July 1986 until July 2005) relevant studies were searched.[7] This search was updated to find additional studies. Titles and abstracts were screened for potentially relevant clinical trials and existing guidelines investigating treatment for actinic keratosis. Full text versions were assessed using predefined eligibility criteria. References from potential relevant studies were checked to find additional evidence. The search and selection resulted in a set of articles used in all three methods (Table 4). To facilitate uniform background information for the three investigated methods a set of 43 articles regarding epidemiology, risk factors and aetiology was selected from a search performed homogenous to a systematic review.[69] Additional literature brought in by working group members was allowed, but it was obligatory to document study details.

Level of evidence
1 = one systematic review (A1) or at least 2 studies of level A2 which were carried out independently 2 = at least 2 independent studies of level B 3 = one study of level A2 or B or studies of level C 4 = expert opinion, for example the members of the working group
For articles concerning intervention (prevention or therapy)
A1 = systematic reviews involving at least 2 studies on A2-level, of which the results of separate studies are consistent; A2 = randomised comparative clinical studies of good quality (randomised, double-blind controlled trails) of sufficient size and consistency; B = randomised clinical trials of mediocre quality of insufficient size or other comparative studies (non randomised, comparative cohort study, case-control study); C = non-comparative studies; D = expert opinion, for example the members of the working group.

Table 2: Classification of literature according to the level of evidence (levels according to the Dutch Institute for Healthcare Improvement CBO)

Outcomes

Each recommendation reported as such in the three different draft guidelines was compared per therapy. The primary outcome was the extent of accordance or discordance of each recommendation and underlying evidence across the three guidelines. The underlying scientific evidence for each recommendation was analysed by checking all the references, mentioning additional articles used and mentioning the level of evidence as formulated by the participants of the different methods (Table 2). A panel of three investigators (RB, JvE, CB) graded the recommendations. A grading system with three categories of perceived obligation to follow a recommendation was used: either negative or positive (Table 3).[70,71] Discrepancies were discussed until an agreement was reached. Factors attributing to possible differences were described. Secondary reported outcomes were total costs (for example retrieval of articles, working group expenses, salary and catering) and time period necessary to make a draft guideline (including preparatory work like searching databases and retrieving full text articles).

Low level of obligation	**Moderate level of obligation**	**High level of obligation**
May	**Should**	**Must**
Including the positive formulations: -can be considered -is an option -can be a place for -can be useful	*Including the positive formulations:* -is recommended -is advised -is a preferred treatment -strive for	*Including the positive formulations:* -is strongly recommended -is first choice -is indicated -is standard therapy
May not	**Should not**	**Must not**
Including the negative formulations: -maybe no place for -does not seem useful -is restraint necessary	*Including the negative formulations:* -is not recommended -is discouraged -deserves no recommendation	*Including the negative formulations:* -is strongly discouraged -is no choice -is contraindicated -is no place for

Table 3: Grading of recommendation according to perceived level of obligation, including the Dutch CBO formulations used to make a weak, moderate or strong recommendation.

Statistical analysis

Statistical analysis was performed using SPSS statistical software, version 17.0. Statistical significance was defined as a p value less than 0.05. Recommendations and level of evidence were regarded as ordinal variables varying in value from 1.0 (highly negative recommendation) to 6.0 (highly positive recommendation) and 1.0 (level of evidence 1) to 4.0 (level of evidence 4). The correlation between the recommendations and the levels of evidence was tested using the intraclass correlation coefficient with an absolute agreement definition.

Treatment		Method 1	Method 2	Method 3
Surgical techniques	**Curettage (including supplementary electrodessiction)**	There is a place for treatment with electrodessication provided that there is enough knowledge and skill available (LOE 4)	Reasonable alternative for hyperkeratotic lesions (LOE 4)	Can be used for (multiple) solitary hyperkeratotic keratoses. At larger body surfaces non-surgical treatments are preferred (LOE 4)
		Add. lit.: -	Add. lit.: -	Add. lit.: -
		Grade: low positive	Grade: low positive	Grade: low positive
	Dermabrasion	There is a place for treatment with dermabrasion provided that there is enough knowledge and skill available (LOE 3)	*	Can be used (LOE 4)
		Add. lit.: 1 C study[8]	Add. lit.: -	Add. lit.: -
		Grade: low positive	-	Grade: low positive
	(Shave) Excision	No place in the treatment of actinic keratosis, unless there is doubt about the diagnosis and histology is desirable (LOE 3)	Not recommended; unless histology is desirable (LOE 4)	Can be used for (multiple) solitary hyperkeratotic keratoses. At larger body surfaces non-surgical treatments are preferred (LOE 4)
		Add. lit.: 1 C study[9]	Add. lit: -	Add. lit.: -
		Grade: high/low negative**	Grade: moderate/low negative**	Grade: low positive

Treatment		Method 1	Method 2	Method 3
Physical techniques	**Cryo-therapy** 10-14	Cryotherapy is a suitable treatment for a few solitary actinic keratoses (LOE 2)	Cryotherapy with liquid nitrogen, preferably in 2 frost-taw sessions, is a good choice (LOE 2)	For multiple, solitary and superficial keratoses cryotherapy can be chosen because of cost-effectiveness and cosmetic result (LOE 3)
		Add. lit.: 1 B study[15]	Add. lit.: -	Add. lit.: -
		Grade: moderate positive	Grade: moderate positive	Grade: low positive
	PDT (photo dynamic therapy) 13 14 16-28	PDT can be used as a treatment for actinic keratosis in case of extensive skin defects over a large area. The added value compared to simpler and cheaper treatments for solitary abnormalities has not been demonstrated (LOE 2)	A single treatment with PDT, followed by a possible second treatment for non-responders after 8-12 weeks, is a good choice; especially for patient with multiple or extensive actinic keratosis (LOE 2/3)	MAL (methyl aminolevulinate)-PDT is a suitable treatment for multiple actinic keratoses or an extensive lesion surface. Because of cost-effectiveness and practicability treatment of surfaces until 100 cm2 is recommended (LOE 2/3)
		Add. lit: 1 B study[15]	Add. lit.: -	Add. lit.: -
		Grade: low positive	Grade: moderate positive	Grade: moderate positive
	Laser 29-31	In experienced hands there is a place for CO2/Erbium-YAG laser in the treatment of actinic keratosis and actinic cheilitis. We recommend this treatment only in patients with extensive diffuse actinic keratosis, in patients resistant to other treatments, and patients with multiple malignant lesions in their history (LOE 3)	Due to the many and serious side effects laser treatment should be reserved for patients with extensive actinic keratosis located on the scalp and unresponsive to other therapies (LOE 2)	There is yet insufficient place for laser in treating actinic keratosis because of side effects and insufficient evidence for effectiveness (LOE 3)
		Add. lit.: 1 B study[32]	Add. lit: 1 B[32] and 1C[33] study	Add. lit: -
		Grade: low positive	Grade: low positive	Grade: moderate negative****

Treatment		Method 1	Method 2	Method 3
Chemical techniques	**Jessner/tri-chloric acid/ glycolic acid peel**[34-36]	A single chemical peeling with Jessner/Tri-chloric acid 35% is an effective treatment for actinic keratosis of the face. The treatment is therefore recommended as an alternative to 5-FU treatment, especially in patients who desire a single treatment and patients whose compliance is estimated to be low (LOE 3)	Jessner/Tri-chloric acid 35% is a good alternative, especially for extensive keratosis (LOE 3)	There are no benefits for chemical peeling compared to 5-FU. Therefore chemical peeling is not recommended (LOE 3)
		Add. lit.: -	Add. lit.:	Add. lit.: -
		Grade: moderate positive	Grade: moderate positive	Grade: moderate negative

Drugs	**5-FU (5-fluorouracil)** [12,17,22,34-42]	5-FU is an effective therapeutic, particularly for the treatment of a greater number of actinic keratoses (LOE 2)	5% 5-FU cream 2 times a day for 3 to 4 weeks is a good choice (LOE 1)	For large body surfaces treatment with 5-FU 1 time a day for 4 weeks is a useful choice of therapy (LOE 1/3)
		Add. lit: -	Add. lit.: -	Add.lit.: -
		Grade: moderate positive	Grade: moderate positive	Grade: moderate positive
	Imiquimod [12,38,43-52]	Imiquimod 5% is an effective therapeutic especially for the treatment of a larger number of actinic keratoses and when subclinical actinic keratoses are presumed (LOE 1)	Treatment with imiquimod has additional value, only when other treatments have failed (LOE 1/3)	Imiquimod is a good and effective alternative for patients who can not be treated with 5-FU. However 5-FU is recommended as first choice because of cost-effectiveness (LOE 1/3)
		Add. lit.: -	Add. lit: 1 A1 study[53]	Add. lit.: -
		Grade: moderate positive	Grade: low positive	Grade: moderate positive

Drugs	**Diclofenac** [47,54-59]	Diclofenac 3% in hyaluronacid 2.5% is not a first choice therapeutic. Diclofenac 3% in hyaluronacid 2,5% can be used as an non-registered alternative in patients with >1 keratoses and objections to other therapies (LOE 1)	Diclofenac is little effective and not recommended (LOE 2)	Diclofenac gel is not recommended (LOE 1)
		Add. lit: -	Add. lit.: -	Add. lit: -
		Grade: low negative/low positive***	Grade: moderate negative	Grade: moderate negative
	Retinoids [60-62]	The treatment of actinic keratosis with topical retinoids is not a first choice alternative in view of the varying effectiveness and side effects reported. Treatment with systemic retinoids is not routinely recommended and should be reserved as an alternative for patients at high risk of developing actinic keratosis and other skin malignancies (LOE 3)	There is no place for local retinoids (LOE 2/3)	Primary treatment with retinoids is not recommended (LOE 3)
		Add. lit: -	Add. lit: -	Add. lit: 1 B study[63]
		Grade: low negative/low positive***	Grade: high negative	Grade: moderate negative
	Colchicine [64,65]	The various other therapies are not recommended because of a lack in clear studies and variable efficacy data. Further randomized and prospective studies are desirable (LOE 2)	Further research to compare local colchicine with the current standard therapies is desirable (LOE 3)	There is yet no place for colchicine cream (LOE 2/3)
		Add. lit.: -	Add. lit.: -	Add. lit.: -
		Grade: moderate negative	Grade: moderate negative	Grade: high negative****

Treatment		Method 1	Method 2	Method 3
Drugs	**Masoprocol** [66]	The various other therapies are not recommended because of a lack in clear studies and variable efficacy data. Further randomized and prospective studies are desirable (LOE 3)	Masoprocol has no place in the treatment of actinic keratosis, because it is not available in the Netherlands and there is only 1 study with a limited follow-up period (LOE 3)	There is yet no place for masoprocol (LOE 3)
		Add. lit.: -	Add. lit: -	Add lit.: -
		Grade: moderate negative	Grade: high negative	Grade: high negative****
	Others [67,68]	The various other therapies are not recommended because of a lack in clear studies and variable efficacy data. Further randomized and prospective studies are desirable (LOE 3)	*	*
		Add. lit.: -	Add. lit.: -	Add. lit.: -
		Grade: moderate negative	-	-

Table 4: Recommendations of the three different guideline methods per treatment. Additional literature used per treatment and method. Grading of perceived obligation to follow a recommendation per treatment and method. * = no recommendation formulated. ** = two grades to underscore the weight of the word 'unless'. *** = two grades because of the bivalent word choice. **** = the word 'yet' in the recommendation was ignored to reflect the status of the recommendation at the time the draft guideline was made. Add. lit. = additional literature. LOE = level of evidence.

RESULTS

The three guideline development methods gave the results depicted in table 4 per treatment. The level of evidence behind the recommendation differed for several treatments: dermabrasion, (shave) excision, cryotherapy, photodynamic therapy (PDT), laser, 5-FU (5-fluorouracil), diclofenac, retinoids and colchicine. However, the correlations between the appointed levels of evidence remained highly significant ($p<0.003$ and $p<0.001$) (table 5). The members of method 1 used five additional references, method 2 used three and method 3 used one additional reference (Table 4). Regarding the grading of recommendations, differences across the three methods were present in the recommendations for (shave) excision, cryotherapy, PDT, laser, chemical peeling, diclofenac, imiquimod, retinoids, colchicine and masoprocol (Table 4). The correlation coefficient between method 1 and method 2 was highly significant at 0.755 ($p<0.001$). There was a significant, but lower, correlation coefficient between method 1 and 3 and between method 2 and 3 (respectively 0.493 ($p<0.026$) and 0.673 ($p<0.007$) (Table 5). With regard to the total costs and time period necessary to prepare a draft guideline, method 3 was the cheapest and quickest (€24 770 and 4months) and method 1 was the most expensive and slowest (€48 100 and 14 months) (Table 6).

Correlations		Intraclass Correlation Coefficient
Level of evidence	Method 1 versus Method 2	.735 (95% CI 0.300-0.916) p = 0.003**
	Method 1 versus Method 3	.826 (95% CI 0.534-0.943) p < 0.001**
	Method 2 versus Method 3	.859 (95% CI 0.585-0.957) p < 0.001**
Recommendations	Method1 versus Method 2	.755 (95% CI 0.372-0.922) p = 0.001**
	Method 1 versus Method 3	.493 (95% CI 0.001-0.806) p = 0.026*
	Method 2 versus Method 3	.673 (95% CI 0.199-0.893) p = 0.007**

Table 5: Correlation between the levels of evidence and recommendation grades of the three different guideline methods.

* Intraclass correlation is significant between the 0.05 and 0.01 level (2-way).

** Intraclass correlation is significant at least at the 0.01 level (2-way).

95% CI = 95 % confidence interval.

	Method 1	Method 2	Method 3
Total costs (euro)*	€ 48,100	€ 39,630	€ 24,770
Time period necessary for draft guideline**	14 months	5 months	4 months

Table 6: Costs and time necessary for the draft guidelines.
* The costs for the preparatory work needed for all the guideline methods was added to each method. ** Including the preparatory work; exclusion of the months waiting for the external approval to start making the guideline.

DISCUSSION

A comparison of the three guideline development methods showed a highly significant correlation between the recommendations formed by the group that followed a traditional evidence-based approach with multiple sessions and one that followed a 2-day conference method, despite the differences in formulation. The third method consisting of a systematic evidence report and recommendations made by a single dermatologist was also significantly correlated with the other methods; although significance and correlation coefficient were lower (Table 5). When considering the development process, there are various benefits and disadvantages in the characteristics of the three methods. The benefit of the standard multiple meetings method (method 1) was that there was time to adapt and reconsider texts and members could consult colleagues and experts between guideline development sessions; probably enhancing support for implementation of the final guideline.[72,73] However, more sessions, time and costs were required (Table 6) and the progress was often dependent on the effort of single group members. The conference method 2 was found to be highly satisfactory by its members, partly because of the direct discussion possible, and was time- and cost-efficient. Time, however, was also a crucial factor since a quality guideline needed to be prepared in just 2 days. Furthermore, the quality of preparatory work was essential. Still, preparatory work was perhaps more controllable due to the uniform and transparent way of searching, selecting and extracting data. A disadvantage could be the large amount of preparatory work needed, so correct scheduling is necessary. Method 3 is quick and relatively inexpensive (table 5), but potentially more susceptible to interpretative bias as it depends on the expertise of a single dermatologist. In addition, the field acceptance of method 3 is highly debatable.[74] Moreover, an epidemiologist often has no clinical experience regarding the topic. On the other hand, a uniform review of the evidence by a single epidemiologist could reduce variation in the appointed level of evidence, because the comparison showed several differences. Additional literature

could explain a part of the variation in the level of evidence for dermabrasion, excision, cryotherapy, PDT and laser therapy.
In four out of eight cases, the additional literature was distilled from existing guidelines.[75-77] The other literature, not identified by the systematic literature search, was brought in by members of the working groups; a benefit for a group method over a single dermatologist. Apart from differences in assessing the underlying literature, variation in recommendations is probably caused by different considerations for clinical practice beyond the evidence (Table 7).

Limitations

Although the working group of method 1 was blinded about the existence of other methods, the working group of method 2 knew that they were part of a comparison as soon as the topic was revealed at the beginning of the 2-day meeting. Also, the members participating in method 3 knew that they were part of a study. This could have resulted in the so-called Hawthorne effect by being part of a study one feels special and puts in more effort to make something special of it. Positive results could therefore be overvalued.[78] Moreover, the individual members of the working groups are a likely source of bias. By matching several characteristics (Table 1) we tried to enhance comparability, as randomisation was not feasible. Still individual opinions and expertise vary and individual group members can play an important role in setting subtle textual differences in a guideline, which emphasises the importance of a rigorous way to develop consensus and address considerations beyond the evidence.[79] Actinic keratosis is a well-described skin disease with a relatively small amount of published studies (some 100 therapeutic trials). Thus, one could say that actinic keratosis is most suitable for guideline development in one meeting, because a draft guideline can be written in a few days. Perhaps for a more intensively trialled, multidisciplinary or controversial guideline topic, a combination of methods 1, 2 and 3 will be the most suitable. Particularly, the combination of an externally generated, methodological conscientious evidence report with subsequent analysis and translation into clinical recommendations in a meeting of clinicians would be interesting for further research.
Another limitation is the way recommendations are being compared and formulated. There are various systems for grading the strength of a recommendation.[80] We graded terms in a recommendation according to the perceived level of obligation to follow that so-called deontic term.[70] Even if perceived obligation to follow a recommendation is a practical measure, recommendations often differ in a more subtle way. For instance, a more specific recommendation with information about

dosage is more likely to be followed than a vague, broader recommendation.[73] Major steps in setting a high quality standard in how to formulate recommendations in clinical practice guideline have been made with the introduction of the Appraisal of Guidelines Research & Evaluation (AGREE)instrument,[81] the formation of (inter)national platforms (such as the Guideline International Network)[82] and in promoting transparent guideline development and implementation methods (Grading of Recommendations Assessment, Development and Evaluation (GRADE), Guideline Implementability Appraisal (GLIA)).[83-85] However, there is no such thing as a golden standard and much remains to be elucidated. Adoption of a single guideline system in the future can make comparison across guidelines easier. Also, adaptation of already existing guidelines (collaboration for ADAPTation of existing guidelines (ADAPTE)), another way to potentially reduce costs and effort, would be simpler.[86] Matching deontic terms to grades of recommendation strength can help to further standardise the use of such expressions by guideline developers.[70]

CONCLUSION

The therapeutic recommendations and appointed levels of evidence for the treatment of actinic keratosis differed on several components when comparing three different guideline development methods. Despite the differences, the methods rendered a significant correlation compared with each other. Consequently, the value of a guideline using a 2-day conference development method, a method in which the preparatory work is carried out by an epidemiologist and finished by a single dermatologist or a combination of these two seems to at least equal that of a guideline composed in multiple sessions, that is for topics with a monodisciplinary character and a relatively small number of conducted trials.
In addition, the presented alternative methods were more time- and cost efficient. The equality in guideline quality remains questionable for a more intensively trialled or controversial guideline topic. Further research and experience on this matter is warranted.

Treatment		Method 1	Method 2	Method 3
Surgical techniques	**Curettage (including supplementary electro-dessiction)**	Advantages: -histological diagnosis possible (although less suitable for differentiation between actinic keratosis and squamous cell carcinoma) -easy to use -cheap Disadvantages: -impaired wound healing, scarring and pigmentation changes possible -only suitable for a few keratoses	-effective and quick for solitary keratoses; especially when histological examination is needed -acceptable to good cosmetic results	-not commonly used in daily practice -painful; local anesthesia is usually necessary -risk of infection and other impaired wound healing -other methods are less destructive and more patient friendly
	Dermabrasion	**	-effective and quick for solitary keratoses; especially when histological examination is needed -acceptable to good cosmetic results	-not commonly used in daily practice -painful; local anesthesia is usually necessary -risk of infection and other impaired wound healing -other methods are less destructive and more patient friendly
	(Shave) Excision	Advantages: -definitive histological diagnosis -high curation rates Disadvantages: -impaired wound healing possible -scarring -only suitable for a few keratoses	-effective and quick for solitary keratoses; especially when histological examination is needed -acceptable to good cosmetic results	-not commonly used in daily practice -painful; local anesthesia is usually necessary -risk of infection and other impaired wound healing -other methods are less destructive and more patient friendly

Treatment		Method 1	Method 2	Method 3
Physical techniques	**Cryo-therapy**	Advantages: -easy to use -cheap -high compliance -low treatment burden for patients -high curation rate Disadvantages: -variability in used treatment technique -risk of scarring -risk of changes in pigmentation -only for solitary keratoses -painful during and direct after treatment	-cheap -easy to use -suitable for solitary or a couple of lesions -therapy not standardized -cosmetic result depends heavily on location of the lesion	-commonly used in daily practice -cheap -easy to use -acceptable cosmetic results -low treatment burden for patients -no standardization of contact time
	PDT (photo dynamic therapy)	Advantages: -high curation rates -no scarring -treatment of a large amount of actinic keratoses possible -excellent cosmetic results -treatment of non-visible keratosis -short treatment interval -non-compliance is not an issue Disadvantages: -often painful -erythema and crustae for 10 to 14 days -relatively expensive -availability of the equipment	-expensive treatment -time consuming; high treatment burden for patients -good cosmetic results	-not significant more effective than cryotherapy or 5-FU -good cosmetic results -low treatment burden for patients -expensive -more time consuming for patients than 5-FU or cryotherapy -large body surfaces are more unfeasible; maximum treatable body surface is 200 cm2

Treatment		Method 1	Method 2	Method 3
Physical techniques	**Laser**	Advantages: -easy to use -high curation rates -suitable for treating extensive actinic keratosis Disadvantages: -local anesthesia necessary -painful -impaired wound healing, scarring and pigmentation changes possible -availability of the equipment -relatively expensive	-effective -expensive -high treatment burden for patients -not available in every clinic -local anesthesia necessary	-remarkable side-effects -effectiveness/efficacy uncertain -expensive
Chemical techniques	**Jessner/tri-chloric acid/ glycolic acid peel**	Advantages: -cheap -large body surfaces are treatable -high compliance (treatment in clinic) Disadvantages: -effect depends on the level of experience with the treatment -painful (acute) -impaired wound healing and pigmentation changes possible	-effective -relatively expensive -time consuming for patients -sometimes high treatment burden for patients -usually good cosmetic result	-there are no advantages of chemical peeling over 5-FU when treating actinic keratosis

Treatment		Method 1	Method 2	Method 3
Drugs	**5-FU (5-fluorouracil)**	Advantages: -cheap -no scarring -treatment of a large amount of actinic keratoses possible -treatment of non-visible keratosis Disadvantages: -high compliance necessary -irritating substance -allergic reaction possible -painful -effect depend on the level of penetration in the skin -extensive treatment duration	-suitable for extensive and/or a large amount of actinic keratosis -treatment can be carried out at home -relatively extensive treatment duration -side-effects; irritation, secondary infections of the skin, changes in pigmentation -cosmetic results acceptable to good -only 5-FU 5% is available in the Netherlands	-easy to use -treatment of large body surfaces possible -cheap -high compliance necessary -side-effects: irritation and crustae
	Imiquimod	Advantages: -treatment of a large amount of actinic keratoses is possible -treatment of non-visible lesions -eventually good cosmetic results -better results than with 5-FU, diclofenac gel and cryotherapy possible. Disadvantages: -high compliance necessary -second treatment of 4 weeks may be necessary -many local side-effects -extensive treatment duration and expensive -long treatment interval (3 weeks)	-effective -relatively expensive -sometimes high treatment burden for patients -cosmetic results acceptable to good	-effective; superiority imiquimod over 5-FU unclear -treatment of large body surfaces possible -side-effects

Treatment		Method 1	Method 2	Method 3
Drugs	**Diclofenac**	Advantages: -no scarring -treatment of large body surfaces possible -treatment of non-visible lesions Disadvantages: -high compliance necessary -irritative -allergic reaction possible -50% non-responder -not registered in the Netherlands	-available over the counter -few side-effects -patients can use it at home	-highly variable effectiveness/efficacy (never above 50%) -effect only after extensive treatment
	Retinoids	-variable effectiveness/efficacy -variable side-effects	-no sufficient studies regarding oral retinoids	-oral etretinate and topical isotretinoin are not registered in the Netherlands -retinoids seem to only influence the clinical characteristics (keratinisation) and have no effect on the malignant transformation risk.
	Colchicine	-no unambiguous studies -variable efficacy/effectiveness	-local colchicine is not available in the Netherlands	-local colchicine is not available in the Netherlands -studies with small patient groups and potential bias; effectiveness/efficacy therefore not proven

Drugs	**Masoprocol**	-no unambiguous studies -variable efficacy/effectiveness	-only one study available -limited follow-up -not available in the Netherlands	-not available in the Netherlands -effectiveness/efficacy not proven
	Others	-no unambiguous studies -variable efficacy/effectiveness	*/**	*/**

Table 7: Considerations, other than the conclusion from scientific literature, in making the clinical recommendations of the three different guideline methods. * = no recommendation formulated. ** = no considerations given.

REFERENCES

1 Lebwohl M. Actinic keratosis: epidemiology and progression to squamous cell carcinoma. Br J Dermatol 2003;149:31e3.

2 Holmes C, Foley P, Freeman M, et al. Solar keratosis: epidemiology, pathogenesis, presentation and treatment. Australas J Dermatol 2007;48:67e74.

3 Salasche SJ. Epidemiology of actinic keratoses and squamous cell carcinoma. J Am Acad Dermatol 2000;42:4e7.

4 Dutch Clinical Practice Guideline Actinic Keratosis. http://www.huidarts.info/documents/?v¼2&id¼193 (accessed 24 Nov 2010).

5 Grimshaw J, Russel I. Achieving health gain through clinical guidelines: current activities and future directions. Ann Intern Med 1990;113:709e14.

6 Burgers JS, Grol R, Klazinga NS, et al. Towards evidence-based clinical practice: an international survey of 18 clinical guideline programs. Int J Qual Health Care 2003;15:31e45.

7 Quaedvlieg PFJ, Ostertag JU, de Kort JML, et al. Evidence based treatments of actinic keratoses: a systematic review of efficacy. J Dermatol Treat. In press.

8 Coleman WP III, Yarborough JM, Mandy SH. Dermabrasion for prophylaxis and treatment of actinic keratoses. J Dermatol Surg 1996;22:17e21.

9 Emmett AJ, Broadbent GD. Shave excision of superficial solar skin lesions. Plast Reconstr Surg 1987;80:47e54.

10 Zouboulis CC, Rohrs H. Kryochirurgische behandlung aktinischer keratosen mit evidenzbasierter ubersicht. Hautarzt 2006;56:353e8.

11 Morton C, Campbell S, Gupta G, et al. Intraindividual, right-left comparison of topical methyl aminolaevulinate-photodynamic therapy and cryotherapy in subjects with actinic keratoses: a multicentre, randomized controlled study. Br J Dermatol 2006;155:1029e36.

12 Krawtchenko N, Roewert HJ, Ulrich M, et al. A randomised study of topical 5% imiquimod vs. topical 5-fluorouracil vs. cryosurgery in immunocompetent patients with actinic keratoses: a comparison of clinical and histological outcomes including 1-year follow-up. Br J of Dermatol 2007;157(Suppl 2):34e40.

13 Szeimies RM, Karrer S, Radakovic-Fijan S, et al. Photodynamic therapy using topical methyl 5-aminolevulinate compared with cryotherapy for actinic keratosis: a prospective, randomized study. J Am Acad Dermatol 2002;47:258e62.

14 Freeman M, Vinciullo C, Francis D, et al. A comparison of photodynamic therapy using topical methyl aminolevulinate (Metvix) with single cycle cryotherapy in patients with actinic keratosis: a prospective, randomized study. J Dermatolog Treat 2003;14:99e106.

15 Szeimies RM, Radakovic S, Calzavara-Pinton PG, et al. A prospective, randomized study comparing photodynamic therapy with Metvix to cryotherapy in actinic keratoses. J Eur Acad Dermatol Venereol 2000;14:235.

16 Dragieva G, Prinz BM, Hafner J, et al. A randomized controlled clinical trial of topical photodynamic therapy with methyl aminolaevulinate in the treatment of actinic keratoses in transplant recipients. Br J Dermatol 2004;151:196e200.

17 Gilbert DJ. Treatment of actinic keratoses with sequential combination of 5-fluorouracil and photodynamic therapy. J Drugs Dermatol 2005;4:161e3.

18 de Haas ER, de Vijlder HC, Sterenborg HJ, et al. Fractionated aminolevulinic acid-photodynamic therapy provides additional evidence for the use of PDT for non-melanoma skin cancer. J Eur Acad Dermatol Venereol 2008;22:426e30.

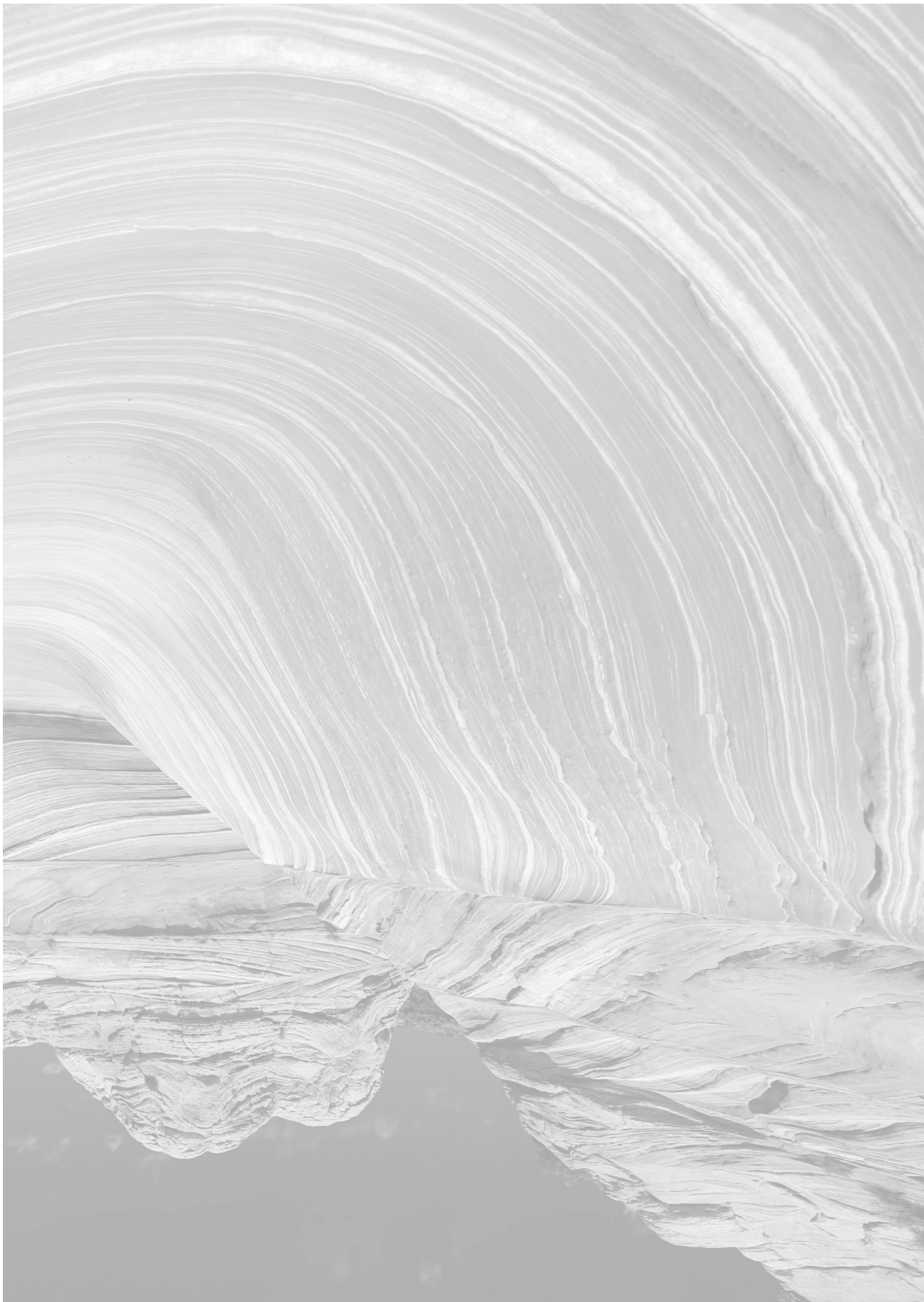
19 Jeffes EW, McCullough JL, Weinstein GD, et al. Photodynamic therapy of actinic keratoses with topical aminolevulinic acid hydrochloride and fluorescent blue light. J Am Acad Dermatol 2001;45:96e104.

20 Pariser DM, Lowe NJ, Stewart DM, et al. Photodynamic therapy with topical methyl aminolevulinate for actinic keratosis: results of a prospective randomized multicenter trial. J Am Acad Dermatol 2003;48:227e32.
21 Piacquadio DJ, Chen DM, Farber HF, et al. Photodynamic therapy with aminolevulinic acid topical solution and visible blue light in the treatment of multiple actinic keratoses of the face and scalp: investigator-blinded, phase 3, multicenter trials. Arch Dermatol 2004;140:41e6.
22 Smith S, Piacquadio D, Morhenn V, et al. Short incubation PDT versus 5-FU in treating actinic keratoses. J Drugs Dermatol 2003;2:629e35.
23 Tarstedt M, Rosdahl I, Berne B, et al. A randomized multicenter study to compare two treatment regimens of topical methyl aminolevulinate (Metvix)-PDT in actinic keratosis of the face and scalp. Acta Derm Venereol 2005;85:424e8.
24 Touma D, Yaar M, Whitehead S, et al. A trial of short incubation, broad-area photodynamic therapy for facial actinic keratoses and diffuse photodamage. Arch Dermatol 2004;140:33e40.
25 Tschen EH, Wong DS, Pariser DM, et al. Photodynamic therapy using aminolaevulinic acid for patients with nonhyperkeratotic actinic keratoses of the face and scalp: phase IV multicentre clinical trial with 12-month follow up. Br J Dermatol 2006;155:1262e9.
26 Varma S, Wilson H, Kurwa HA, et al. Bowen's disease, solar keratoses and superficial basal cell carcinomas treated by photodynamic therapy using a large-field incoherent light source. Br J Dermatol 2001;144:567e74.
27 Wiegell SR, Haedersdal M, Philipsen PA, et al. Continuous activation of PpIX by daylight is as effective as and less painful than conventional photodynamic therapy for actinic keratoses; a randomized, controlled, single-blinded study. Br J Dermatol 2008;158:740e6.
28 Zelickson B, Counters J, Coles C, et al. Light patch: preliminary report of a novel form of blue light delivery for the treatment of actinic keratosis. Dermatol Surg 2005;31:375e8.
29 Iyer S, Friedli A, Bowes L, et al. Full face laser resurfacing: therapy and prophylaxis for actinic keratoses and non-melanoma skin cancer. Lasers Surg Med 2004;34:114e19.
30 Jiang SB, Levine VJ, Nehal KS, et al. Er:YAG laser for the treatment of actinic keratoses. Dermatol Surg 2000;26:437e40.
31 Wollina U, Konrad H, Karamfilov T. Treatment of common warts and actinic keratoses by Er:YAG laser. J Cutan Laser Ther 2001;3:63e6.
32 Ostertag JU, Quaedvlieg PJF, van der Geer S, et al. A clinical comparison and long term follow up of topical 5-FU versus laser resurfacing in the treatment of widespread actinic keratoses. Lasers Surg Med 2006;38:731e9.
33 Ostertag JU, Quaedvlieg PJ, Neumann MH, et al. Recurrence rates and long-term follow-up after laser resurfacing as a treatment for widespread actinic keratoses on the face and scalp. Dermatol Surg 2006;32:261e7.
34 Lawrence N, Cox SE, Cockerell CJ, et al. A comparison of the efficacy and safety of Jessner's solution and 35% trichloroacetic acid vs 5% fluorouracil in the treatment of widespread facial actinic keratoses. Arch Dermatol 1995;131:176e81.
35 Marrero GM, Katz BE. The new fluor-hydroxy pulse peel. A combination of 5-fluorouracil and glycolic acid. Dermatol Surg 1998;24:973e8.
36 Witheiler DD, Lawrence N, Cox SE, et al. Long-term efficacy and safety of Jessner's solution and 35% trichloroacetic acid vs 5% fluorouracil in the treatment of widespread facial actinic keratoses. Dermatol Surg 1997;23:191e6.
37 Epstein E. Does intermittent "pulse" topical 5-fluorouracil therapy allow destruction of actinic keratoses without significant inflammation? J Am Acad Dermatol 1998;38:77e80.

38 Gupta AK, Davey V, Mcphail H. Evaluation of the effectiveness of imiquimod and 5-fluorouracil for the treatment of actinic keratosis: Critical review and meta-analysis of efficacy studies. J Cutan Med Surg. 2005;9:209e14.
39 Jorizzo J, Stewart D, Bucko A, et al. Randomized trial evaluating a new 0.5% fluorouracil formulation demonstrates efficacy after 1-, 2-, or 4-week treatment in patients with actinic keratosis. Cutis 2002;70:335e9.
40 Jorizzo J, Weiss J, Furst K, et al. Effect of a 1-week treatment with 0.5% topical fluorouracil on occurrence of actinic keratosis after cryosurgery: a randomized, vehicle-controlled clinical trial. Arch Dermatol 2004;140:813e16.
41 Loven K, Stein L, Furst K, et al. Evaluation of the efficacy and tolerability of 0.5% fluorouracil cream and 5% fluorouracil cream applied to each side of the face in patients with actinic keratosis. Clin Ther 2002;24:990e1000.
42 Weiss J, Menter A, Hevia O, et al. Effective treatment of actinic keratosis with 0.5% fluorouracil cream for 1, 2, or 4 weeks. Cutis 2002;70:22e9.
43 Alomar A, Bichel J, McRae S. Vehicle-controlled, randomized, double-blind study to assess safety and efficacy of imiquimod 5% cream applied once daily 3 days per week in one or two courses of treatment of actinic keratoses on the head. Br J Dermatol 2007;157:133e41.
44 Chen K, Yap LM, Marks R, et al. Short-course therapy with imiquimod 5% cream for solar keratoses: a randomized controlled trial. Australas J Dermatol 2003;44:250e5.
45 Falagas ME, Angelousi AG, Peppas G. Imiquimod for the treatment of actinic keratosis: a meta-analysis of randomized controlled trials. J Am Acad Dermatol 2006;55:537e8.
46 Korman N, Moy R, Ling M, et al. Dosing with 5% imiquimod cream 3 times per week for the treatment of actinic keratosis: results of two phase 3, randomized, double-blind, parallel-group, vehicle-controlled trials. Arch Dermatol 2005;141:467e73.
47 Kose O, Koc E, Erbil AH, et al. Comparison of the efficacy and tolerability of 3% diclofenac sodium gel and 5% imiquimod cream in the treatment of actinic keratosis. J Dermatolog Treat 2008;19:159e63.
48 Lebwohl M, Dinehart S, Whiting D, et al. Imiquimod 5% cream for the treatment of actinic keratosis: results from two phase III, randomized, double-blind, parallel group, vehicle-controlled trials. J Am Acad Dermatol 2004;50:714e21.
49 Stockfleth E, Meyer T, Benninghoff B, et al. A randomized, doubleblind, vehicle-controlled study to assess 5% imiquimod cream for the treatment of multiple actinic keratoses. Arch Dermatol 2002;138:1498e502.
50 Szeimies RM, Gerritsen MJ, Gupta G, et al. Imiquimod 5% cream for the treatment of actinic keratosis: results from a phase III, randomized, double-blind, vehicle-controlled, clinical trial with histology. J Am Acad Dermatol 2004;51:547e55.
51 Torres A, Storey L, Anders M, et al. Microarray analysis of aberrant gene expression in actinic keratosis: effect of the Toll-like receptor-7 agonist imiquimod. Br J Dermatol 2007;157:1132e47.
52 Vender RB, Goldberg O. Innovative uses of imiquimod. J Drugs Dermatol 2005;4:58e63.
53 Hadley G, Derry S, Moore RA. Imiquimod for actinic keratosis: systematic review and meta-analysis. J Invest Dermatol 2006;126:1251e5.
54 Fariba I, Ali A, Hossein SA, et al. Efficacy of 3% diclofenac gel for the treatment of actinic keratoses: a randomized, double-blind, placebo controlled study. Indian J Dermatol Venereol Leprol 2006;72:346e9.
55 Gebauer K, Brown P, Varigos G. Topical diclofenac in hyaluronan gel for the treatment of solar keratoses. Australas J Dermatol 2003;44:40e3.
56 McEwan LE, Smith JG. Topical diclofenac/hyaluronic acid gel in the treatment of solar keratoses. Australas J Dermatol 1997;38:187e9.

57 Rivers JK, Arlette J, Shear N, et al. Topical treatment of actinic keratoses with 3.0% diclofenac in 2.5% hyaluronan gel. Br J Dermatol 2002;146:94e100.
58 Smith SR, Morhenn VB, Piacquadio DJ. Bilateral comparison of the efficacy and tolerability of 3% diclofenac sodium gel and 5% 5-fluorouracil cream in the treatment of actinic keratoses of the face and scalp. J Drugs Dermatol 2006;5:156e9.
59 Wolf JE Jr, Taylor JR, Tschen E, et al. Topical 3.0% diclofenac in 2.5% hyaluronan gel in the treatment of actinic keratoses. Int J Dermatol 2001;40:709e13.
60 Alirezai M, Dupuy P, Amblard P, et al. Clinical evaluation of topical isotretinoin in the treatment of actinic keratoses. J Am Acad Dermatol 1994;30:447e51.
61 Kang S, Goldfarb MT, Weiss JS, et al. Assessment of adapalene gel for the treatment of actinic keratoses and lentigines: a randomized trial. J Am Acad Dermatol 2003;49:83e9.
62 Moriarty M, Dunn J, Darragh A, et al. Etretinate in treatment of actinic keratosis. A double-blind crossover study. Lancet 1982;1:364e5.
63 de Seveaux RG, Smit JV, de Jong EM, et al. Acitretin treatment of premalignant and malignant skin disorders in renal transplant recipients: clinical effects of a randomized trial comparing two doses of acitretin. J Am Acad Dermatol 2003;49:407e12.
64 Akar A, Bulent Tas, tan H, Erbil H, et al. Efficacy and safety assessment of 0.5% and 1% colchicine cream in the treatment of actinic keratoses. J Dermatol Treat 2001;12:199e203.
65 Grimaitre M, Etienne A, Fathi M, et al. Topical colchicine therapy for actinic keratoses. Dermatology 2000;200:346e8.
66 Olsen EA, Abernethy ML, Kulp-Shorten C, et al. A double-blind, vehicle-controlled study evaluating masoprocol cream in the treatment of actinic keratoses on the head and neck. J Am Acad Dermatol 1991;24:738e43.
67 Alberts DS, Dorr RT, Einspahr JG, et al. Chemoprevention of human actinic keratoses by topical 2-(difluoromethyl)-dl-ornithine. Cancer Epidemiol Biomarkers Prev 2000;9:1281e6.
68 Cafardi JA, Elmets CA. T4 endonuclease V: review and application to dermatology. Expert Opin Biol Ther 2008;8:829e38.
69 Quaedvlieg PJ, Tirsi E, Thissen MR, et al. Actinic keratosis: how to differentiate the good from the bad ones? Eur J Dermatol 2006;16:335e9.
70 Lomotan EA, Michel G, Lin Z, et al. How “should” we write guideline recommendations? Interpretation of deontic terminology in clinical practice guidelines: survey of the health services community. Qual Saf Health Care. Published Online First: 10 August 2010. doi:10.1136/qshc.2009.032565.
71 Dutch CBO Guideline Manual. http://www.cbo.nl/thema/Richtlijnen/EBRO-handleiding/6-Schrijven-van-teksten/ (accessed 24 Nov2010).
72 Wakkee M, Lugtenberg M, Spuls PI, et al. Knowledge, attitudes and use of the guidelines for the treatment of moderate to severe plaque psoriasis among Dutch dermatologists. Br J Dermatol 2008;159:426e32.
73 Cabana MD, Rand CS, Powe NR, et al. Why don’t physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999;282:1458e65.
74 Grol R, Grimshaw J. Evidence-based implementation of evidence-based medicine. Jt Comm J Qual Improv 1999;25:503e10.
75 Berker de D, McGregor JM, Hughes BR, et al. Guidelines for the management of actinic keratoses. Br J Dermatol 2007;156:222e30.
76 Braathen LD, Szeimies RM, Basset-Seguin N, et al. Guidelines on the use of photodynamic therapy for nonmelanoma skin cancer: An international consensus. J Am Acad Dermatol 2007;56:125e43.

77 Stockfleth E, Kerl H. Guidelines for the management of actinic keratoses. Eur J Dermatol 2006;16:599e606.
78 Franke RH, Kaul JD. The Hawthorne experiments: first statistical interpretation. Am Sociol Rev 1978;43:623e43.
79 Raine R, Sanderson C, Black N. Developing clinical guidelines: a challenge to current methods. BMJ 2005;331:631e3.
80 West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence. Evid Rep Technol Assess (Summ) 2002;47:1e11.
81 Cluzeau FA, Burgers JS, Brouwers M, et al. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 2003;12:18e23.
82 http://www.g-i-n.net/ (accessed 24 Nov 2010).
83 Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924e6.
84 Guyatt GH, Oxman AD, Kunz R, et al. Going from evidence to recommendations. BMJ 2008;336:1049e51. Erratum in: BMJ 2008;336(7658).
85 Shiffman RN, Dixon J, Brandt C, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak 2005;5:23.
86 Fervers B, Burgers JS, Voellinger R, et al. Guideline adaptation: an approach to enhance efficiency in guideline development and improve utilisation. BMJ Qual Saf 2011;20:228e36.



6

GUIDELINE-BASED CLINICAL DECISION SUPPORT

Guideline-based clinical decision support in acne patients receiving isotretinoin; improving adherence and cost-effectiveness

Published as
Borgonjen RJ, de Lange JA, van de Kerkhof PCM.
Guideline-based clinical decision support in acne patients receiving isotretinoin; improving adherence and cost-effectiveness.
J Eur Acad Dermatol Venereol. 2017;31(10):e440-e442

SUMMARY

Background clinical practice guidelines are intented to improve quality of care. However guidelines are often published as long texts and implementation in daily practice is challenging.

Objectives testing the impact of active and passive guideline based clinical decision support (CDS) on laboratory monitoring of acne patients receiving isotretinoin. Aiming to potentially increase guideline adherence and achieve a reduction in clinical practice variation and costs.

Methods all acne patients receiving isotretinoin in the study center were included prospectively, 51 during the 5 months before the CDS introduction and 43 patients in the 5 months after the CDS introduction. We compared adherence to guideline recommended laboratory testing intervals (passive support) and tests (active support) before and after CDS using chi square tests.

Results guideline adherence to laboratory testing intervals was 57% before CDS and 70% after CDS (p = 0.197). Adherence to the recommended tests was 0% before CDS and 63% after (p = <0,0005). Costs of laboratory testing per patient were reduced more than 2.5 fold.

Conclusion Use of guideline based CDS appears to improve adherence to recommended laboratory tests in acne patients receiving isotretinoin and reduces overtesting. Timely laboratory test intervals increased non significantly. Hence, active support was more effective than passive support. Passive CDS remains subject to barriers preventing implementation.

INTRODUCTION

Clinical practice guidelines (CPG) intend to assist practitioners about appropriate healthcare for specific clinical circumstances.[1] They have proven to be a tool for condensing the overwhelming quantity of (new) medical knowledge into guidance for clinical practice. A CPG can therefore improve quality of care, but there are barriers to implementation and adherence such as a lack of awareness and inertia of previous practice.[2,3]

There are also external barriers relating to the guideline itself. Often CPGs are still extensive texts, that can be difficult and cumbersome to read and do not support healthcare practitioners in finding patient-specific recommendations.[2,4] Those barriers stimulated the development of clinical decision support (CDS) systems, which could ultimately be incorporated into electronic health records (EHR). CDS uses passive or active support to modify clinician behavior through recommendations of specific actions. Not the healthcare professional needs to decide whether consultation of a guideline is necessary, but support is directly and automatically delivered in daily practice.[5,6,7]

Implementing guidelines with the help of CDS might improve its acceptance and application. Several studies showed the possibility to reduce costs and increase clinician adherence to CPGs.[8,9] However, there are mixed results, especially regarding the effects of CDS on timely test ordering.[10,11,12]

To achieve proper care for acne patients, extensive guidelines have been created by various dermatological associations.[13,14] To-date, no CDS was used to implement them. The European and Dutch acne guidelines state that laboratory testing is advised according to a certain schedule when prescribing isotretinoin (Fig.1); a drug of choice in many forms of acne.[13,14] A review and a questionnaire carried out during the 2014 annual meeting of the American Academy of Dermatology showed variation in laboratory testing.[15] As an example, of the 2343 dermatologists who completed the questionnaire 33% ordered monthly complete blood count testing. We want to present the results of a prospective study embedding the acne guidelines on laboratory testing to remind physicians about guideline recommendations in their daily workflow. Adherence to the recommended laboratory testing in acne patients receiving isotretinoin may lead to a reduction of practice variability and costs, while minimizing the patient burden of laboratory testing associated with isotretinoin treatment.

METHODS

Design

A prospective pre- versus post-intervention comparison design was used. Guideline based CDS was embedded in the laboratory order form in the EHR system. When clinicians order laboratory testing after introduction of the guideline based CDS, the support appears. It uses both active components (recommended laboratory tests are already ticked) as well as passive components (repeating the textual guideline recommendation (Fig.1)). Intervention goal is improvement in recommended laboratory testing adherence, both in frequency (passive support) as in the tests used to monitor the treatment (active support).

Data

Data were extracted from the EHR system of the study center (Radboud university medical center).

Subjects

All laboratory orders in acne patients receiving isotretinoin between October 15th 2015 and August 15th 2016 (5 months before and after intervention) were included. Every dermatologist (in training) was counted in. Date of intervention was March 15th 2016 (Fig.1). Patients receiving isotretinoin for an (off-label) indication other than acne were excluded.

Outcomes

The primary outcome was guideline adherence, defined as laboratory testing after 0, 1 and 4 months of isotretinoin therapy and every 3 months thereafter. Shorter or longer intervals were defined as non-adherence. We allowed a window of grace of 7 days and 14 days for recommended testing intervals after 1 month and every 3 months respectively.

The following laboratory tests are recommended: liver enzymes (aspartate aminotransferase and/or alanine aminotransferase) and lipids (triglyceride, low-density lipoprotein, high-density lipoprotein). Any additional tests were regarded as non-adherence. Reasons for non-adherence were analyzed since not every non-adherence is automatically overtesting and more specific tests could be necessary depending on clinical factors (e.g. co-morbidities and/or co-medication).

Secondary outcomes were total cost reduction and number of non-recommended extra laboratory tests and visits per patient before and after the intervention.

Laboratory testing costs were calculated using the 2016 laboratory pricelist of the study center.

Statistics

Data collection and analysis was performed using Statistical Package for Social Sciences (SPSS®) for Windows, version 22.0. Categorical variables were expressed as numbers and percentages. P-values of <.05 were considered statistically significant. We compared guideline adherence between the groups using chi-square tests.

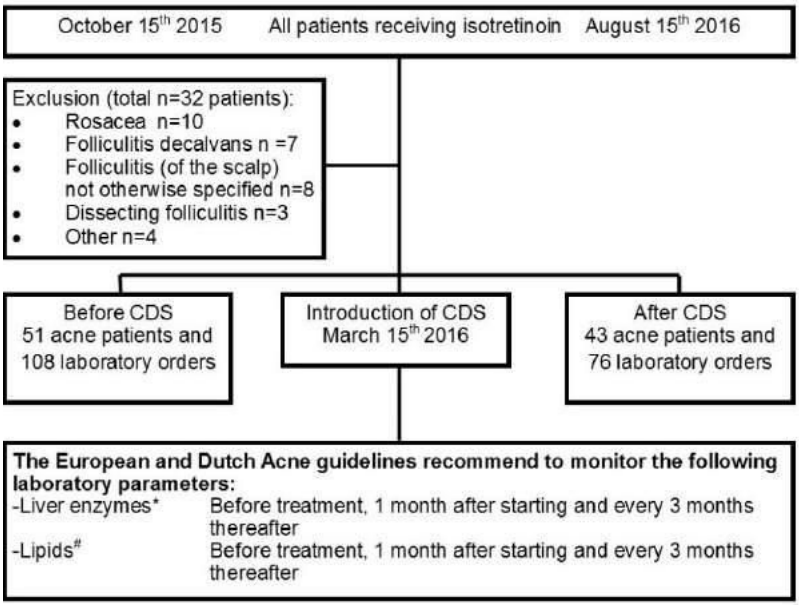


Figure 1: Study design, flow chart and clinical decision support text physicians see when initiating an isotretinoin laboratory order. The recommended laboratory tests are already ticked.[13,14] * aspartate aminotransferase (AST), and/or alanine aminotransferase (ALT)). # triglycerides (TG), low-density lipoprotein (LDL), high-density lipoprotein (HDL).

RESULTS

We identified 108 laboratory orders in 51 acne patients receiving isotretinoin before the introduction of the guideline based CDS and 76 laboratory orders in 43 patients after the introduction. In total 32 patients receiving isotretinoin were excluded (Fig.1). There were 15 different physicians in the analysis. The majority of physicians contributed 2-4 patients.

Primary outcomes

The data shows an increase in guideline adherence regarding recommended laboratory testing, both in frequency (passive support) as in the tests used to monitor the treatment (active support) (Table 1). However, the increase in adherence with passive support is not statistically significant (Table 1). Adherence is depicted per patient and per laboratory order. Undertesting occurred two times before and three times after the CDS introduction (included in non-adherence; Table 1). Documented reasons for non-adherence to the recommended laboratory tests were co-morbidity (one patient with diabetes and one with Gilbert's syndrome) and co-medication (one patient with terbinafine and one with prednisone).

	Adherence passive clinical decision support		**Adherence active clinical decision support**	
	Patients % (no.)	Laboratory orders % (no.)	Patients % (no.)	Laboratory orders % (no.)
Before CDS	57% (29)	73% (80)	0% (0)	0% (0)
After CDS	70% (30)	77% (61)	63% (27)	75% (57)
Chi-square	0,197	0,321	<0,001	<0,001

Table 1: Guideline adherence before and after passive clinical decision support (timely laboratory testing) and active clinical decision support (recommended laboratory tests)

Secondary outcomes

Total costs for only the laboratory tests were € 2207 before and € 686 after the guideline based CDS. That is a more than 2.5 fold reduction per patient within five months (€ 43 to € 16). The reduction in non-adherent testing is portrayed in Fig.2. There was however no significant reduction in non-recommended extra visits or venipuncture per patient. The mean non-recommended extra visits per patient was 0,43 before implementation of the guideline based CDS tool and 0,32 after. For venipunctures this was 0,59 versus 0,42.

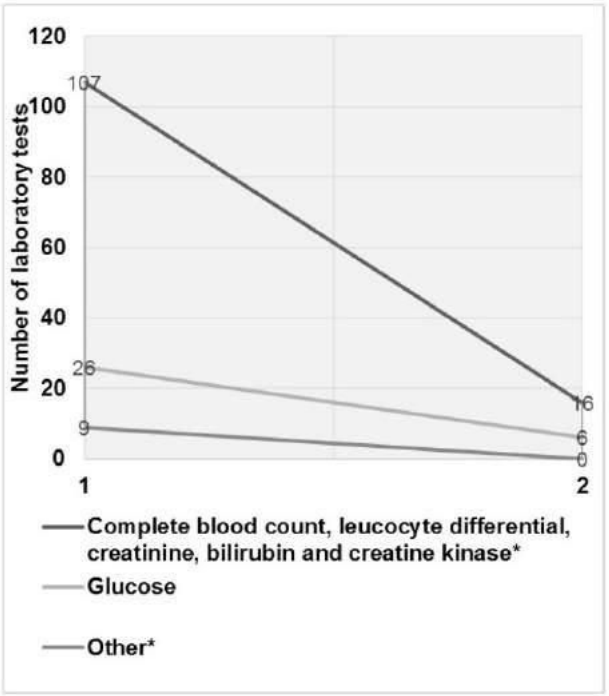


Figure 2: Frequency of non-adherent testing before (1) and after (2) implementation of the guideline based clinical decision support tool in the EHR system.
*The category 'Other' consists of ureum (2), lactate dehydrogenase (2), other liver enzymes (4) and glycated haemoglobin (1).

DISCUSSION

Notwithstanding the absence of statistical significance in the passive CDS part of this study, we are able to show improvement in adherence to recommended laboratory tests and a reduction in costs from test ordering. Our data underlines the importance of improving laboratory testing adherence because, without CDS, none of the laboratory orders were consistent with guideline recommendations and 43% of the patients experienced non-timely laboratory testing. Assuming that the acne guidelines themselves are beneficiary in daily patient care, the results promise a benefit of active CDS.

Being just an addition to an otherwise conventional laboratory order, the guideline based CDS in this study enters daily practice through normal workflow screens, imposes no specific burdens of its own, and is scalable to each patient one desires to reach. Another advantage of the support is the absence of pop-up screens. These screens can have unintended consequences, such as alert desensitization. Perhaps this is the reason why our findings are in contrast with Roshanov et al. This review of

factors associated with success of CDS found that 58% of the randomized controlled trials showed some benefit in clinical process or patient outcomes. Yet, systems that presented advice within EHR or order entry systems were less likely to be effective.[16] The effect among dermatologists validates the idea that CDS can improve clinical performance. Still, the adherence after the implementation of the passive textual support was not statistically significantly different. A discrepancy between active and passive support is most likely caused by a barrier in the physician itself. As noted before by Cabana et al. inertia of previous practice seems to be one of the most important barriers to implementation.[2]

Strengths and limitations

As our study is not randomized, better adherence could be attributed to chance association. On the other hand, our study took place in daily practice by those who are the target of the guidelines: dermatologists caring for acne patients. An advantage of the study design is that all dermatologists received the same support; also those who are convinced that their actions are already in line with guideline standards and decide that there is no need to consult the guideline. Finally, certain possible biases were inevitable. We are aware of the fact that non-reporting is not the same as non-adherence. Clinicians can have other reasons to monitor a patient using (extra) laboratory tests (e.g. a marginal non-elevated value), though those reasons were not given. Another noticeable difference and possible bias is that there are fewer patients included after the CDS intervention than before. A likely explanation may be a reluctance to start isotretinoin during the summer because of the associated photosensitivity.

CONCLUSION

The guideline based CDS in this prospective study appears to improve adherence to recommended laboratory tests in acne patients receiving isotretinoin. It might, in some cases, lead to better timeliness of laboratory test ordering. Also guideline based CDS seems a promising tool to reduce costs. However, active support was far more effective than passive support. Passive CDS remains subject to barriers preventing implementation such as inertia of previous practice. In summary, simple guideline based active CDS that is incorporated into daily workflow can diminish overtesting. Further research will hopefully identify more barrier-free guideline based CDS to reduce nondirected- and overtesting in other conditions.

REFERENCES

1 Institute of Medicine (US) Committee to Advise the Public Health Service on Clinical Practice Guidelines. Clinical Practice Guidelines: Directions for a New Program. Field MJ, Lohr KN, eds. Washington (DC): National Academies Press (US); 1990.

2 Cabana MD, Rand CS, Powe NR et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999;282(15):1458-65.

3 Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 1993;342(8883):1317-22.

4 Kilsdonk E, Peute LW, Riezebos RJ, Kremer LC, Jaspers MW. Uncovering healthcare practitioners' information processing using the think-aloud method: From paper-based guideline to clinical decision support system. Int J Med Inform 2016;86:10-9.

5 Garcia D, Moro CM, Cicogna PE, Carvalho DR. Method to integrate clinical guidelines into the electronic health record (EHR) by applying the archetypes approach. Stud Health Technol Inform 2013;192:871-5.

6 Jones JB, Stewart WF, Darer JD, Sittig DF. Beyond the threshold: real-time use of evidence in practice. BMC Med Inform Decis Mak 2013;13:47.

7 Lobach DF, Hammond WE. Computerized decision support based on a clinical practice guideline improves compliance with care standards. Am J Med 1997;102(1):89-98.

8 Pestotnik SL, Classen DC, Evans RS, Burke JP. Implementing antibiotic practice guidelines through computer-assisted decision support: clinical and financial outcomes. Ann Intern Med 1996;124(10):884-90.

9 Lu MT, Rosman DA, Wu CC et al. Radiologist Point-of-Care Clinical Decision Support and Adherence to Guidelines for Incidental Lung Nodules. J Am Coll Radiol 2016;13(2):156-62.

10 Boggan JC, Schulteis RD, Donahue M, Simel DL. Guideline-based decision support has a small, non-sustained effect on transthoracic echocardiography ordering frequency. BMJ Qual Saf 2016;25(1):57-62.

11 Tapper EB, Sengupta N, Lai M, Horowitz G. A Decision Support Tool to Reduce Overtesting for Ceruloplasmin and Improve Adherence With Clinical Guidelines. JAMA Intern Med 2015;175(9):1561-2.

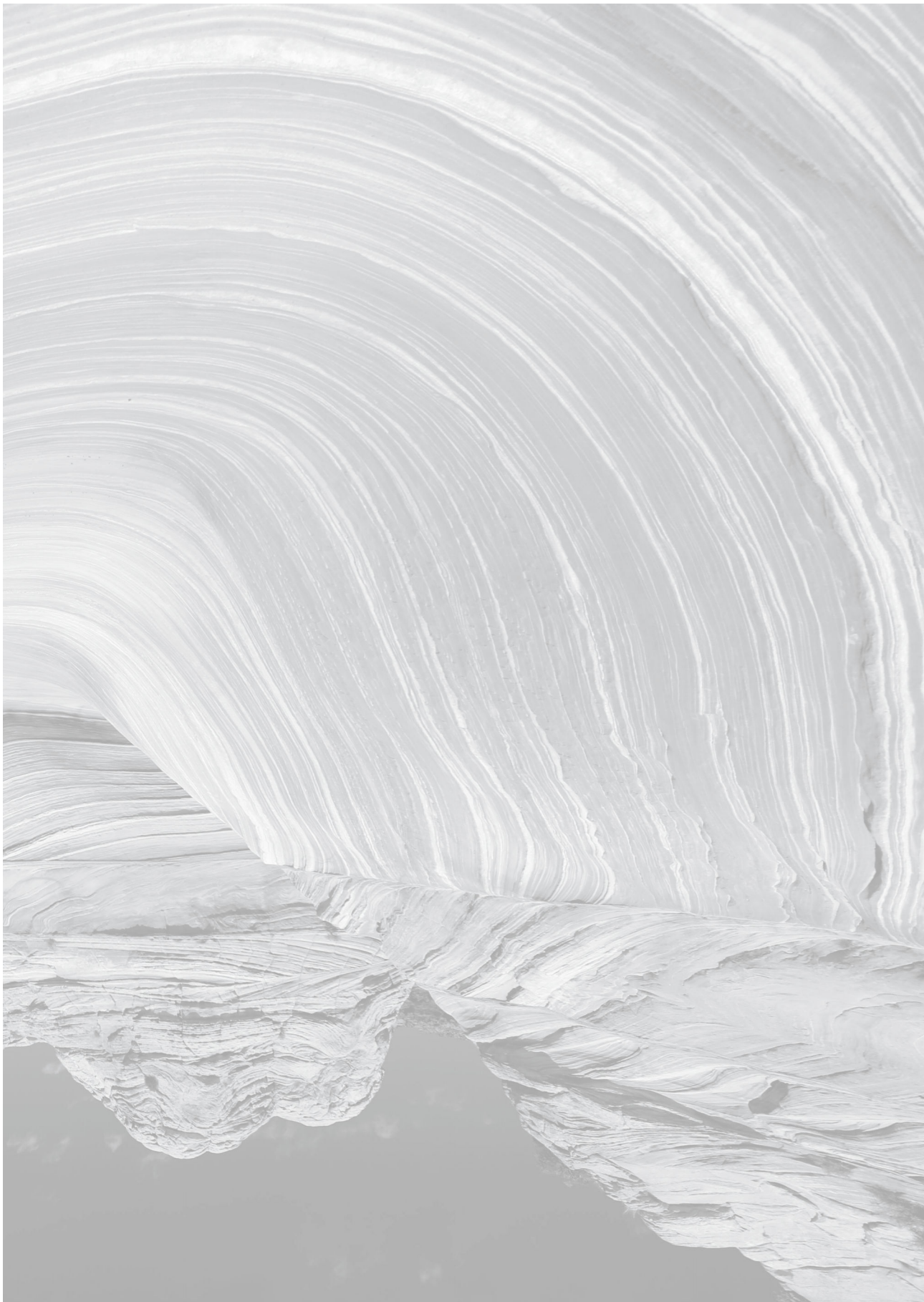
12 Bright TJ, Wong A, Dhurjati R et al. Effect of clinical decision-support systems: a systematic review. Ann Intern Med 2012;157:29–43.

13 Nast A, Dréno B, Bettoli V et al. European evidence-based (S3) guidelines for the treatment of acne. J Eur Acad Dermatol Venereol 2012;26 Suppl 1:1-29.

14 van de Kerkhof PCM, Borgonjen RJ. Acne. In: Samenvatting richtlijnen Dermatologie 2012. Bik CMJM, Van Everdingen JJE, eds. Haarlem: DCHG medische uitgeverij, 2012; 11-22.

15 Lee YH, Scharnitz TP, Muscat J, Chen A, Gupta-Elera G, Kirby JS. Laboratory Monitoring During Isotretinoin Therapy for Acne: A Systematic Review and Meta-analysis. JAMA Dermatol 2016;152(1):35-44. Erratum in: JAMA Dermatol. 2016;152(1):114.

16 Roshanov PS, Fernandes N, Wilczynski JM et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. BMJ 2013;346:f657.



7

IMPROVING THE APPLICABILITY OF GUIDELINES

Improving the applicability of guidelines on nonmelanoma skin cancer in frail older adults: a multidisciplinary expert consensus and systematic review of current guidelines

Published as

Lubeek SF, Borgonjen RJ, van Vugt LJ, Olde Rikkert MG, van de Kerkhof PC, Gerritsen MJ. Improving the applicability of guidelines on nonmelanoma skin cancer in frail older adults: a multidisciplinary expert consensus and systematic review of current guidelines.

Br J Dermatol. 2016;175(5):1003-1010.

SUMMARY

Background Balancing treatment decisions in frail older adults with nonmelanoma skin cancer (NMSC) can be challenging. Clinical practice guidelines (CPGs) could provide assistance.
Objectives To collect and prioritize items related to frail older adults with NMSC for integration into CPGs and to assess the current extent of this integration.
Methods Items were collected and prioritized by a multidisciplinary working group (29 members) using a modified Delphi procedure and a five-point Likert scale. To assess current integration of these items in CPGs, a systematic review was subsequently performed by two independent reviewers using five medical databases (PubMed, Embase, Cochrane Library, SUMsearch and Trip Database), websites of guideline developers/databases, and (inter)national dermatological societies.
Results Prioritization of a final 13-item list showed that 'limited life expectancy' (4.5 ± 0·9) and 'treatment goals other than cure' (4.4 ± 0.7) were most desired to be integrated into CPGs; both were included in six (46%) of the CPGs found (n = 13). Attention to 'tumour characteristics' and 'comorbidities' were included in CPGs most often (100% and 77%, respectively).
Conclusion More attention to items related to frail older adults in NMSC CPGs is broadly desired, but CPG integration of these items is currently limited. More integration might stimulate more holistic, personalized and patient-centred care in frail older adults.

INTRODUCTION

Frailty is mostly described as the vulnerability of an individual to poor homeostatic resolution after a stressor event, which is considered the result of a cumulative decline in the functioning of various physiological systems with increasing age.[1] Medical decision making will be challenging within the growing population of frail older adults worldwide, especially since factors such as limited life expectancy, comorbidities, and logistical limitations (e.g. travel distance, availability of support by relatives) have to be taken into account.[2] Furthermore, guidance from clinical evidence is limited, since frail older adults are often excluded or underrepresented in clinical trials.[3-5]

A common dermatological condition among frail older adults in which these challenges are frequently faced is nonmelanoma skin cancer (NMSC).[6-8] On the one hand, most NMSC generally have a relatively low malignant potential compared to many other types of cancer. This might favor conservative management, especially in case of a limited life expectancy. On the other hand, NMSC can cause considerable morbidity (e.g. pain, secondary infection or functional impairment) and in the longer term even mortality.[9-11] Therefore, weighing potential benefits and risks of treatment together with preferences from patients and/or their family is essential, but not always easily performed.

Clinical practice guidelines (CPGs) are defined as a set of general recommendations intended to optimize patient care, which assist healthcare providers in daily clinical practice.[12] Obviously, it is not possible to include specific guideline recommendations for every possible situation in every individual patient. Therefore, it could be desirable and in the best interest of a patient to depart from a CPG in some circumstances. However, balancing the extensiveness of the content and the specificity of recommendations in a CPG with generally accepted or self-evident aspects of daily clinical practice might be a challenge. Various studies have shown the potential benefits of integration of certain topics, including healthcare structures, physician's behavior, and patient outcomes in CPGs.[13] Nevertheless, implementation and adherence to CPGs is often suboptimal and influenced by many factors like appropriate inclusion of the needs and preferences of all relevant professional groups, feasibility and the flexibility to individualize recommendations.[14-15] Previous studies have shown that guidance from CPGs in frail older adults is limited. Firstly, CPGs (together with most health care systems) often focus on curative care and pay little or no attention to other treatment goals, like palliation or preservation of function. Furthermore, CPGs often fail to address issues related to frail older adults and therefore, provide little guidance in the care for this group of patients.[16-20]

In the present study, we aimed to collect items which are considered important in medical decision making in frail older adults with NMSC and prioritize them based on whether they should be integrated in NMSC CPGs. Additionally, we assessed how these items are included in the current NMSC CPGs worldwide and provide recommendations for the future.

METHODS

Multidisciplinary working group formation

A multidisciplinary working group of Dutch experts was formed. Age, gender, years of experience and health care setting were taken into account to ensure a well-balanced working group.[21] An overview of the working group formation process and represented parties is given in Supplementary Table S1.

Composition and prioritization of a list of items considered important

A conceptual list of 10 items influencing medical decision making in frail older adults with NMSC composed by the authors was discussed and adapted by a condensed version of the multidisciplinary working group (including at least one expert per medical specialty), using a modified Delphi procedure to reach consensus.[22] Consensus was defined as agreement to include an item by all condensed working group members. During the consensus procedure, 5 items were added to the list and 2 items were removed. Furthermore, 2 items were rephrased to enhance distinctness. Consensus was reached after two rounds and 13 items were included on the final item list. In addition, to prioritize items for integration in CPGs, participants in the extended working group were asked to provide the items with a score using a five-point Likert scale. A higher score indicates the item is more desired to be included in CPGs.

Systematic review of literature

A comprehensive review of literature was performed independently by two authors (SL and LvV) to identify CPGs concerning NMSC. An initial search was performed using PubMed, EMBASE (Excerpta Medica Database) and the Cochrane Library. In addition, websites of guideline developers/databases and (inter)national dermatological societies as well as SUMsearch and Trip Database were searched for relevant guidelines not yet identified. An overview of the search strategy, selection criteria, and data sources used is presented in Supplementary Tables S2 and S3. This review was conducted and reported according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement.[23]

Guideline assessment

Included CPGs were assessed for inclusion of each item from the developed item list (yes or no) and the place in the CPG (in recommendations, only in plain text, or not further specified). Furthermore, overall quality and reporting of the CPGs was assessed using the Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument, an internationally recognized and validated instrument to evaluate the quality of guidelines.[24] The AGREE II instrument consists of 23 items distributed over six domains. Every item is scored by a 7-point Likert-scale (1 = strongly disagree, 7 = strongly agree) after which domain scores could be calculated and presented as proportions, following the user manual (http://www.agreetrust.org). Guideline assessment was performed by two independent reviewers (SL and RB), both experienced in guideline appraisal. All available appendices and supplementary documents were included in the assessment.

Statistics

Statistics were reported following the "Statistical Analyses and Methods in the Published Literature" (SAMPL) guidelines.[25] Categorical variables were expressed as numbers and percentages. Continuous variables were expressed as means and standard deviations (SD) or medians and ranges, dependent on the distribution of data. Representativeness of the working group members was tested by separately comparing the distribution in age, sex and health-care setting for the members working in primary care (general practitioners and elderly care physicians) and secondary (or tertiary care) with the corresponding overall populations of physicians in the Netherlands.[26-27] Data collection and analysis was performed using Statistical Package for Social Sciences (SPSS®) for Windows, version 22.0 (IBM Corporation, Armonk, NY).

RESULTS

Multidisciplinary working group characteristics

The condensed and extended working group, the first for composition and the latter for prioritization of the item list, consisted of 11 and 29 members respectively. Working group characteristics are given in Table 1. A relatively large number of medical specialists in the working group works in academic hospitals as compared to the overall population of medical specialists (57% vs. 20%, respectively).

Characteristic	Mean (SD) or n (%)	
Age, mean (SD)	50.1 (7.4)	
Male sex, n (%)	18 (62%)	
Discipline, n (%)	Dermatologist	13 (45%)
	Plastic surgeon	2 (7%)
	Otorhinolaryngologist	2 (7%)
	Oral and maxillofacial surgeon	2 (7%)
	Radiation therapist	2 (7%)
	Geriatrician	2 (7%)
	Elderly care physician	3 (10%)
	General practitioner	3 (10%)
Years of experience[1], mean (SD)	17.3 (7.0)	
Health care setting, n (%)[2]	General hospital	11 (38%)
	Academic hospital	13 (45%)
	Other	9 (31%)

Table 1: Multidisciplinary working group characteristics (n=29) [1]: residency years excluded; [2]: some working group members work in multiple health care settings; no missing data.

Prioritization of the final list

The final list of prioritized items to be integrated in CPGs is shown in Table 2. The overall items score was 3.9 (0.9). Highest scores were given to the items "limited life expectancy" and "treatment goals other than curation": 4.5 (0.9) and 4.4 (0.7), respectively. Lowest scores were given to the items "performance of skin self-examination" and "legal status / impaired capacity to give informed consent": 3.3 (0.7) and 3.3 (0.9), respectively. No scores were missing.

Guideline search and selection

The initial search for NMSC CPGs resulted in 6878 results, of which 3682 remained after removal of duplicates. After the additional search and screening by title and abstract, 37 full-text articles were assessed for eligibility and eventually 13 articles were included.[28-40] Figure 1 shows the guideline selection process.

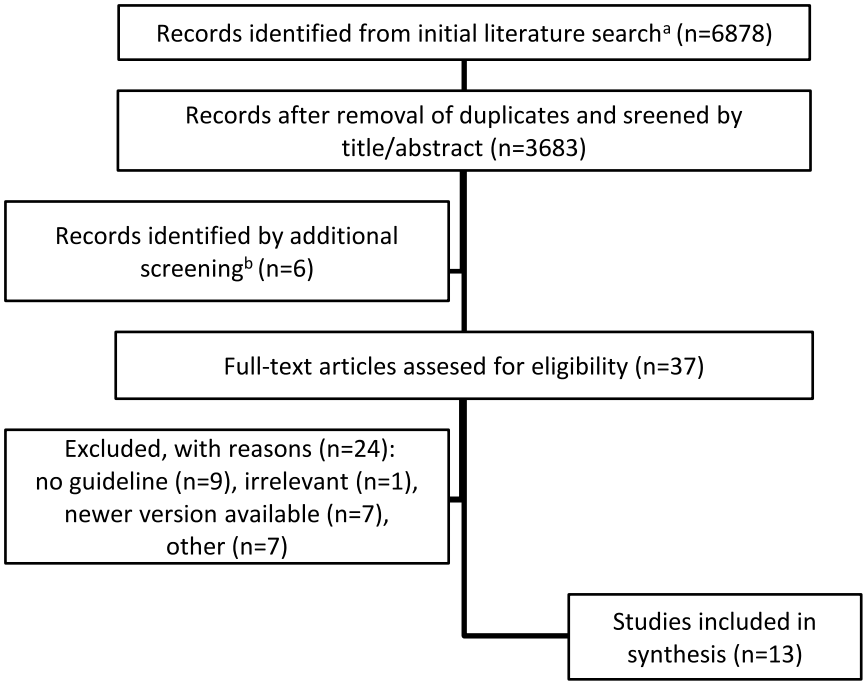


Figure 1: An overview of the guideline selection process in this systematic review on nonmelanoma skin cancer guidelines. [a] Pubmed, EMBASE and the Cochrane library; [b] Websites of guideline developers/databases, (inter)national dermatological societies, SUMsearch, and the Trip Database.

Guideline assessment

Most guidelines are from Europe (n=11) and less than five years old (n=7). Squamous cell carcinoma CPGs were most often found (n=6), followed by basal cell carcinoma CPGs (n=5) and CPGs combining the both (n=2). The median AGREE II overall score was 50% (range 36%-86%) and the median number of items included per guideline was 3 (23%; range 1-9). Guideline characteristics together with the AGREE II scores and the number of items included per guideline are summarized in Table 3.

All of the individual items from the prioritized list were included in at least one guideline. Two items were included in the majority of the CPGs: tumor characteristics (n=13, 100%) and comorbidities (n=10, 77%). The two items with the highest prioritization scores ("limited life expectancy" and "treatment goals other than curation") were both included in 6 (46%) CPGs. All other items were included less frequently. Most items were only mentioned in plain text (n=20, 49%), followed by recommendations (n=15, 31%) and "not further specified" (n=10, 20%). A full overview of the individual items and their CPG-inclusion is shown in Table 2.

Item	Prioritization score, mean (SD)	Included in guidelines, n (%)	Place in CPGs, n (%)		
			In recommendations	Only in plain text	Not further specified
Limited life expectancy	4.5 (0.9)	6 (46)	0 (0)	4 (67)	2 (33)
Treatment goals other than curation	4.4 (0.7)	6 (46)	1 (17)	4 (67)	1 (17)
Comorbidities	4.2 (0.7)	10 (77)	0 (0)	8 (80)	2 (20)
Tumor characteristics	4.1 (0.8)	13 (100)	10 (77)	0 (0)	3 (23)
Cognitive impairment	4.0 (0.7)	2 (15)	0 (0)	2 (100)	0 (0)
Limited possibilities to visit hospital/clinic	4.0 (0.9)	4 (31)	2 (50)	1 (25)	1 (25)
Option of follow-up by primary care physician	4.0 (0.9)	1 (8)	0 (0)	1 (100)	0 (0)
Family involvement in medical decision making	3.9 (1.0)	1 (8)	0 (0)	1 (100)	0 (0)
Functional impairment	3.9 (0.6)	1 (8)	0 (0)	1 (100)	0 (0)
Treatment compliance	3.8 (0.9)	2 (16)	1 (50)	1 (50)	0 (0)
Primary care physicians involvement in medical decision making	3.7 (1.0)	1 (8)	1 (100)	0 (0)	0 (0)
Legal status / impaired capacity to give informed consent	3.3 (0.9)	1 (8)	0 (0)	1 (100)	0 (0)
Performance of skin self-examination	3.3 (0.7)	1 (8)	0 (0)	0 (0)	1 (100)

Table 2: Items desired to be integrated in nonmelanoma skin cancer clinical practice guidelines (CPGs) related to medical decision making in frail older adults, prioritized by multidisciplinary working group (n=13). Higher scores indicate the item is more desired to be included in clinical practice guidelines; values may not add up due to rounding. SD = standard deviation, n = number

First author (year*)	Reference	Country / region	Initiating organi-zation(s)	Guideline topic (BCC/SCC/ both)	AGREE II Domain scores, %					AGREE II Overall score, %	Items[1] included in guideline, n (%)
Marks (2008)	28	Australia	CCA and ACN	Both	69	79	38	90	59	61	64	6 (46)
Stratigos (2015)	29	Europe	EDF, EADO and EORTC	SCC	64	38	49	60	18	68	50	4 (31)
Trakatelli (2012)	30	Europe	EDF	BCC	48	31	31	71	27	68	43	2 (15)
Bonerandi (2009)	31	France	SFD	SCC	88	81	77	93	45	82	86	9 (69)
Dandurand (2004)	32	France	SFD and ANAES	BCC	76	57	41	74	39	14	50	3 (23)
Breuningen (2012)	33	Germany	DKG and DDG	SCC	31	36	26	71	21	61	36	2 (15)
Hauschild (2012)	34	Germany	DKG and DDG	BCC	31	36	24	71	21	61	36	4 (31)
Kelleners (2014)	35	Netherlands	NVDV	BCC	88	76	88	90	54	64	86	5 (38)
Krekels (2010)	36	Netherlands	NVDV	SCC	90	83	80	88	38	39	79	2 (15)

First author (year*)	Reference	Country / region	Initiating organi-zation(s)	Guideline topic (BCC/SCC/ both)	AGREE II Domain scores, %						Items[1] included in guideline, n (%)
										AGREE II Overall score, %	
Proby (2014)	37	Scotland	SIGN	SCC	98	83	79	88	95	75	86	3 (23)
Motley (2009)	38	United Kingdom	BAD	SCC	71	33	30	60	41	14	43	3 (23)
Telfer (2008)	39	United Kingdom	BAD	BCC	60	36	34	71	43	43	50	4 (31)
Miller (2012)	40	United States	NCCN	Both	57	45	25	69	38	50	43	1 (8)
				Median:	67	55	48	77	41	54	50	3
				Range	31-98	31-83	24-88	60-93	18-95	14-82	36-86	1-9

Table 3: Overview of included guidelines and results of guideline assessment (n=13)*: year of authorization; BCC: basal cell carcinoma; SCC: squamous cell carcinoma (of the skin); AGREE II Domains include (1) scope and purpose, (2) stakeholder involvement, (3) rigour of development, (4) clarity of presentation, (5) applicability, and (6) editorial independence; [1]: items related to frail older adults in relation to nonmelanoma skin cancer care; CCA: Cancer Council Australia; ACN: Australian Cancer Network; EDF: European Dermatology Forum; EADO: European Association of Dermato-Oncology; EORTC: European Organization for Research and Treatment of Cancer; SFD: French Dermatology Society; ANAES: French National Agency for Accreditation and Evaluation in Healthcare; DKG: German Cancer Society; DDG: German Society of Dermatology; NVDV: Dutch Society of Dermatology and Venereology; SIGN: Scottish Intercollegiate Guidelines Network; BAD: British Association of Dermatologists; NCCN: National Comprehensive Cancer Network.

DISCUSSION

In this study we collected items considered important in medical decision making in frail older adults with NMSC and prioritized them based on whether they should be integrated in NMSC CPGs. It was shown that many items were desired to be integrated in NMSC CPG, but currently the overall CPG-integration is limited.

The importance of frailty among older adult patients with cancer and integration of appropriate considerations into CPGs is increasingly acknowledged and multiple organizations and taskforces have published position papers and guidelines on this topic.[41-43] However, in agreement with this study, previous studies in various fields of medicine showed that attention to frailty-related items like comorbidity and treatment goals other than curation in CPGs is still limited.[16-20,44-45]

The items from this study show a high level of agreement with existing literature on important domains in geriatric oncology.[43,46-47] The expert panel in geriatric oncology in the study from O'Donovan et al. rated functional status, followed by comorbidities and cognition as the most important domains influencing treatment decisions, which were also included in this study although differently prioritized.[47] On the other hand, other domains like mood and nutritional status were not included in this study. Although most of these domains were initially discussed in the working group, they were not included eventually because they were considered of limited additional relevance in the context of treatment of NMSC.

Of course, the items within this study do not solely apply to the population of frail older adults and should be integrated in medical decision making in every patient. However, we specifically focused on frail older adults because in our experience medical decision making is more frequently influenced by the mentioned items in this population, compared to the younger, less frail and often less dependent population.

Based on the outcomes of this study we suggest to include at least the following items in a CPG on NMSC: (1) limited life expectancy, (2) treatment goals other than curation, (3) comorbidities, and (4) tumor characteristics. These items could be integrated in CPG in several ways, e.g. in a dedicated section on general considerations or implementation of the CPG in daily practice or in the discussion of each individual recommendation. The items should be clearly described and easily identifiable.

The other items could certainly be of added value, but we think these should be weighed in relation to the objectives and health questions covered by the guideline (e.g. specific focus on prevention, one treatment option), the target population, the target users of the CPG (e.g. dermatologists, general practitioners, multiple disciplines), and the characteristics of the health-care system(s) in which the guideline will be applied (e.g. the existence and role of general practitioners).

Furthermore, we think that participation of a geriatrician and/or elderly care physician as well as involvement of patients in the guideline development process will assist in providing proper attention to frail older adults.

Integration of frailty-related items within CPGs might stimulate a more holistic, personalized and patient-centered care instead of the more tumor-centered care which seems to have become the current standard. Also, integration of more frailty-related items in CPGs might lead physicians to feel more empowered to deviate from regular protocols in the best interest of a patient, especially since CPGs are more and more used as quality indicators and performance incentives by health policy makers.[48] Nevertheless, we acknowledge that integration of frailty-related items in CPGs does not necessarily mean a change in daily clinical practice and/or patient outcomes. Therefore, attention to other factors which could lead to the mentioned change of perspective in health care (e.g. education of health-care providers, evaluation and adaptation of current healthcare systems, multidisciplinary approaches, usage of screening instruments on frailty etc.) remains important as well.

Also, we would like to emphasize the importance of paying enough attention to patient preferences and shared-decision-making in busy daily dermatological practice. This is closely related to several items collected in this study, like cognitive impairment and an impaired capacity to give informed consent. Cognitive impairment is typically defined as the loss of one or more cognitive functions (e.g. memory or problem-solving), which could be seen as a continuum ranging from mild (e.g. mild cognitive impairment) to severe (e.g. end-stage dementia).[49] It might be difficult to estimate a patient's cognitive functioning and the ability to provide informed consent. Therefore, we think in case of doubt the threshold to consult close relatives or other health-care providers related to the patient should be low.

The recognition of frailty in daily dermatological practice can be challenging and screening tools to detect frailty and assist in medical decision making might be helpful. An increasing body of literature is available on detecting frailty in oncology by geriatric assessment (GA) and related screening tools.[50] The goals of GA is to determine the medical, psychosocial and functional capabilities of an older adult patient to compose an integrated treatment plan.[43,51] GA can assist in identifying frailty, help making tailored treatment decisions, and might optimize patient outcomes.[43,50] Several methods are described to perform GA, ranging from multidisciplinary but time-consuming comprehensive geriatric assessment (CGA) to less extensive, derivative screening methods like the Vulnerable Elders Survey-13 (VES-13), Geriatric 8 (G8), or Groningen Frailty Index (GFI).[43,46] However, no studies on NMSC were found and extrapolation of available data from studies on other types of

cancer seems limited by factors such as the relatively low malignant natural course of NMSC and the generally less intense and less complex treatment regimes with little adverse events compared to many other types of cancer. Still, the general principles and goals of GA and frailty screening are universal and could be useful in the field of NMSC. Future research should focus on the potential effects and added value of GA and derivative screening tools in frail older adults with NMSC.
The results and recommendations in this study could be of interest to CPG development in general, since most items are not restricted to NMSC.

Limitations
Despite the thorough conceptualization and discussion of the different items within the working group, some overlap as well as potential misinterpretation could have occurred. The study was performed in the Netherlands including Dutch physicians, which should be kept in mind when generalizing the results. No patients were included in the working group. We acknowledge the importance of including patient preferences and shared decision making and we would like to suggest future research should focus on this in more detail. The working group formed for this study included a relatively large number of medical specialists working in academic medicine. This could be explained by the large number of medical specialists from academic hospitals directly participating in the activities of medical societies (for instance national guideline committees) and who were therefore delegated to the working group in this study. However, comparison of the working group members with the overall population of physicians showed no important differences considering age and gender for both medical specialists and primary care physicians, indicating an appropriate representativeness regarding these characteristics. Furthermore, it may be that assessed CPGs received low scores on item inclusion despite of the inclusion of frailty-related content, when this content could not be classified into the assessed items on the list. Nonetheless, neither of the reviewers noticed any important (mis)classification bias.

CONCLUSION
More attention for items related to frail older adults in NMSC CPGs is broadly desired and integration of these items in current CPGs is limited. More integration might stimulate a more holistic, personalized and patient-centered care.

Working group formation:	
A multidisciplinary working group of experts was formed consisting of representatives from several medical specialties (as shown below). Age, gender, years of experience and health care setting were taken into account to ensure a well-balanced working group. All Dutch medical societies related to the mentioned medical specialties were asked to delegate a representative with special affinity regarding nonmelanoma skin cancer and/or frail older adults. Moreover, representatives of the national dermato-oncology working group, the national guideline committees (in both primary and secondary care), and a national network of dermatologists with special interest in geriatric dermatology were also included. Finally, the working group was completed and further balanced with members from several regional health-care settings.	
Medical specialties included:	**Representation from:**
Dermatology	Dutch Society of Dermatology and Venereology[1] and national network of dermatologists with interest in geriatric dermatology
Plastic surgery	Dutch Society for Plastic Surgery
Otorhinolaryngology	Dutch Society of Otorhinolaryngology and Head & Neck Surgery
Oral and maxillofacial surgery	Dutch Society of Oral and Maxillofacial Surgery
Radiation therapy	Dutch Society of Radiotherapy and Oncology
Geriatric medicine	Dutch Geriatric Society
Elderly care (nursing home) medicine	Dutch Association of Elderly Care Physicians and Social Geriatricians
General practitioner care	Dutch College of General Practitions[2]

Supplementary Table S1: Overview of represented parties included in a working group considering nonmelanoma skin cancer care in frail older adults [1]: including the national dermato-oncology working group and national guideline committees (basal cell carcinoma and squamous cell carcinoma) and; [2]: including national guideline committee (nonmelanoma skin cancer); the working group was completed and balanced with members from several regional health-care settings.

Medical search engines searched:
Pubmed (http://www.ncbi.nlm.nih.gov/pubmed)
EMBASE (Excerpta Medica Database) (http://ospguides.ovid.com)
Cochrane Library (http://www.cochranelibrary.com)
SUMsearch (http://sumsearch.org)
Trip Database (https://www.tripdatabase.com)
Guideline developers/databases searched:
Guidelines International Network (http://www.g-i-n.net)
National Guidelines Clearinghouse / Agency for Healthcare Research and Quality (http://guideline.gov and http://www.ahrq.gov)
Ärztliches Zentrum für Qualität in der Medizin (http://www.aezq.de/)
National Institute for Clinical Excellence (https://www.nice.org.uk/)
New Zealand Guidelines Group (http://www.health.govt.nz)
Scottish Intercollegiate Guidelines Network (http://www.sign.ac.uk)
Dutch Association of Comprehensive Cancer Centers (http://www.oncoline.nl)
Kwaliteitskoepel Medisch Specialisten (http://www.kwaliteitskoepel.nl)
Dutch College of General Practitioners (http://www.nhg.artsennet.nl)
(Inter)national dermatological societies searched:
American Academy of Dermatology (https://www.aad.org)
Australasian College of Dermatologists (https://www.dermcoll.edu.au)
British Association of Dermatologists (http://www.bad.org.uk)
Canadian Dermatology Association (http://www.nzdsi.org)
Deutsche Dermatologische Gesellschaft (http://www.derma.de)
Dutch Society of Dermatology and Venereology (http://www.huidarts.info)
European Dermatology Forum (http://www.euroderm.org)

Supplementary Table S2: Overview of the databases and websites searched in this systematic review on nonmelanoma skin cancer guidelines

Key word selection and search limitations:
The key words (and synonyms) “basal cell carcinoma, “cutaneous squamous cell carcinoma”, or “nonmelanoma skin cancer”, combined with “guideline” (both as keyword and separately as limitation) were used. All related synonyms found in the Medical Subject Headings (MeSH) database were also used. All search terms were limited to title and/or abstract. Complete search strategies are shown below:
Search strings:
*“Basal cell carcinoma “combined with “guidelines” (and synonyms)**
((((((((((((((((((((basal cell carcinoma[Title/Abstract]) OR basal cell carcinomas[Title/Abstract]) OR BCC[Title/Abstract]) OR BCCs[Title/Abstract]) OR sBCC[Title/Abstract]) OR sBCCs[Title/Abstract]) OR nBCC[Title/Abstract]) OR nBCCs[Title/Abstract]) OR rodent ulcer[Title/Abstract]) OR rodent ulcers[Title/Abstract]) OR basalcell carcinoma[Title/Abstract]) OR basalcell carcinomas[Title/Abstract]) OR basal cell neoplasm[Title/Abstract]) OR basal cell neoplasms[Title/Abstract]) OR basal cell tumor[Title/Abstract]) OR basal cell tumour[Title/Abstract]) OR basal cell tumors[Title/Abstract]) OR basal cell tumours[Title/Abstract]) OR basalioma[Title/Abstract]) OR basaliomas[Title/Abstract])) AND ((((((((guideline[Title/Abstract]) OR guidelines[Title/Abstract]) OR recommendation[Title/Abstract]) OR recommendations[Title/Abstract]) OR protocol[Title/Abstract]) OR protocols[Title/Abstract]) OR standard[Title/Abstract]) OR standards[Title/Abstract]); with and without “guidelines” as limitation
*“Cutaneous squamous cell carcinoma” combined with “guidelines” (and synonyms)**
((((((((skin[Title/Abstract]) OR cutaneous[Title/Abstract]) OR dermis[Title/Abstract]) OR epidermis[Title/Abstract]) OR dermal[Title/Abstract]) OR epidermal[Title/Abstract])) AND ((((((((((((((((((squamous cell carcinoma[Title/Abstract]) OR squamous cell carcinomas[Title/Abstract]) OR SCC[Title/Abstract]) OR SCCs[Title/Abstract]) OR SqCC[Title/Abstract]) OR SqCCs[Title/Abstract]) OR squamouscell carcinoma[Title/Abstract]) OR squamouscell carcinomas[Title/Abstract]) OR squamous cell neoplasm[Title/Abstract]) OR squamous cell neoplasms[Title/Abstract]) OR squamous cell tumor[Title/Abstract]) OR squamous cell tumour[Title/Abstract]) OR squamous cell tumors[Title/Abstract]) OR squamous cell tumours[Title/Abstract]) OR planocellular carcinoma[Title/Abstract]) OR planocellular carcinomas[Title/Abstract]) OR epidermoid carcinoma[Title/Abstract]) OR epidermoid carcinomas[Title/Abstract])) AND ((((((((guideline[Title/Abstract]) OR guidelines[Title/Abstract]) OR recommendation[Title/Abstract]) OR recommendations[Title/Abstract]) OR protocol[Title/Abstract]) OR protocols[Title/Abstract]) OR standard[Title/Abstract]) OR standards[Title/Abstract]); with and without “guidelines” as limitation

*"Nonmelanoma skin cancer" combined with "guidelines" (and synonyms)**
((((nonmelanoma[Title/Abstract]) OR non melanoma[Title/Abstract]) OR NMSC[Title/Abstract])) AND (((((((guideline[Title/Abstract]) OR guidelines[Title/Abstract]) OR recommendation[Title/Abstract]) OR recommendations[Title/Abstract]) OR protocol[Title/Abstract]) OR protocols[Title/Abstract]) OR standard[Title/Abstract]) OR standards[Title/Abstract]); with and without "guidelines" as limitation
Article selection criteria:
Title, abstract, full-text screenings were consecutively and independently performed by two authors (SL and LvV). Relevant articles published up to 15-05-2015 were included. When full-text articles were not available online, the Dutch Central Catalogue (PiCarta) and/or the first author was contacted to obtain the article. Exclusion criteria were: full-text not available, non-English/Dutch/German publications, and guidelines written on one therapeutic option (for instance exclusively on photodynamic therapy or Mohs micrographic surgery). Both guidelines regarding only one type of NMSC (basal cell carcinoma or squamous cell carcinoma) as well as guidelines regarding NMSC in general were included. Only the most recent version of each guideline was included. Inconsistencies between the two reviewers were discussed until consensus could be reached. In case no consensus could be reached a third reviewer (MG) was involved.

Supplementary Table S3: Search strategy and selection criteria used in this systematic review on nonmelanoma skin cancer guidelines *Pubmed version; comparable strategies were used in other databases.

REFERENCES

1 Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet. 2013;381(9868):752-62.

2 United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects: The 2015 Revision. New York: United Nations, 2015. Available from: http://esa.un.org/unpd/wpp/publications/files/key_findings_wpp_2015.pdf.

3 Cherubini A, Oristrell J, Pla X et al. The persistent exclusion of older patients from ongoing clinical trials regarding heart failure. Arch Intern Med. 2011;171(6):550-6.

4 Konrat C, Boutron I, Trinquart L, Auleley GR, Ricordeau P, Ravaud P. Underrepresentation of elderly people in randomised controlled trials. The example of trials of 4 widely prescribed drugs. PloS One. 2012;7(3):e33559.

5 Zulman DM, Sussman JB, Chen X, Cigolle CT, Blaum CS, Hayward RA. Examining the evidence: a systematic review of the inclusion and analysis of older adults in randomized controlled trials. J Gen Intern Med. 2011;26(7):783-90.

6 Rogers HW, Weinstock MA, Feldman SR, Coldiron BM. Incidence Estimate of Nonmelanoma Skin Cancer (Keratinocyte Carcinomas) in the US Population, 2012. JAMA Dermatol. 2015;151(10):1081-6.

7 Musah A, Gibson JE, Leonardi-Bee J, Cave MR, Ander EL, Bath-Hextall F. Regional variations of basal cell carcinoma incidence in the U.K. using The Health Improvement Network database (2004-10). Br J Dermatol. 2013;169(5):1093-9.

8 Eisemann N, Waldmann A, Geller AC et al. Non-melanoma skin cancer incidence and impact of skin cancer screening on incidence. J Invest Dermatol. 2014;134(1):43-50

9 Lewis KG, Weinstock MA. Trends in nonmelanoma skin cancer mortality rates in the United States, 1969 through 2000. J Invest Dermatol. 2007;127(10):2323-7.

10 Linos E, Berger T, Chren MM. Point: Care of potential low-risk basal cell carcinomas (BCCs) at the end of life: The key role of the dermatologist. J Am Acad Dermatol. 2015;73(1):158-61.

11 Fosko SW. Counterpoint: Limited life expectancy, basal cell carcinoma, health care today, and unintended consequences. J Am Acad Dermatol. 2015;73(1):162-4.

12 Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E, editor(s). Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. Clinical practice guidelines we can trust. Washington (DC): National Academies Press; 2011.

13 Lugtenberg M, Burgers JS, Westert GP. Effects of evidence-based clinical practice guidelines on quality of care: a systematic review. Qual Saf Health Care. 2009;18(5):385-92.

14 Kastner M, Bhattacharyya O, Hayden L et al. Guideline uptake is influenced by six implementability domains for creating and communicating guidelines: a realist review. J Clin Epidemiol. 2015;68(5):498-509.

15 Gagliardi AR, Marshall C, Huckson S, James R, Moore V. Developing a checklist for guideline implementation planning: review and synthesis of guideline development and implementation advice. Implement Sci. 2015;10:19.

16 Mutasingwa DR, Ge H, Upshur RE. How applicable are clinical practice guidelines to elderly patients with comorbidities? Can Fam Physician. 2011;57(7):e253-62.

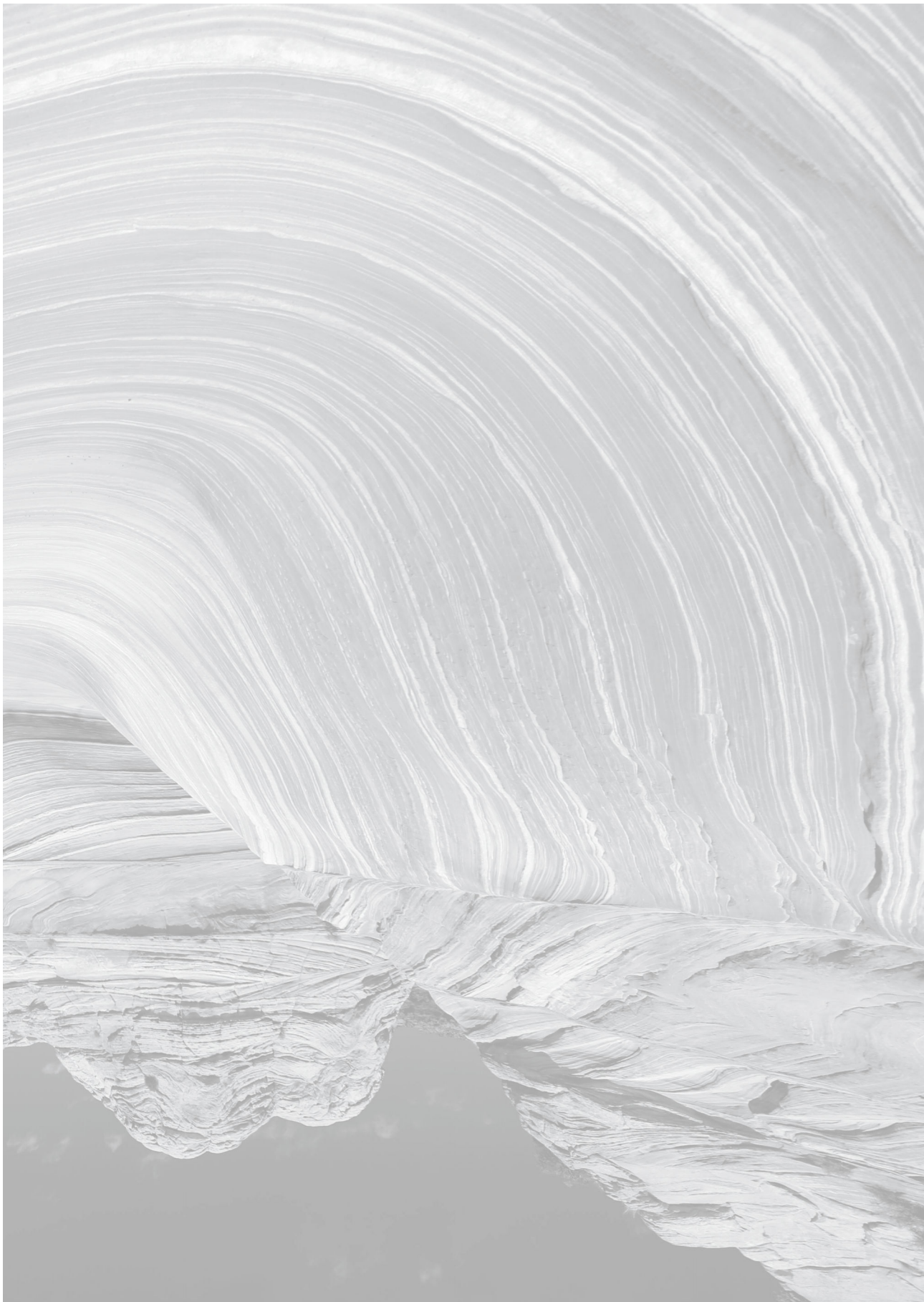
17 Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA. 2005;294(6):716-24.

18 Cox L, Kloseck M, Crilly R, McWilliam C, Diachun L. Underrepresentation of individuals 80 years of age and older in chronic disease clinical practice guidelines. Can Fam Physician. 2011;57(7):e263-9.

19 Hughes LD, McMurdo ME, Guthrie B. Guidelines for people not for diseases: the challenges of applying UK clinical guidelines to people with multimorbidity. Age Ageing. 2013;42(1):62-9.
20 Vitry AI, Zhang Y. Quality of Australian clinical guidelines and relevance to the care of older people with multiple comorbid conditions. Med J Aust. 2008;189(7):360-5.
21 Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. Health Res Policy Syst. 2006;4:15.
22 Linstone HA, Turoff M. Delphi Method: Techniques and Applications. Reading, MA : Addison-Wesley Publishing Company, Boston, USA; 1975.
23 Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.
24 Brouwers MC, Kho ME, Browman GP et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42.
25 Lang T, Altman D. Basic statistical reporting for articles published in clinical medical journals: the SAMPL Guidelines. In: Smart P, Maisonneuve H Polderman A (eds). Science Editors' Handbook, European Association of Science Editors, 2013.
26 Capaciteitsorgaan. Capaciteitsplan 2013, hoofdrapport en deelrapporten 1, 2 en 5. Utrecht, 2013. Available from: http://www.capaciteitsorgaan.nl/Portals/0/capaciteitsorgaan/publicaties/Capaciteitsplan%202013/Capaciteitsplan%202013%20Hoofdrapport.pdf. Dutch.
27 Van Hassel DT, Kasteleijn A, Kenens RJ. Cijfers uit de registratie van huisartsen, peiling 2013. 2014. Available from: http://www.nivel.nl/sites/default/files/Cijfers-uit-de-registratie-van-huisartsen-peiling-jan-2013.pdf. Dutch.
28 Marks R, Barraclough B, Bloom RJ et al. Australian Cancer Network Working Party to revise Management of Non Melanoma Skin Cancer Guidelines (2002). Basal cell carcinoma, squamous cell carcinoma (and related lesions) – a guide to clinical management in Australia. Cancer Council Australia and Australian Cancer Network, Sydney. 2008. Available from: http://www.cancer.org.au/content/pdf/HealthProfessionals/ClinicalGuidelines/Basal_cell_carcinoma_Squamous_cell_carcinoma_Guide_Nov_2008-Final_with_Corrigendums.pdf
29 Stratigos A, Garbe C, Lebbe C et al. On behalf of the European Dermatology Forum (EDF), the European Association of Dermato-Oncology (EADO) and the European Organization for Research and Treatment of Cancer (EORTC). Diagnosis and treatment of invasive squamous cell carcinoma of the skin: European consensus-based interdisciplinary guideline. 2015. Available from: http://www.euroderm.org/edf/index.php/edf-guidelines/category/5-guidelines-miscellaneous
30 Trakatelli M, Morton C, Nagore E et al. Update of the European guidelines for basal cell carcinoma management. 2012. Available from: http://www.euroderm.org/edf/index.php/edf-guidelines
31 Bonerandi JJ, Beauvillain C, Caquant L et al. For the French Dermatology Recommendations Association (aRED). Guidelines for the diagnosis and treatment of cutaneous squamous cell carcinoma and precursor lesions. J Eur Acad Dermatol Venereol. 2011;25 Suppl 5:1-51
32 Dandurand M, Petit T, Martel P et al. ANAES: French National Agency for Accreditation and Evaluation in Healthcare Management of basal cell carcinoma in adults - clinical practice guidelines. 2004. Available from: http://www.has-sante.fr/portail/upload/docs/application/pdf/Basal_cell_carcinoma_guidelines.pdf

33 Breuninger H, Eigentler T, Bootz F et al.. Brief S2k guidelines--Cutaneous squamous cell carcinoma. J Dtsch Dermatol Ges. 2013 Jun;11 Suppl 3:37-45, 39-47.
34 Hauschild A, Breuninger H, Kaufmann R et al. Brief S2k guidelines--Basal cell carcinoma of the skin. J Dtsch Dermatol Ges. 2013 Jun;11 Suppl 3:10-5.
35 Kelleners-Smeets NW, De Haas ER, Beljaards RC et al. Evidence-based richtlijn basaalcelcarcinoom (modulaire update 2014). 2014. Available from: http://www.nvdv.nl/informatie-voor-de-professional-2/informatie-voor-de-professional/richtlijnen-2/. Dutch.
36 Krekels GA, Van Berlo CL, Van Beurden M et al. Richtlijn Plaveiselcelcarcinoom van de huid. 2010. Available from: http://www.nvdv.nl/informatie-voor-de-professional-2/informatie-voor-de-professional/richtlijnen-2/. Dutch.
37 Proby CM, Affleck AG, Bowden P et al. Scottish Intercollegiate Guidelines Network. Management of primary cutaneous squamous cell carcinoma. Edinburgh: SIGN; 2014. (SIGN publication no. 140). [June 2014]. Available from: http://www.sign.ac.uk
38 Motley RJ, Preston PW, Lawrence CM. British Association of Dermatologists. Multi-professional Guidelines for the Management of the Patient with Primary Cutaneous Squamous Cell Carcinoma. 2009. Available from: http://www.bad.org.uk/library-media%5Cdocuments%5CSCC_2009.pdf
39 Telfer NR, Colver GB, Morton CA. British Association of Dermatologists. Guidelines for the management of basal cell carcinoma. Br J Dermatol. 2008;159(1):35-48.
40 Miller SJ, Alam M, Andersen J et al. National Comprehensive Cancer Network (NCCN). Basal cell and squamous cell skin cancers. 2012. Available from: https://www.nccn.org/store/login
41 Pallis AG, Fortpied C, Wedding U et al. EORTC elderly task force position paper: approach to the older cancer patient. Eur J Cancer. 2010;46(9):1502-13.
42 Hurria A, Wildes T, Blair SL et al. Senior adult oncology, version 2.2014: clinical practice guidelines in oncology. J Natl Compr Canc Netw. 2014;12(1):82-126.
43 Wildiers H, Heeren P, Puts M et al. International Society of Geriatric Oncology consensus on geriatric assessment in older patients with cancer. J Clin Oncol. 2014;32(24):2595-603.
44 Mast KR, Salama M, Silverman GK, Arnold RM. End-of-life content in treatment guidelines for life-limiting diseases. J Palliat Med. 2004;7(6):754-73.
45 Jansen J, McKinn S, Bonner C et al. Systematic review of clinical practice guidelines recommendations about primary cardiovascular disease prevention for older adults. BMC Fam Pract. 2015;16:104.
46 Hamaker ME, Jonker JM, de Rooij SE, Vos AG, Smorenburg CH, van Munster BC. Frailty screening methods for predicting outcome of a comprehensive geriatric assessment in elderly patients with cancer: a systematic review. Lancet Oncol. 2012;13(10):e437-44.
47 O'Donovan A, Mohile SG, Leech M. Expert consensus panel guidelines on geriatric assessment in oncology. Eur J Cancer Care (Engl). 2015;24(4):574-89.
48 Garber AM. Evidence-based guidelines as a foundation for performance incentives. Health Aff (Millwood). 2005;24(1):174-9.
49 Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. JAMA. 2014;312(23):2551-61.
50 Puts MT, Santos B, Hardt J et al. Ann Oncol. An update on a systematic review of the use of geriatric assessment for older adults in oncology. 2014;25(2):307-15.
51 Rubenstein LZ, Stuck AE, Siu AL, Wieland D. Impacts of geriatric evaluation and management programs on defined outcomes: overview of the evidence. J Am Geriatr Soc. 1991;39(9 Pt 2):8S-16S; discussion 17S-18S.

7



8

DEVELOPMENT OF A TOOL TO ASSESS GUIDELINE ADHERENCE

A national study on adherence to a basal cell carcinoma guideline; development of a tool to asses guideline adherence

Published as

Borgonjen RJ, van Everdingen JJ, Bruijnzeel-Koomen CA, van de Kerkhof PC, Spuls PI. A national study on adherence to a basal cell carcinoma guideline; development of a tool to assess guideline adherence.
Br J Dermatol. 2015;172(4):1008-13.

SUMMARY

Background Clinical practice guideline implementation may be at variance with actual daily practice, as guideline adherence is a complex process. Feedback regarding adherence is essential to monitor the effect that a guideline has in clinical practice and whether or not the quality of care is raised by implementation.
Objectives Developing a tool for obtaining and giving nationwide feedback regarding adherence.
Methods From February 2010 to June 2013, a 32-item questionnaire was used as an audit tool during committee visits to assess adherence across 37 dermatological centres in The Netherlands. The questions were derived from the recommendations in the Dutch Basal Cell Carcinoma (BCC) guideline. Five selected medical records per dermatologist were audited and the results were discussed with the audited centre. Data were pooled to calculate the compliance with each recommendation across all participating centres.
Results Adherence to recommended actions varied considerably (20.2-100%) across the domains of prevention, diagnostics, treatments and aftercare. Using and reporting surgical margins, giving patient advice, restricting the use of cryosurgery for certain BCCs and reporting on prognostic factors all failed to reach a threshold of 80%. Nonadherence to recommended actions proved to be related to whether or not a dermatologist was directly involved.
Conclusion The findings emphasize the importance of direct feedback to practitioners regarding adherence. Furthermore, together with existing frameworks, the method described could be used by developers in a guideline update to identify and anticipate barriers to successful implementation.

INTRODUCTION

The objective of developing clinical practice guidelines is to improve the quality of care by translating the best available scientific evidence into specific recommendations.[1] To achieve that goal it is critical that a guideline is used and adhered to in daily practice. Developing guidelines and making them available to health care professionals is not enough. To accomplish successful implementation, different factors have to be taken into account. First there are factors relating to implementation strategy and effort, such as dissemination via group meetings, provision of educational materials or computerized decision support.[2-4] Second are factors relating to the guideline itself. The method of development, the level of complexity of recommendations and the strength of evidence all have an influence on the implementation of a guideline.[5-9] Third, the habits of a health care professional can influence adherence, such as a tendency to treat despite the lack of effective interventions, or the willingness to change behavior. Finally, patient beliefs together with organizational, financial (e.g. when recommended therapies are not reimbursed) and cultural factors may also limit guideline adoption.[3,10,11] To see how well complex implementation processes succeed, evaluating adherence to a guideline is pivotal, and can reveal impediments not previously recognized.[12] Moreover, a guideline audit can improve implementation in itself by familiarizing health professionals with its contents and thereby contribute to the quality of care.[13-16] Another reason why evaluating adherence to a guideline is important is the fact that guidelines require regular updating.[17,18] Possible guideline implementation barriers have to be assessed before an update to ensure general acceptance and usage afterwards.[7,19] Previous research on adherence elucidated potential barriers and factors related to successful implementation, and as a result created an initial framework.[3,11,12,15] Few studies, however, have evaluated nationwide adherence to a complete set of recommendations.[20] Previous audits to evaluate the use of dermatological guidelines were mostly focused on a specific part of a guideline and/or assessed the implementation in one center.[21,22] Other audits invited and included only interested dermatologists to participate and/or relied on self-reporting questionnaires.[23-26] This study analyses and discusses the adherence to all recommendations in a basal cell carcinoma guideline, using a method that obtains daily practice feedback on a nationwide level.

METHODS

In 2008 the Dutch Society of Dermatology and Venereology (NVDV) released a guideline for the management and treatment of basal cell carcinoma (BCC; available at www.huidarts.info), because the incidence of BCC is increasing and the management of BCCs represents an important part of the workload of dermatologists.[27] After the BCC guideline was released, authorized and disseminated, all the recommendations in the guideline were transformed into elements of care with their recommended actions (summarized in Table 1). This resulted in a medical audit with 32 yes or no questions with optional textual remarks (Appendix). The questionnaire was divided across domains with two questions on prevention, seven on diagnostics, three for aftercare and twenty questions divided over eight possible treatments. The audit was approved by the visitation committee of the Dutch Society of Dermatology and Venereology (NVDV) for use as a tool during the compulsory five-year visitation for every dermatologist in the Netherlands. Two months before the visitation of a dermatological center supporting staff were asked to identify recent patients with known basal cell carcinoma and randomly pick five medical records per dermatologist. The selected medical records were then audited by an appointed dermatologist in the center concerned. Results were discussed with every dermatologist at the audited center according to the following format:

1. Is there uniformity concerning therapy?
2. Are there deviations from the guideline?
3. Was suboptimal care detected?
4. Does the audit give rise to improvements or alterations in policy?

A report from that discussion and a concept of possible improvements was required to be sent together with the audits to the visitation committee three weeks in advance of the actual visitation. During the visitation the audited medical records were to be made available and were subjected to random checks by the committee. Audit results and the improvement plan were discussed with the committee.

All visitation audits were collected between February 2010 and June 2013. Anonymized data was pooled and the audit results presented as an overall adherence percentage to the recommended action. Adherence below 80% was regarded as non-compliant. In line with accepted standards for clinical record-keeping, non-documentation in the patient's medical record, pathologists report, pathology request form or other available information sources was assumed to indicate that a recommended action had not been fulfilled. In the case of a mixed pathological subtype of BCC, the most unfavorable subtype was reported. When a question box was indecipherable or spoiled, the question was taken as missing data. An

unanswered question was regarded as “unknown”. Missing data and questions that were checked as “not applicable” were not used in the calculation of adherence. The data was analyzed with the Statistical Package for the Social Sciences software (SPSS), using the Chi-square test for checking statistically significant differences when comparing adherence to specific elements of care.

RESULTS

In total, 471 BCCs in patients seen by 103 dermatologists in 37 dermatological centers were audited. 12 patients had more than one treatment in a single visit and were therefore audited for multiple treatments in one audit. 467 patients reported their sex with 55.9% male and 44.1% female. The diagnosis showed a nodular basal cell carcinoma in 24.5% of the patients, a superficial basal cell carcinoma in 11.5%, an infiltrative basal cell carcinoma in 7.6% and in 56.1% the diagnosis BCC was not further specified. Recommended actions with an overall adherence below 80% are highlighted in bold in Tables I and II.

Element of care	Recommended actions	Adherence
Prevention		
Prevention of further excessive exposure to UV radiation	Advice on appropriate sun protective behaviour (n=471) Handing out prevention leaflet (n=471)	**59.7 %** **20.2 %**
Diagnostics		
Histological confirmation BCC	Skin biopsy (n=468)	**63.0 %**
Establishing subtype BCC	Requesting subtype from pathologist (n=462) Reporting subtype to dermatologist (n=461)	**40.3 %** 94.2%
Differentiation between high and low risk BCC	Reporting relevant prognostic factors: -location -size -recurrent BCC? (n=469)	 92.4% **67.3%** **42.4%**
Aftercare		
Adequate follow-up	At least yearly follow-up for high-risk or multiple BCC (n=345)	93.0%
Communication with patient’s general practitioner	General practitioner is informed about diagnosis and therapy (n=468)	94.4%
Patient information	Information about possible recurrence and handing out leaflet (n=468)	**69.9%**

Table 1: Elements of care with recommended actions and adherence

The recommendations are reported as being followed in a majority of possible therapies for BCC, but it is clear that adherence to the recommended actions in case of excision, and partly in case of cryosurgery of the BCC, is below the applied threshold of 80%.

Therapy	Recommended actions	Adherence
Excision (n=353)	Reporting surgical margins: -in medical record -when requesting histopathology Using recommended margins: -BCCs where 3 mm is recommended (example: low-risk BCC) -BCCs where 5 mm is recommended (example: BCC > 2 cm)	 **38.2%** **17.0%** **63.0%** **37.9%**
Mohs surgery (n=35)	Reasoning why treatment option was chosen is known	**74%**
Photodynamic therapy (PDT) (n=44)	Only for primary, superficial BCCs. Reporting of treatment characteristics: -light source, -schedule/dose, -photosensitizer	86% 91% 91% 98%
Cryosurgery (n=13)	Only for nodular or superficial BCCs < 2 cm. Reporting of treatment characteristics: -number of freeze-thaw cycles/freezing time -applied method	**69 %** **23%** 100%
Imiquimod (n=6)	Only for primary, superficial BCCs < 2 cm and not on face. Reporting of treatment characteristics -schedule/dose -adverse events	100% 100% 83%
Curettage and cautery (n=5)	Only for primary BCCs < 0.5 cm with favorable location/subtype	100%
Radiotherapy (=4)	Reasoning why treatment option was chosen is known	100%
Other (n = 19)	Reasoning why treatment option was chosen is known	100%

Table 2: Different therapies with recommended actions and adherence

Comparing adherence to specific elements of care

In cases where explicit patient advice (e.g. on appropriate sun exposure or recurrence) was given or where leaflets were handed out, more surgical margins and prognostic factors were reported and skin biopsies taken (all $p \leq 0.001$). So when elements of care were dependent on a dermatologist there was a significant relation in overall adherence or non-adherence (all $p < 0.01$), whereas a dermatologist-independent factor such as reporting subtype by a pathologist was not significant ($p = 0.24$).

DISCUSSION

Adherence to recommended actions varied considerably (20.2% to 100%) in this study and largely depended on the dermatologist. Using and reporting surgical margins, giving patient advice, restricting the use of cryosurgery for certain BCCs and reporting prognostic factors all clearly failed to meet the applied threshold. Earlier studies mostly found adherence rates around 70%.[28-32] An overall reason adherence rates are somewhat higher in other studies might be an increase in the use of self-reports to estimate actual adherence to guidelines. Self-reports appear to produce gross over-estimation of performance and the extent of bias is often greater than the degree of improvement observed after implementation interventions.[32] Having said that, achieving an overall adherence rate of 100% is not always desirable since a reasonable deviation from guidelines is sometimes necessary to address specific situations where guidance is often lacking, such as in cases of co-morbidity, co-medication and patients with a low life expectancy.[28] In such cases an adherence between 80-90% could be feasible and desirable. For this reason, an 80% adherence rate was used in this study as an acceptable threshold. Some reasons for not reaching 100% adherence in this study were found in the textual comments. Four dermatologists mentioned that they carried out the recommended action but did not document it. This is a limitation in adherence studies because adherence might reflect variation in the adequacy of physician documentation as well as actual variation in adherence. Five centers reported that they arranged with their pathologists to always report the subtype in a pathology report. Such a measure partly explains the low adherence on requesting subtype (40.3%), because making an explicit request becomes unnecessary. The low percentage of adherence to reporting freezing time and number of freeze/thaw cycles (23%) could be explained by bias due to the low number of cases receiving cryosurgery. It could also be a sign that dermatologists find it difficult to standardize this treatment modality. One reason for the low adherence to recommendations regarding patient advice (20.2% to 69.9%) could be the lack of self-efficacy and outcome expectancy. These are barriers often encountered in preventive health

education and counseling, and that often hinder a behavior being initiated and sustained where poor outcomes are – however unjustly – anticipated.[3] The observed related non-adherence in dermatologist-dependent actions could be caused by a lack of agreement with the guideline or the concept of guidelines in general. More probably though there is the habit of previous practice and the lack of motivation to change, or to document, perhaps due to time constraints. The failure to hand out the prevention leaflet may have been down to the fact that nine centers reported that they use their own leaflets or other material available on the internet. It is for the guideline committee to decide in a future review if such behavior constitutes an acceptable alternative.

As emphasized earlier, guideline adherence is a complex process. Several frameworks for assessing possible barriers to implementation underline the importance of the use of daily practice data collection tools, such as audits.[12,15] These frameworks can standardize the reporting of barriers to adherence, as the effectiveness of interventions to improve adherence is dependent not only on the intervention itself but also on the existence and strength of baseline barriers. In addition, feedback about adherence can improve implementation and physician behavior directly by raising guideline awareness and familiarity with its contents.[3,11] Nevertheless, audit and feedback are not enough: the most promising approach to achieve successful implementation is to use a variety of interventions including reminders, educational outreach and provision of other training materials.[2,33] A national audit has the advantage of obtaining adequate sample sizes for reliable generalization and eliminates duplication of effort in audit design. Incorporating audits as part of existing visitation programs is recommended as a measure for reducing costs. The advantage of a system connected with a visitation cycle is guaranteed feedback from all the centers and not just the ones that are willing to participate. Establishing a regular, cycle-based feedback system is interesting for guideline developers, as well as physicians themselves, as local data is directly discussed with peers. Comparing local data with the analysis of the pooled national adherence levels, it is possible to detect local (organizational or cultural) barriers and eventually allows for the benchmarking of quality of care against practice in similar centers. Such a system could very well approximate to the ideal local environment of timely, individualized, and non-punitive feedback to which Hysong et al. refer.[13] Auditing all recommendations in a guideline also identifies possible indicators, facilitating validation and acceptance of indicators since data is most likely already available. Furthermore, continuous evaluation can be used to generate information about when to update a guideline. In the development of a guideline, the scientific literature is sometimes outdated even before the

guideline is implemented. As Shekelle et al. describes, new evidence demonstrating that recommended interventions are inappropriate, ineffective, or superseded by new ones warrants an update of a clinical practice guideline, as does new information on outcomes and performance.[7]
Guideline developers could use the audit information to make less subjective judgments regarding guideline updating, and combine literature search and professional opinion to decide on an update, thus ensuring an efficient use of resources.

Limitations
A limitation regarding the objectivity in this study is the fact that the dermatological center itself was responsible for collecting the data and sometimes fewer than 5 medical records per dermatologist were audited. This could lead to selection bias since the selection process was not checked by the visitation committee. Another question mark hanging over this study is the validity of extrapolating the adherence data to the entire population of patients with basal cell carcinoma seen by a dermatologist. Approximately 30% of all dermatologists in the Netherlands were included so far and future data will disclose if the practices observed in this study can be replicated in all Dutch dermatological centers. An additional discussion point is the choice of a five-year guideline revision period, as Shojania et al. showed that almost 25% of systematic reviews are probably outdated only two years after publication and Shekelle et al. concluded that, in general, guidelines should be re-evaluated at least every three years.[18,24] Further analysis in the future could be interesting to see if behavior in daily practice is altered when a new update of the guideline is implemented, and if a subset of centers or dermatologists can be identified where implementation is carried out quickly and with little effort. Possibilities regarding an automated feedback system including computer-assisted decision making should also be explored.

CONCLUSION
Guideline development can require great effort and expense and the difficulties in persuading professionals to act on evidence-based guidelines are well known. It is essential to gain feedback regarding adherence in daily practice to justify effort and cost, to see if processes are changed, behavior altered, and quality of care finally improved. In this article we described a method for obtaining feedback on adherence to recommendations, using an existing national five-year visitation schedule. The adherence percentages reported in this article varied considerably due to various

possible adherence and implementation barriers. The adherence data and remarks generated have two functions. Firstly, they provide direct feedback to dermatologists and dermatological centers. Secondly, the method described in this article could, together with a framework more focused on factors intrinsic to a guideline like the GuideLine Implementability Appraisal (GLIA)[12], be used by developers of a BCC guideline update to understand and systematically anticipate barriers to successful implementation, and thereby ultimately help improve health outcomes.

APPENDIX

Audit questions per patient:

Prevention

1a. Patient informed about appropriate sun protecting behaviour?

1b. Information leaflet from the Dutch Society of Dermatology and Venereology about sun exposure handed out?

Diagnostics

2. Core biopsy taken prior to therapy?

3. Basal cell carcinoma mentioned in the differential diagnosis on the pathology request form?

4. Request for mentioning the type of basal cell carcinoma on the pathology request form?

5. Histological type of basal cell carcinoma specified in the pathology results?

6. (Photographic) documentation of precise location of basal cell carcinoma?

7. Size reported?

8. Documentation of possible prior treatment for this basal cell carcinoma?

Excision

9a. Surgical margins mentioned in medical record?

9b. Surgical margins mentioned on the pathology request form?

10a. In case of a basal cell carcinoma of ≤ 20 mm: surgical margin of ≥ 3 mm?

10b. In case of a basal cell carcinoma > 20 mm and/or an aggressive type and/or recurrence: surgical margin ≥ 5 mm?

11. Mohs micrographic surgery considered?

Mohs surgery

12. Reasoning why Mohs micrographic surgery was performed mentioned in medical record?

13. Primary superficial basal cell carcinoma?

Photodynamic therapy

14. Lightsource is known?
15. Treatment scheme is known?
16. Photosensitizer is known?

Cryotherapy

17. Superficial or nodular basal cell carcinoma and < 2 cm?
18.a Method of application of cryotherapy is known?
18.b The number of freeze-thaw cycles is documented?
19. The 'open spray' method is applied?

Imiquimod

20. The treated basal cell carcinoma was no recurrence, less than 2 cm in diameter, of the superficial growing type and not in the facial H-zone?
21. Possible to determine treatment scheme and duration of treatment from medical record?
22. Reporting the occurrence or non-occurrence of (local) side effects?

Curettage and cautery

23. Small, superficial or nodular, primary basal cell carcinoma on favourable location?

Radiotherapy

24. Reasoning why radiotherapy was performed is known?

Other treatment

25. Reasoning why a different treatment was applied is known?

Aftercare

26. In case of multiple basal cell carcinoma or basal cell carcinoma in the facial H-zone: follow-up visit took place or is scheduled ≤ 1 year?
27. General practitioner is informed about the diagnosis and treatment of this basal cell carcinoma?
28. Patient (in writing) informed about the possibility of recurrence?

REFERENCES

1 Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 1999;318(7182):527-30.

2 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003;362(9391):1225-30.

3 Cabana MD. Why Don't Physicians Follow Clinical Practice Guidelines?: A Framework for Improvement. JAMA 1999;282(15):1458-65.

4 Balk SJ, Landesman LY, Spellmann M. Centers for disease control and prevention lead guidelines: do pediatricians know them? J Pediatr 1997;131(2):325-7.

5 Greco PJ, Eisenberg JM. Changing physicians' practices. N Engl J Med 1993;329(17):1271-3.

6 Davis DA, Taylor-Vaisey A. Translating guidelines into practice. A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. CMAJ 1997;157:408-16.

7 Shekelle P, Woolf S, Grimshaw JM, Schunemann HJ, Eccles MP. Developing clinical practice guidelines: reviewing, reporting, and publishing guidelines; updating guidelines; and the emerging issues of enhancing guideline implementability and accounting for comorbid conditions in guideline development. Implementat Sci 2012;7:62.

8 Michie S, Johnston M. Changing clinical behaviour by making guidelines specific. BMJ 2004;328(7435):343-5.

9 Shekelle PG, Kravitz RL, Beart J, Marger M, Wang M, Lee M. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. Health Serv Res 2000;34(7):1429-48.

10 Olesen F, Lauritzen T. Do general practitioners want guidelines? Attitudes toward a county-based and a national college-based approach. Scand J Prim Health Care 1997;15(3):141-5.

11 Wakkee M, Lugtenberg M, Spuls PI, de Jong EM, Thio HB, Westert GP, et al. Knowledge, attitudes and use of the guidelines for the treatment of moderate to severe plaque psoriasis among Dutch dermatologists. Br J Dermatol 2008;159(2):426-32.

12 Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak 2005;5:23.

13 Hysong SJ, Best RG, Pugh JA. Audit and feedback and clinical practice guideline adherence: making feedback actionable. Implement Sci 2006;1:9.

14 Foy R, MacLennan G, Grimshaw J, Penney G, Campbell M, Grol R. Attributes of clinical recommendations that influence change in practice following audit and feedback. J Clin Epidemiol 2002;55(7):717-22.

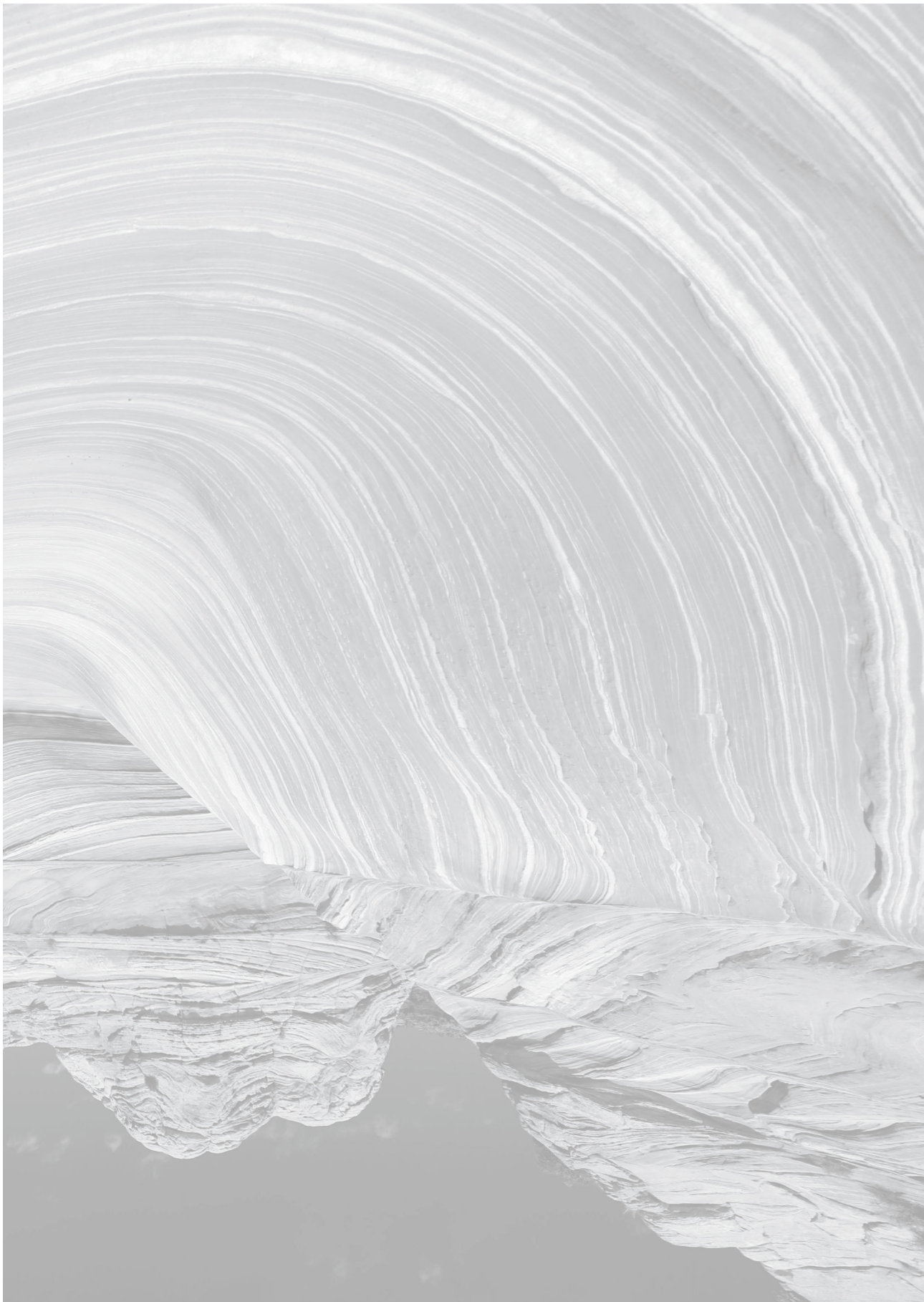
15 Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. Implement Sci 2011;6:26.

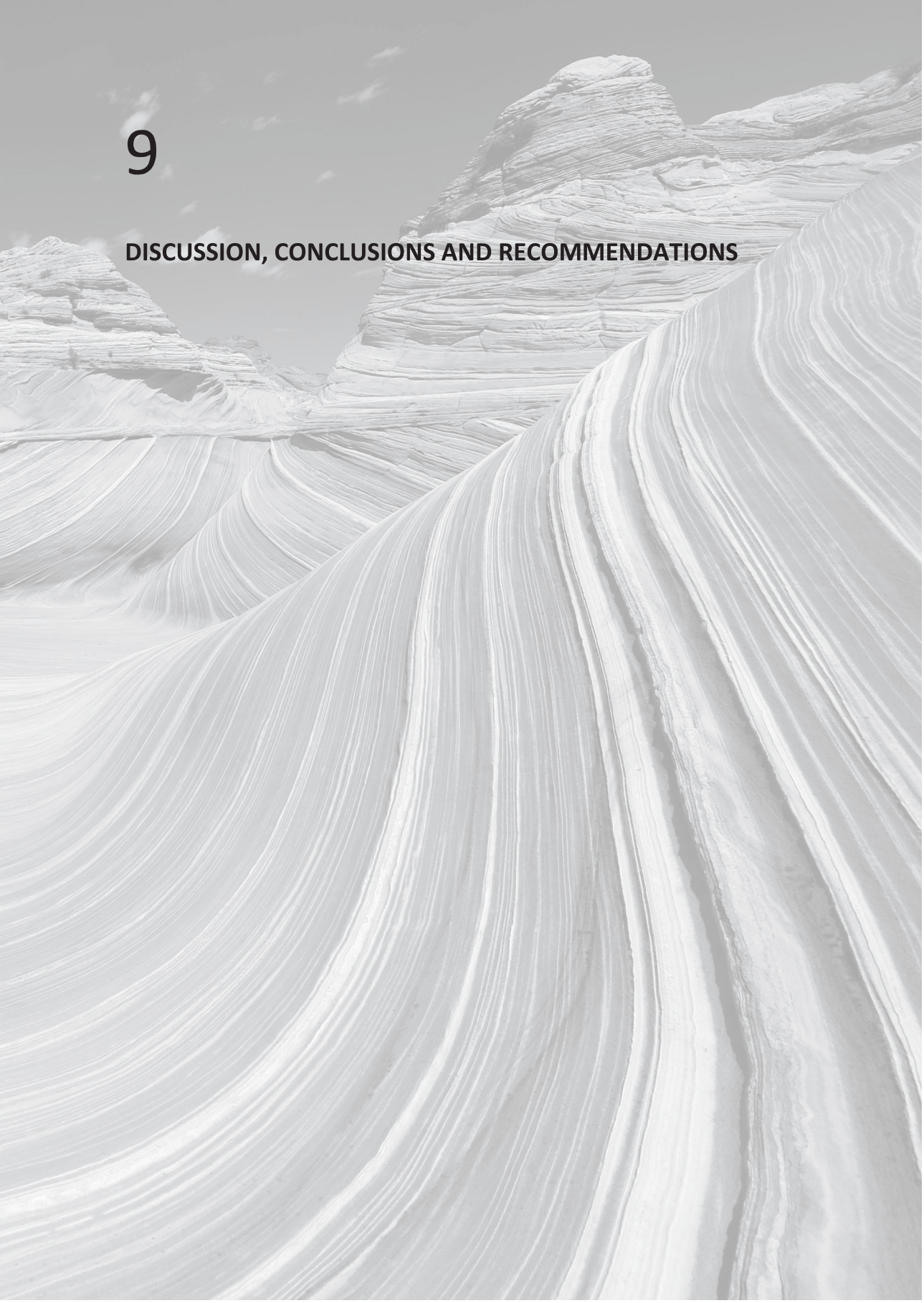
16 Madan V, Walker SL. An audit of the management of tinea capitis using the British Association of Dermatologists' guidelines. Int J Dermatol 2007;46(5):547.

17 Shekelle P, Eccles MP, Grimshaw JM, Woolf SH. When should clinical guidelines be updated? BMJ 2001;323(7305):155-7.

18 Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 2001;286(12):1461-7.

19 van Everdingen JJ. Van consensus naar CBO-richtlijn. Ned Tijdschr Geneeskd 1999;143(42):2086-9.
20 Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, et al. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. BMJ 2004;329(7473):999.
21 Yones SS, Palmer RA, Kuno Y, Hawk JL. Audit of the use of psoralen photochemotherapy (PUVA) and narrowband UVB phototherapy in the treatment of psoriasis. J Dermatol Treat 2005;16(2):108-12.
22 MacDonald A, Berry C, Holmes S. An Audit of the Management of Melanoma Patients at Glasgow Royal Infirmary 1998-2003. Scott Med J 2006;51(1):30-3.
23 Bewley A, Cerio R, Clement M, Hunt S, Lucke T, Ratnavel R, et al. Current application of National Institute for Health and Clinical Excellence (NICE) guidance in the management of patients with severe psoriasis: a clinical audit against NICE guidance in seven National Health Service specialist dermatology units in England. Clin Exp Dermatol 2011;36(6):602-6.
24 Gudi V, Ormerod AD, Dawn G, Green C, MacKie RM, Douglas WS, et al. Management of basal cell carcinoma by surveyed dermatologists in Scotland. Clin Exp Dermatol 2006;31(5):648-52.
25 Raj G, Bell HK. A multi-centre audit on genital Lichen sclerosus in the North West of England. J Eur Acad Dermatol Venereol Epub 2013/04/27.
26 Berends MA, de Jong EM, van de Kerkhof PC, Gerritsen MJ. Dermatologists' adherence to the guideline of the Dutch Society of Dermatology and Venereology with respect to the treatment with methotrexate for severe chronic plaque psoriasis: results from a Dutch survey. Dermatol 2007;215(1):45-52.
27 Flohil SC, de Vries E, Neumann HA, Coebergh JW, Nijsten T. Incidence, prevalence and future trends of primary basal cell carcinoma in the Netherlands. Acta Derm Venereol 2011;91(1):24-30.
28 Burgers JS, van Everdingen JJE. Beyond the evidence in clinical guidelines. Lancet 2004;364(9432):392-3.
29 Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care 2001;39(8 Suppl 2):II46-54.
30 Lugtenberg M, Burgers JS, Besters CF, Han D, Westert GP. Perceived barriers to guideline adherence: a survey among general practitioners. BMC Fam Pract 2011;12:98.
31 Grol R, Dalhuijsen J, Thomas S, Veld C, Rutten G, Mokkink H. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ 1998;317(7162):858-61.
32 Adams AS, Soumerai SB, Lomas J, Ross-Degnan D. Evidence of self-report bias in assessing adherence to guidelines. Int J Qual Health Care 1999;11(3):187-92.
33 Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care 2001;39(8 Suppl 2):II2-45.
34 Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med 2007;147(4):224-33.





9

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

DISCUSSION

The research described in this thesis aimed to explore the quality of dermatological guidelines and how to further improve that quality. Aspects of quality were reviewed focusing on different levels of the evidence based guideline development cycle in order to get more insight into the necessary steps that are required to reach a high quality guideline. Thereby the most important adagio to remember is that guidelines are a means to an end.

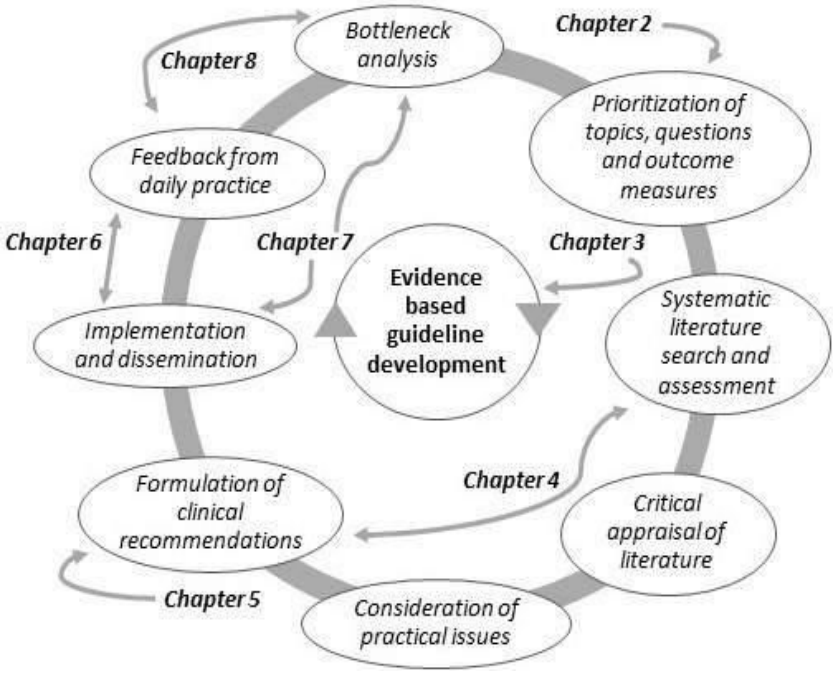


Figure 1: Evidence based guideline development cycle. Arrows represent the parts of the cycle covered in that particular chapter.

In **Chapter 1** a number of relevant research questions were postulated. These questions address various difficulties related to guideline development. The remaining part of this chapter discusses several aspects related to these questions by following the guideline development cycle, after which a number of recommendations are presented to promote high quality dermatological guidelines.

HOW TO ENSURE THE QUALITY OF A GUIDELINE?

To answer that question a definition of a high quality guideline is essential. High quality guidelines are guidelines that deliver relevant support to healthcare professionals and users based on valid scientific research complemented with structured expertise and experience. To guarantee quality, guidelines should be developed within a coordinated program according to the principles of evidence-based medicine. Explicit reporting of the methods and procedures followed will improve guideline quality[1,2] On the other hand a structured approach does not warrant a relevant and valid guideline.[3] Burgers et al. already identified this issue when looking at guideline quality. One has to distinguish between the methodological quality of a guideline or in other words the rigour of development and the quality of the content.[1]

AGREE

The call for rigour of development gave rise to many 'guidelines for guidelines' and several tools to validate guidelines. These validation tools score the various items that should ideally be incorporated into guidelines.[4] The AGREE collaboration succeeded to develop an instrument which incorporated most of the items that dermatological guideline developers find important.[5] The AGREE instrument can act as a guide for guideline developers to take into account all the important matters when creating a guideline. A high AGREE score also guarantees the healthcare professional and health care user that the methodological quality is sound.[6] **Chapter 3** investigated the Dutch dermatological guidelines with the AGREE II instrument by rating 23 AGREE II items across 6 domains on a 7-point Likert scale (1–strongly disagree to 7–strongly agree). Dutch dermatological guidelines achieved high scores in the domains scope and purpose, stakeholder involvement, rigour of development and clarity of presentation. The domain scores in an review of European guidelines were quite similar.[7] The domains applicability and editorial independence received lower scores. The domain applicability pertains to the likely barriers and facilitators to implementation, strategies to improve uptake, and resource implications of applying the guideline. This is discussed in length underneath the subheading 'Implementation and dissemination'.

Competing interests

Editorial independence is concerned with the formulation of recommendations not being unduly biased with competing interests.[8] Rosumeck et al. also recognized a low score in the domain editorial independence when reviewing 38 German dermatological guidelines. Data about financing of the guidelines were included in 12 of 38 guidelines (32 %) only. Conflicts of interest were stated in no more than 7 of the 38 guidelines (18 %).[9] With the introduction and ongoing development of biologics, financial interests in dermatology increased. Probably this boosted possible conflicts of interests too. Clinical practice guidelines have to be trustworthy instruments. Therefore it is crucial to handle competing interests in a strict manner.[10] In current dermatological guidelines details about the views of the funding body and the influence thereof on the content of the guideline are stated insufficiently. Competing interests of guideline development group members are unsatisfactorily recorded and addressed. One possibility to improve (the trust in) editorial independence is a specific discussion about competing scientific, commercial, personal and/or financial interests at the beginning of guideline development. Additionally in case of potential conflicts of interest during the guideline development process abstaining from participation in specific questions is a solution. Nonparticipation in guideline topics of people with known potentially crucial competing interests might be the most rigorous answer.

Preparatory phase of guideline development

The editorial independence mentioned before, touches upon one of the most crucial factors in guideline development, namely who initiates and creates a guideline. In the Netherlands the government holds the professional groups responsible for defining the quality of their healthcare. In evidence based guideline development a working group is installed to create a guideline. Usually, an appointed chairman steers the working group.[11] In **chapter 5** we experimented with the composition of the working group to see if guideline recommendations were altered (discussed more beneath subheading 'Formation of clinical recommendations'). The working group should represent all relevant professional groups.[8,11] This is already vital for future support when implementing a guideline.

Prioritizing a topic

The composition of a (multidisciplinary) guideline working group highly depends on the topic of a guideline. So before a working group is formed, one has to choose a topic. Topics ideally arise from daily practise**.** However, a structured approach to prioritize topics was generally lacking. We learned in **chapter 1** that in the last decades

guideline development has turned into an industry and making guidelines became more complex because of the widening scope and increasing demand for rigour of development. Consequently high quality guideline development is costly and demands resources. Hence one has to prioritize. Moreover, because once a guideline is developed it has to be kept up to date in order to remain valid and useful for the target users.

How to prioritize

Criteria why certain topics are more important than other often remain implicit. **Chapter 2** showed that dermatological guidelines are available for all the topics dermatologists find most important. Widespread, costly topics with significant morbidity and wide variation in current care are often considered for a guideline.[12] However, defining criteria is strongly dependent on the actors involved in the prioritization process.[13] Dermatologists for instance find costs and mortality somewhat less important criteria (**chapter 2**). Making criteria more explicit can facilitate the process of prioritization across stakeholders, but determining cut-off points remains a challenge.[13]

Who should prioritize

C**hapter 3** identified that the domain stakeholder involvement focuses on the extent to which a guideline was developed by the appropriate stakeholders, represents the views of its intended users and targets the right public. Appropriate stakeholders are first of all healthcare professionals dealing with the topic. In **chapter 3** each of the Dutch dermatological guidelines had practising general hospital representation in the authorship and was at least targeted at dermatologists in secondary care. Still, in **chapter 7** additional participation of a geriatrician and/or elderly care physician in the guideline development process was recommended to give appropriate attention to frail older adults. As a professional group it is fundamental to objectively determine the priority of topics in order to establish a correct timeline and allocate funds and other resources to the right project. Nevertheless, other stakeholders have to be involved.[12,14]

Involving healthcare users

What is especially essential for a high quality guideline that is capable of improving healthcare, is that views and preferences of the target healthcare users have been sought. An international movement that promotes the collaboration between healthcare professionals and users to focus on relevant topics is the James Lind

Alliance (JLA). The JLA method is designed to raise awareness of research questions which are of direct relevance and potential benefit to healthcare users and the professionals who treat them.[15] There are several ways to integrate user views into a guideline. Healthcare users can for instance be part of the working group, but such minimal participation might inhibit their decision-making power and contribution. Meaningful active involvement could be a parallel user group trajectory feeding into guideline development.[16] The psoriasis guideline analysed in **chapter 3** had the most promising, but also extensive way. This way includes a patient representative in the working group and a literature search. Since the literature search revealed no conclusive evidence about patient preferences and issues in treatment decisions, a focus group was carried out to identify critical views and preferences. These critical views and preferences were validated through a questionnaire across thousands of patients.[17] Further international guidance on how to involve healthcare users is delivered by a Toolkit on Patient and Public Involvement in Guidelines from G-I-N PUBLIC, a working group of the Guidelines International Network (G-I-N) (http://www.g-i-n.net/working-groups/gin-public). The toolkit collects international experiences and best practices of patient involvement. It aims to support guideline developers by discussing, for example, how to conduct public consultation, how to include qualitative research on healthcare user views in guidelines and how to recruit and support healthcare users in guideline development.

Involving other stakeholders

Other major stakeholders in the Netherlands are hospital managers, healthcare insurers and governmental institutions such as the Ministry of Health, Welfare and Sport and the Dutch Health Care Inspectorate. These stakeholders recognised the possibilities of guidelines for their own agenda. Yet, cooperation is a must for eventual implementation of the recommendations and reimbursement. Including other stakeholders and acknowledging the fact that guidelines are important documents for these stakeholders as well, could also initiate a meaningful discussion about who is responsible for financial support. A discussion that is hampered by the wish to remain independent and responsible for the content of a guideline as a professional group on one hand and the ongoing need for funding to keep up with increasing demands for extensive state-of the-art guidelines on the other hand. Internationally, independent non-profit organisations like Cochrane (formerly known as the Cochrane Collaboration) also try to find ways to effectively collaborate with a broad-spectrum of stakeholders in order to deliver priority Cochrane reviews.[18]

Turning a topic into questions

The topics mentioned in **chapter 2** are very broad and translation into questions is necessary to achieve guidance where it is needed. A clinical practice guideline hopes to improve healthcare by answering specific questions. It is therefore helpful to structure a guideline in terms of answerable questions. A question should include a target population, an intervention or exposure, possibly a comparison, an outcome and a health care setting or context.[8] Nevertheless, when a question arises one has to determine whether a guideline is the appropriate instrument to achieve an answer. Guidelines are needed when there is uncertainty. If the answer is clear beforehand, guidelines do not have additional value compared with a review or textbook. By focussing on the relevant questions budget and effort can be saved.

Determining relevant outcomes and outcome measures

Once the appropriate questions are specifically formulated a guideline working group has to consider the relevant outcomes and outcome measures. Again involving healthcare users is crucial to introduce treatment goals that illustrates and ensures their optimal care. Research on the effects of treatments frequently overlooks the shared interests of healthcare users and professionals and fail to address important questions that provide meaningful answers.[15] The lack of patient reported outcomes and outcome measures (PRO(M)) is also acknowledged in several dermatological publications. A review by Heinl et al. found that only about one fifth of eczema trials include a quality of life measure as outcome. Many different instruments were used, limiting the possibilities of comparing and synthesising individual trials' findings.[19] Luckily, there is increasing attention for patient reported outcomes and outcome measures (PRO(M)). An initiative to integrate PROM in trials and reduce heterogeneity is COMET. The COMET (Core Outcome Measures in Effectiveness Trials) initiative promotes the development and application of agreed standardised sets of outcomes, known as core outcome sets (COS). These sets represent the minimum that should be measured and reported in all clinical trials of a specific condition. Thus making it easier for the results of trials to be compared and combined.

Dermatological COS are developed for eczema and vitiligo.[20,21] More COS with PROM are underway, for example in hidradenitis suppurativa and rosacea.[22,23] Stimulating and important herein is the foundation of the Cochrane Skin Group - Core OUtcome Set INitiative (CSG-COUSIN) in 2015. Its mission to develop and implement core outcome sets in dermatology in order to improve and standardise outcome measurement in clinical trials and to make trial evidence more useful.[24]

GRADEing outcomes

Before the introduction of GRADE, studies where analysed and rated as a whole. The Grades of Recommendation, Assessment, Development and Evaluation Working Group (GRADE Working Group) has developed a system for grading the quality of evidence and in GRADE outcome measures are analysed across individual studies.[25,26] The GRADE system requires an assessment of the quality of a body of evidence for each individual outcome. Thus it is imperative to prioritize outcome measures beforehand.

In **chapter 4** GRADE was used to summarize the available literature on off-label use of cyclosporin in dermatology. A systematic review is more useful by considering all outcomes that are important or critical to healthcare users for decision making. In addition, they may require consideration of outcomes that are important to the other stakeholders mentioned before. If evidence is lacking for an important outcome, this should be acknowledged, rather than ignoring the outcome. GRADE specifies three categories of outcomes according to their importance for decision-making; critical, important but not critical and of limited importance. The relative importance given to health outcomes should reflect the perspective of those who are affected and input from healthcare users is important in prioritization. This is recognized in **chapter 7** were integration of frailty-related items within clinical practice guidelines is advocated to stimulate a more holistic, personalized and patient-centred care instead of the sometimes encountered tumour-centred care.

To facilitate ranking of outcomes according to their importance guideline developers may choose to rate outcomes numerically on a 1 to 9 scale (7 to 9 – critical; 4 to 6 – important; 1 to 3 – of limited importance) to distinguish between importance categories. Critical and important outcomes shape guideline recommendations, the third will not. Ranking outcomes by their relative importance can help to focus attention and save time and money by discarding non relevant outcomes beforehand. Nonetheless, in **chapter 4** it turned out that for most of the questions evidence was lacking. The relative importance of outcomes measures were reconsidered in light of this lacking evidence. We chose to consider all the (surrogate) outcomes in order to provide the guideline working group with all the possible information to weigh the desirable versus undesirable outcomes. In case of surrogate outcomes it was acknowledged that the surrogates were a substitute for more important outcomes. This in turn led to the necessity to rate down the quality of the evidence because of indirectness (table 1).

Systematic literature search and selection

After the questions in PICO format are set and the appropriate outcomes and outcome measures are determined, a systematic search has to be carried out. Searching the right databases for existing scientific evidence is paramount. Methodological experts in the form of clinical librarians are indispensable in this process. Searching for existing guidelines in databases like the Trip Database, National Guidelines Clearinghouse or Scottish Intercollegiate Guidelines Network is obligatory. Also disease specific databases such as the GREAT database (Global Resource for EczemA Trials; www.greatdatabase.org.uk) are useful. Other databases to consider are trial databases. They can help guideline developers identify recent progresses.

International collaboration and adaptation

In **chapter 7** we identified many guidelines on non-melanoma skin cancer developed by different organizations. Earlier studies recognized this potential duplication of effort.[27] International collaboration is a way to minimize (double) efforts and to maximize efficient use of resources by sharing some elements of the expensive and time-consuming process of guideline development.[28] The founding of the Guidelines International Network made sharing knowledge and data easier.[29] In the field of dermatology herein the Cochrane Skin Group and the European Dermatology Forum are noteworthy.[30,31]

But successful international collaboration, in order to use an existing guideline, has a few conditions that have to be met. First of all, the search strategy and selection criteria have to be available in full.[32] With an available search strategy and selection criteria it is possible to repeat a search to update the scientific literature or to see if relevant databases or populations to answer the question were missed. Second the methodological quality of a guideline has to be sufficient. AGREE II is a suitable instrument to rate a possible useful guideline.[6,8] Third the adoption of a single high quality guideline system can make comparison across guidelines easier. The GRADE system seems to be promising since it is embraced by a lot of (inter)national organisations.[25,33]

However, the eventual recommendations still have to be adjusted to the local cultural and organizational context, even if the evidence base is the same. Guidelines produced in one setting or country may not necessarily be appropriate for another.[34] The former collaboration for ADAPTation of Existing guidelines (ADAPTE)) have developed a manual consisting of 24 steps to make a guideline fit for use after adaptation.[35] GRADE also offers a stepped model for adaptation.[36] In order to promote international collaboration and reduce costs and efforts Dutch guidelines

should have an English summary and preferably English evidence tables as well. This prevents the fact that Dutch guidelines are not being found by other organisations worldwide in international databases, a problem encountered in **chapter 3**.

Critical appraisal literature

Once all the relevant literature is selected or adapted from existing guidelines the critical appraisal begins. A systematic review of the evidence should be conducted for every outcome that is considered critical and important. One of the methods to create a guideline in **chapter 5** was to review the literature by a trained epidemiologist. This saved costs and guarantees a uniform way to summarize the literature and appoint a level of evidence, because the comparison between working groups in **chapter 5** showed several differences. The resulting systematic review can be published internationally and thus can offer an incentive for PhD students to carry out the labour. In this, dermatological Cochrane systematic reviews offer a reliable robust review method.[30]

Shortcomings in literature

The heterogeneity and lack of evidence encountered in **chapter 4** is common when appraising the available literature. Reporting standards are a way to reduce these shortcomings. The EQUATOR (Enhancing the QUAlity and Transparency Of health Research) Network is an international organisation founded in 2006 that promotes transparent reporting and wider use of robust reporting standards.[37] The EQUATOR network seeks to minimise avoidable waste of financial and human investments in health research projects. It does so by maintaining and further developing an online collection of health research reporting standards (http://www.equator-network.org/library/). The AGREE instrument is the reporting standard for clinical practice guidelines. The RIGHT (Reporting Items for practice Guidelines in HealThcare) statement is a recently derived 22 item checklist to further assist in reporting clinical practice guidelines.[38] Furthermore, outcome reporting bias (the results-based selection for publication of a subset of the original recorded outcome variables), is a problem in randomized trials.[39] Last but not least, we already mentioned the importance of trial registers to discover publication bias. International databases of prospectively registered systematic reviews like PROSPERO aim to reduce the opportunity for reporting bias by enabling comparison between a completed review with what was planned in the protocol.

Presentation of evidence

The GRADE method depicts the results from the critical literature appraisal in evidence profiles; just as in **chapter 4**. These profiles and the Summary of Findings tables offer a detailed assessment of the underlying confidence in an estimate of effect by outcome and is then combined with an actual analysis of what the effects are. Members of the Cochrane GRADEing method group conducted research on the applicability and the presentation of evidence. Probably the most successful project was DECIDE, a 5-year (2011-2015) EU co-funded project by the GRADE Working Group. Key DECIDE tools developed or improved in the project are a GRADE software app and guideline development tool (GRADEpro), GRADE Evidence to Decision frameworks and interactive Evidence to Decision tools and Summary of Findings tables (http://www.decide-collaboration.eu/). All aimed to reduce effort.

GRADEing evidence

The GRADE method forces to meticulously asses the available literature.

Those who review the evidence will grade the confidence in the estimates of effect of a body of evidence for each outcome in four categories; high-, moderate-, low-, or very low-quality evidence. Randomization is considered the best method to protect against bias and confounding and the initial quality of a body of evidence from randomized control trials usually starts as high quality. Still there are 5 factors that lower the quality and, usually, for observational studies, 3 factors that can increase the quality (method section **chapter 4** and Table 1). GRADE addresses many of the perceived limitations of other models of evidence evaluation. Publication bias for instance is one of the reasons for downgrading.[40] Another limitation was the tendency to power a study for the main outcome only and not for all the subsequent sub-analyses. In this case GRADE allows to downgrade because of imprecision.

GRADE shows methodological flaws within the component studies, consistency of results across different studies, and generalisability of research results to the wider patient base.[33,41] However, problems like the availability of randomized controlled trials for new or pharmaceutical industry sponsored treatments and the lack thereof for older treatments or investigator driven research remain.

GRADE system	
Type of evidence	Randomized trial = high Observational study = low Any other evidence = very low
Decrease* grade if	- Important inconsistency - Uncertainty about directness - Imprecise or sparse data - High probability of reporting bias - Limitation to study quality
Increase grade if	- Strong evidence of association - significant RR of > 2 (< 0.5) based on consistent evidence from two or more observational studies, with no plausible confounders (+1) - Very strong evidence of association—significant RR of > 5 (<0.2) based on direct evidence with no major threats to validity (+2) - Evidence of a dose response gradient (+1) - All plausible confounders would have reduced the effect (+1)
*Each quality criterion can reduce the quality by one or, if very serious, by two levels.	

Table 1: Factors to down- or upgrade according to GRADE. RR = relative risk.

Consideration of practical issues

A guideline is not a systematic review. Subsequent translation is necessary to deliver meaningful recommendations. Guidelines offer the opportunity to include balanced thoughts behind the recommendations. Issues that need to be considered are the quality of the evidence from the critical appraised literature, the balance between benefits and down sides, values and preferences from health care users and professionals and use of resources. Therefore, variation in recommendations between guidelines might be wishful and only reflecting differences in available resources, organizational aspects and disease perception. Still the reasons for this variation should be made explicit in the guideline. Moreover, because as noted in **chapter 5** differences can also reflect individual opinions and expertise. Individual group members can play a role in setting textual differences. This emphasises the importance of a rigorous way to talk about considerations beyond the evidence.[42] Detailed information about other considerations is a better alternative to nonspecific recommendations that could potentially harm patients.[43] Grading of Recommendations, Assessment, Development and Evaluations (GRADE) is now widely seen as the most effective method of linking evidence-quality evaluations to clinical recommendations.[33]

The resulting transparancy also enhances confidence from other stakeholders in the quality of guideline recommendations. Especially the balance between benefits and

down sides and the resources implied receive ever more interest from governmental organisations and healthcare insurers. A guideline working group has a task to promote the most cost effective care. In order to be clear, prices, organisational impacts and benefits of the different options have to be given. Increasingly experts in cost-effectiveness analysis and health technology assessments are needed to deliver information when literature is lacking. Yet, decisions to reimburse a treatment or procedure are not in the scope of a guideline or guideline working group.

Formulation of clinical recommendations

Next recommendations are formed using the evidence from the critical appraised literature, the balance between benefits and down sides, values and preferences from health care users and professionals and use of resources. Recommendations offer an answer to the questions raised. In **chapter 5** we used the perceived level of obligation to follow a deontic term to analyse differences in recommendations. A comparison of the three guideline development methods showed significant but varying correlations between the recommendations formed by the group that followed a traditional evidence-based approach with multiple sessions, one that followed a 2-day conference method and a third method consisting of a systematic evidence report and recommendations made by a single dermatologist. The methods all had pros and cons. The standard multiple meetings method offers time to adapt and reconsider texts. The conference methods presents a possibility for direct discussion and was time- and cost-efficient. The single dermatologist method is quick but is more susceptible to bias and impeded acceptance.[44]

The GRADE method used in **chapter 4** eventually offers 4 ways to grade a recommendation. Two directions of the recommendation (either for or against an intervention or diagnostic test or strategy) and a strength of this recommendation by either determining that it is a strong or a conditional/weak recommendation. It is recommended to formulate recommendations in a clear and unambiguous way using standardized wording, such as using the term recommend for strong recommendations and suggest for conditional or weak recommendations or other terminology such as the “should” and “may” in **chapter 5**.

Broad versus specific recommendations

Even if grading a recommendation and using standard wording is a practical measure, the impact of a recommendation often differs in other ways. For instance, a more specific recommendation with information about dosage is more likely to be followed

than a broader recommendation.[45] Discussion about this matter among members of a guideline working group is advisable.
A related problem is the implementation of international guidelines on an institutional level. As we have discussed before, sharing time and resources is an incentive to make guidelines and recommendations sufficiently broad. The pitfall of the increased attention from other stakeholders for guideline recommendation is also to make vague recommendations that leave considerable room for interpretation. Guideline developers often fear that a more strict formulation confines health care professionals. Probably also due to the growing value the Dutch Healthcare Inspectorate and Medical Disciplinary College impute to guidelines. This dilemma can partially be solved by highlighting the definition of a guideline, a non-binding document with the goal to support and improve healthcare. Thus a guideline is not a collection of minimum standards. Furthermore a guideline is especially of value where there is room for improvement and is not a substitute for a standard care protocol definition. Another solution would be to be even more transparent and portray the goal of a recommendation. Is it meant to be (minimum) norm or a strive worthy goal? This prevents illegitimate use of a guideline recommendation as a (safety) norm for legislation and law enforcement.

ARE GUIDELINES DOING WHAT THEY INTEND TO DO?

A high quality guideline is useless and a waste of time and money, when it does not deliver the support and improvement it intends to do in daily practice. Thus guideline development is not an aim in itself, but should be seen as part of a quality circle in which implementation and evaluation also play an important role. The second part of this thesis focussed on this essential part of the cycle. As noted previously in **chapter 1,** it is not easy to answer the above mentioned question. A systematic review evaluating the effects of Dutch evidence-based guidelines showed that guidelines can be effective in improving the process and structure of care. However, high variation in effects across recommendations suggested that implementation strategies tailored to single recommendations are needed for relevant healthcare improvements.[46] Nevertheless, just as in the first part of this thesis there are step that can be taken to endorse guideline quality and usage.

Implementation and dissemination

Implementation and dissemination are key actions to achieve the healthcare improvements guidelines strive for.[46,47] In **chapter 3** we found that the domain

Applicability received a lower score. This emphasises that implementation of a guideline often does not receive the attention it needs, as discussed before in **chapter 1**. The so called Guideline for Guidelines developed by a group of experts regarding guideline development already advocated the attention for implementation during the entire guideline development cycle.[48] Besides, if implementation of guideline recommendations is to be successful, several strategies have to be used.[49] Consequently implementation already begins in the early phases of guideline development. For instance when prioritizing a guideline topic, since implementing a much wanted guideline is easier. Also including the appropriate professional groups probably already enhances support for a future guideline.

Barriers to implementation

The effectiveness of interventions to promote implementation is dependent not only on the intervention itself but also on the existence and intensity of baseline barriers. By making use of frameworks to assess barriers in advance, guideline developers can revise a recommendation before it is disseminated.[50] As we have learned before, barriers can come in many shapes and sizes. They can be intrinsic to the guideline itself. We already discussed the wording of a recommendation, but the form of a guideline document is another possible barrier. Barriers can also be healthcare professional related or extrinsically related to financial and organizational aspects.[45,46]

Guideline derivatives to promote implementation

In the introduction of this thesis we mentioned that originally guidelines were developed as long textual documents. One of the problems with presenting guidelines as textual documents to healthcare professionals is that the professional must decide whether consultation of a guideline is necessary.[51] Nowadays a guideline should be a collection of recommendations tailored to healthcare users and professionals in different settings. By now nearly all Dutch dermatological guidelines include a leaflet for healthcare users and key recommendations are easily visible in a summary (**chapter 3**). Additional translation into decision aids might empower healthcare users even more and could promote shared decision making. For healthcare professionals an accompanying flowchart can ease the translation of a guideline into a local protocol. Mentionable is the MAGIC project, a multi-layered presentation format for clinical practice guidelines and electronic GRADE-based decision aid prototype (MAGICapp;www.magicapp.org) and the Dutch app builders everywhereIM (everywhereim.com). Other helpful products are educational, audit (**chapter 8**) and computer assisted tools (**chapter 6**).

It is clear that guideline implementation is a complex process. Eventually an implementation plan might be able to shed light on all the actors and factors.[48] An implementation plan depicts possible barriers beforehand in a transparent way, in order to develop a combination of tactics to tackle them. Cooperation between stakeholders is essential to deal with potential organisational and financial barriers. The most hopeful combination of tactics to promote implementation included audit and feedback, reminders and education.[49] An audit such as the one used in **chapter 8** can improve implementation by raising familiarity with the guideline and more important its contents.[45,50,52] By integrating audits together with other aspects of quality enhancement and safeguarding, like visitation and continuous medical education, costs and efforts might be reduced. Furthermore in such an environment of constant recaps through various interactive ways guideline usage may be fortified even more.

Guidelines are also useful as educational tool in continuous medical education. Development of a guideline based online test might raise awareness, familiarity and implementation of guideline recommendations. The test should be accredited and part of the continued medical education of a dermatologist. Feedback from the scientific rationale could be incorporated in the test as well. Further research is needed in this matter.

External comments

Subsequently to internal debate about implementation, guideline recommendations have to be exposed to a broader audience then the guideline working group. Subjecting guideline recommendations to comments from peers, healthcare users, (methodological) experts and other stakeholders is both a way to encourage implementation and a way to generate feedback. It can counter physician related barriers like 'guidelines represent the opinion of a limited group of colleagues'.[45] A description about how comments are handled is valuable. Formalized ways to receive comments are online fora or pilot testing of a guideline. Big national conferences are often costly and time consuming and are probably less suitable. Comments can give even more insight into potential barriers for implementation. Where possible these barriers should be addressed before further dissemination.

Feedback from daily practice

Feedback is needed in every Deming plan-do-check-act cycle. A worthwhile open way to produce feedback is an online commentary option for guideline users. Comments from guideline users could eventually lead to a decision to revise a guideline.

When to revise a guideline

In **chapter 3** it was noted that a precise revision procedure was often missing. In order to remain trustworthy, valid and reflect state-of-the art healthcare a guideline has to be up to date and guideline working groups have to discuss a valid and quick revision procedure. An update of the underlying scientific literature is sometimes needed even before the guideline is implemented.[53] Shojania et al. similarly noticed that up to 25% of systematic reviews are outdated 2 years after publication.[54] **Chapter 8** states that guideline developers could also use audit information to make less subjective judgments regarding guideline updating. An example is the low adherence on requesting the subtype of a basal cell carcinoma due to arrangements with pathologists in **chapter 8.** An audit can thus identify possible obsolete items that warrant updating. Accordingly new evidence in combination with information on outcomes and performance decides if an update is necessary.[55]

Partial revision of a guideline

The GRADE approach encouraged in this thesis forces guideline developers to pose specific questions with specific outcomes. These detailed questions are making it possible to partially revise a guideline. The search strategy can be used per question to automatically monitor new evidence in databases. An automatically set (email)alert delivers new literature to an appointed person. The working group or a professional group can also make the decision to create an online environment where considerations about (partial) revisions are directly visible for everyone. Partial revisions of a guideline have the potential to save a lot of money and efforts. Thus ensuring that keeping abreast the existing body of guidelines is manageable.

Guidelines in the electronic health record system

There are also other more formalized ways to receive feedback. In **chapter 6** we experimented with incorporation of guideline recommendations in the electronic health record system. Many organizations are developing computer-based guidelines as well as guideline based decision support systems.[56] Modern electronic health record systems offer data-mining opportunities to facilitate implementation and generate information to improve clinical processes. Using electronic health record

data in real-time for decision support and process automation has the potential to both reduce costs and improve the quality of patient care.[56-58] This is in concordance with the findings in **chapter 6**. In **chapter 6** a passive reminder was less effective in improving adherence to recommended laboratory testing then active decision support. We talked about the possibility of a physician related barrier such as inertia of previous practice as a cause. Also, computerized guidelines should enters daily practice through normal workflow screens, impose no specific burdens of its own and avoid pop-up screens. The effect of a physician related barrier stresses the need for a variety of implementation and feedback approaches discussed in the paragraph above. Incorporating guidelines in the electronic health record environment is not enough. Perceived usefulness of computerized clinical practice guidelines and social influence are critical factors influencing physicians' intention to use computerized clinical practice guidelines.[59] To overcome such barriers ideally a non-punitive environment should exist. In an environment like that physicians can be confronted with their behaviour in relation to others around them. Such automated digital feedback about individual clinician behaviour turned out to be effective.[52]

Physician adherence

Lasting attention to the supporting non-binding role of a (computerized) guidelines is vital. Physician related barriers like 'guidelines oversimplify medical practice' and 'guidelines challenge the autonomy of healthcare professionals' are always existing. In **chapter 8** varying and lower adherence mainly occurred in items which were dependent on a dermatologist to take action. However, achieving an adherence rate of 100% is not desirable since a reasonable deviation from guidelines is necessary to address specific situations where guidance is often lacking (e.g. co-morbidity, co-medication and patients with low life expectancy).[60] Awareness for deviations and documenting them is necessary, because another study by Lubeek et al. showed that in case of a deviation from a guideline, a reason for this deviation was stated in only 4.5%.[61]

Bottleneck analysis

Eventually all information gathered during and after implementation feed into a new guideline cycle. A bottleneck analysis is necessary to arrange the input. A bottleneck analysis might result in a total guideline revision, a partial guideline revision, a change in scope or addition to a guideline, a new guideline or a change in the strength or wording of a recommendation. A bottleneck analysis can also identify gaps in current knowledge. Gaps that in an ideal world would become part of the research agenda as

well. In **Chapter 7** the bottleneck analysis found several important items for medical decision making in frail older adults with non-melanoma skin cancer. Current integration of these items in clinical practice guideline is limited. Broadening the scope of a guideline is a dilemma. Integration of more frailty-related items might lead physicians to feel more empowered to deviate in the best interest of a patient.[62] On the other hand expanding a guideline requires resources. The items in **chapter 7** were ranked and limited life expectancy, treatment goals other than curation, comorbidities and tumour characteristics are items that should be integrated in medical decision making in every non melanoma skin cancer patient. Thus prioritization remains obligatory. An utopian all-embracing guideline could cover all possible questions, populations and circumstances. This is not only impossible but also not desirable. Guidelines should remain non-binding exemplary documents that offer support in a complex environment.

THESIS LIMITATIONS, EXTRAPOLATION, IMPLICATIONS AND FUTURE RESEARCH

Evidence based guideline development remains closely linked to evidence based medicine in general. Hence, it is important to place continuing research regarding evidence based guideline development in the bigger picture of advances in evidence based medicine.[63] A couple of limitations to this thesis can hamper extrapolation into this bigger picture. In **chapter 2** topic prioritization was performed by healthcare professionals only, in this case dermatologists. Likewise criteria in **chapter 7** were determined using an expert group without direct input from healthcare users. As a framework the results from this studies are definitely useful, but integrating preferences from healthcare users is necessary in the future to keep updating and improving once procedure.[64] The limited use of research publications on patient preferences in healthcare policy-making is also noted by Utens et al. Their qualitative study among Dutch stakeholders shows challenges on how to integrate evidence on patient preferences in different decision contexts.[65] Furthermore, evidence about user preferences is sometimes lacking. For example, there is a scarcity of data about non melanoma skin cancer and patient reported appearance concerns.[66] Therefore, addressing the mismatch between what researchers want or are stimulated to research, and the practical information that is needed day-to-day by patients and health professionals is vital.

Extrapolating the findings in **chapter 6 and 8** could be difficult. Research focussing on professional behaviour and computerised decision support is often performed in a well-controlled academic research setting with a limited number of decision points.[67]

In a recent Dutch study, implementation of an audit and feedback strategy was tested on both test ordering and prescribing behaviour in 'normal' primary care. However, they did not see a decrease in volumes of tests ordered or of drugs prescribed.[68] Barriers that can hinder large-scale use of computerized decision support are irrelevant alerts, too high intensity of alerts, a lack of flexibility and learning capacity of the system, a negative effect on patient communication, and the additional time and work it requires to use the system.[67] Consequently, using the possibilities of integration into modern electronic health record is essential.

This thesis touches upon two fields of research. On one hand research about guidelines in general. The Guideline International Network and the Dutch GENEVER alliance for excellent guidelines (GEzamenlijk NEderlands Verbond voor Excellente Richtlijnen) provides ongoing support and collaboration, assisting members in reducing duplication of effort and promoting best practice.
On the other hand research methods like split face trails and patient reported outcomes that justify the many psychological aspects of skin diseases are unique to dermatology. Dermatology specific research is needed. The Cochrane Skin Group provides assistance according to the principles of evidence-based dermatology.[69]
The cyclic character of this thesis is no coincidence. Guideline development is regarded as a cyclic process that is never finished. A guideline and its spin off products is also called an evidence ecosystem. Completing a cycle from primary studies into systematic reviews, translation into guidelines and into feedback generating spin offs like support in electronical health records systems and decision aids for shared decision making, must be both timely and methodologically trustworthy.[70] As with every ordinary living ecosystem, it is a continuing challenge.

THESIS RECOMMENDATIONS

Preparatory phase of guideline development

-Create guidelines that cover top priority topics.

-Report the procedure to prioritize guideline topics.

-Make prioritization criteria explicit to facilitate the process of prioritization across stakeholders.

-Use AGREE II as a guide when creating a guideline.

-Add a specific discussion about competing scientific, commercial, personal and/or financial interests at the beginning of guideline development.

-Report the results of the above mentioned discussion and add individual conflict of interests statements.

-Ensure that the guideline development group represents all relevant professional groups.

-Integrate the views and preferences of target healthcare users into a guideline.

-Structure guideline questions according to PICO format.

-Further prioritize questions into relevant (patient reported) outcomes and outcome measures.

Systematic literature search and selection

-Methodological experts in the form of clinical librarians should be part of designing a systematic literature search.

-Searching for existing guidelines in guideline databases is obligatory.

-The search strategy and selection criteria should to be reported completely.

-Use AGREE II to rate the (methodological) quality of a useful existing guideline.

Critical appraisal literature

-Use an uniform way, such as GRADE, to summarize and appraise the literature.

-Add an English summary and English evidence tables to Dutch guidelines to prevent duplication of efforts.

Consideration of practical issues

-Use GRADE to translate and link evidence-quality evaluations to balanced clinical recommendations.

Formulation of clinical recommendations

-Look for alternatives like a 2-day conference method to formulate recommendations in order to speed up the process and reduce costs.

-Formulate recommendations in a clear and unambiguous way using behaviourally specific language.
-Formulate recommendations as specific as possible and specify the goal of a recommendation.

Implementation and dissemination
-Use a combination of implementation tools.
-Frameworks to assess and address implementation barriers in advance should be used to conceive an implementation plan with solutions.
-It is recommended that guidelines should provide spin off products like a summary, decision aids or other guideline derived information for health care users or a flowchart to facilitate uptake in local electronic health record systems.
-Use a guideline derived audit to raise familiarity with guideline contents and to assist in peer feedback.
-Development of guideline based (online/interactive) continuous medical education is recommended.

Feedback from daily practice
-Create guidelines with a cyclic modular guideline maintenance programme.
-An online commentary option for guideline users should be encouraged.
-Add a detailed and clear revision procedure to every guideline.
-The search strategy should be used per question to automatically monitor new evidence in databases.
-Using electronic health record data in real-time for guideline based decision support is recommended.
-Address barriers to guideline based decision support.

Bottleneck analysis
-A bottleneck analysis is recommended to organize feedback and identify knowledge gaps.

REFERENCES

1 Burgers JS. Guideline quality and guideline content: are they related? Clin Chem 2006;52:3-4.

2 Burgers JS, Cluzeau FA, Hanna SE, Hunt C, Grol R. Characteristics of high quality guidelines: evaluation of 86 clinical guidelines developed in ten European countries and Canada. Int J Technol Assess Health Care 2003;19:148–57.

3 Watine J, Friedberg B, Nagy E, Onody R, Oosterhuis W, Bunting PS et al. Conflict between guideline methodologic quality and recommendation validity: a potential problem for practitioners. Clin Chem 2006;52:65-72.

4 Graham ID, Calder LA, Hebert PC et al. A comparison of clinical practice guideline appraisal instruments. Int J Technol Assess Health Care 2000;16:1024–38.

5 Cox NH, Williams HC. The British Association of Dermatologists therapeutic guidelines: can we AGREE? British Journal of Dermatology 2003;148:621–625.

6 AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care. 2003;12(1):18-23.

7 Werner RN, Marinović B, Rosumeck S et al. The quality of European dermatological guidelines: critical appraisal of the quality of EDF guidelines using the AGREE II instrument. J Eur Acad Dermatol Venereol 2016;30(3):395-403.

8 AGREE Next Steps Consortium (2009). The AGREE II Instrument [Electronic version]. Retrieved 04.04.2017 from http://www.agreetrust.org.

9 Rosumeck S, Sporbeck B, Rzany B, Nast A.Disclosure of potential conflicts of interest in dermatological guidelines in Germany - an analysis - status quo and quo vadis. 2011;9(4):297-304.

10 Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice Guidelines; Graham R, Mancher M, Miller Wolman D, et al., editors. Clinical Practice Guidelines We Can Trust. Washington (DC): National Academies Press (USA); 2011.

11 Van Everdingen JJE, Dreesens DHH, Burgers JS, Swinkels JA, van Barneveld TA, van der Weijden (red). Evidence-Based Richtlijnontwikkeling. Een leidraad voor de praktijk. Houten: Bohn Stafleu van Loghum, 2014.

12 Oxman AD, Schunemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 2. Priority setting. Health Res Policy Syst. 2006;4:14.

13 Oortwijn WJ, Vondeling H, van Barneveld T, et al. Priority setting for health technology assessment in The Netherlands: principles and practice. Health policy. 2002;62(3):227-42.

14 Reveiz L, Tellez DR, Castillo JS, et al. Prioritization strategies in clinical practice guidelines development: a pilot study. Health Res Policy Syst. 2010;8:7.

15 Partridge N, Scadding J. The James Lind Alliance: patients and clinicians should jointly identify their priorities for clinical trials. Lancet 2004;364(9449):1923-1924.

16 Tong A, Lopez-Vargas P, Howell M, Phoon R, Johnson D, Campbell D et al. Consumer involvement in topic and outcome selection in the development of clinical practice guidelines. Health Expect. 2012;15(4):410-23.

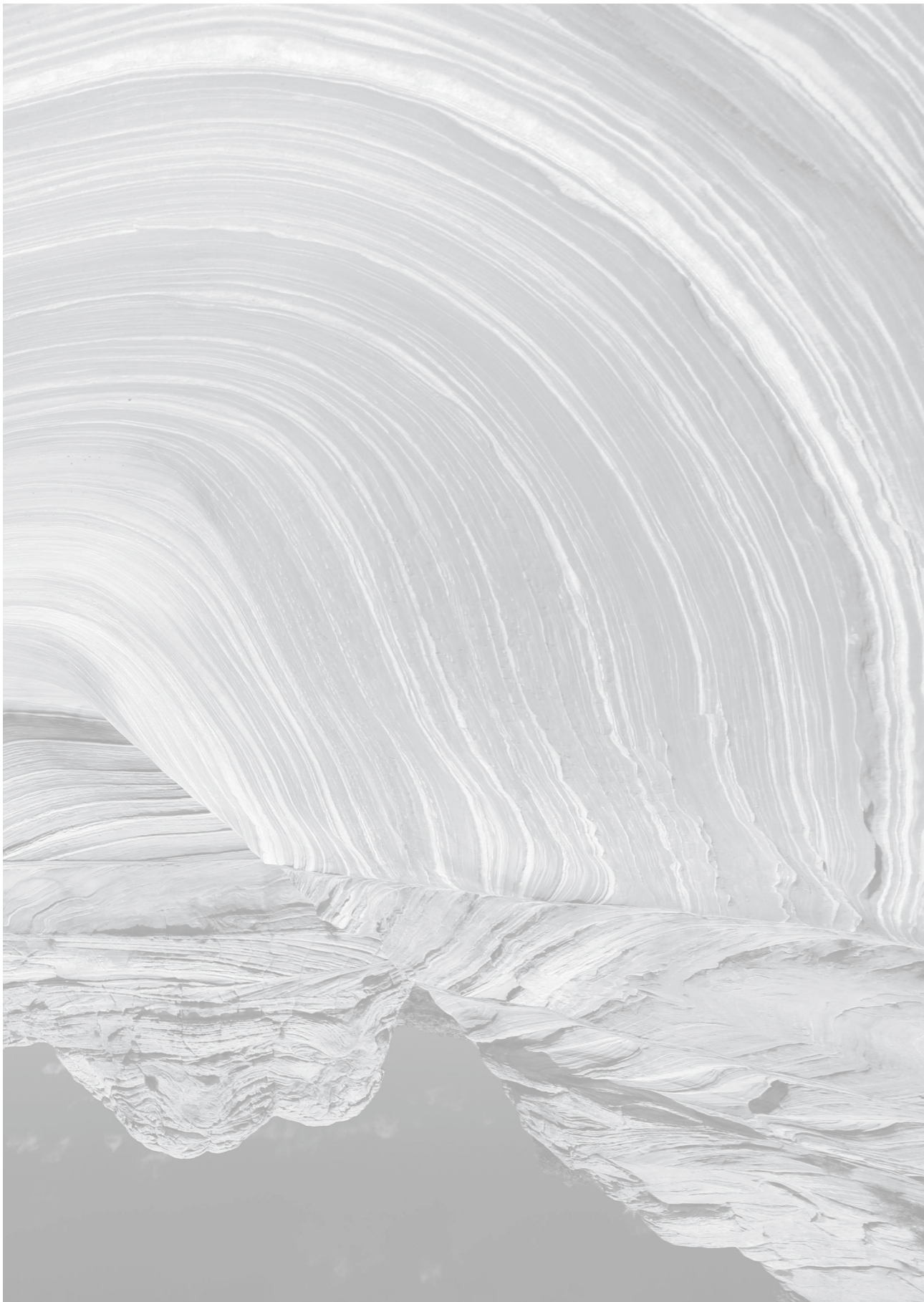
17 van Cranenburgh OD, de Korte J, Sprangers MA, de Rie MA, Smets EM. Satisfaction with treatment among patients with psoriasis: a web-based survey study. Br J Dermatol. 2013;169(2):398-405.

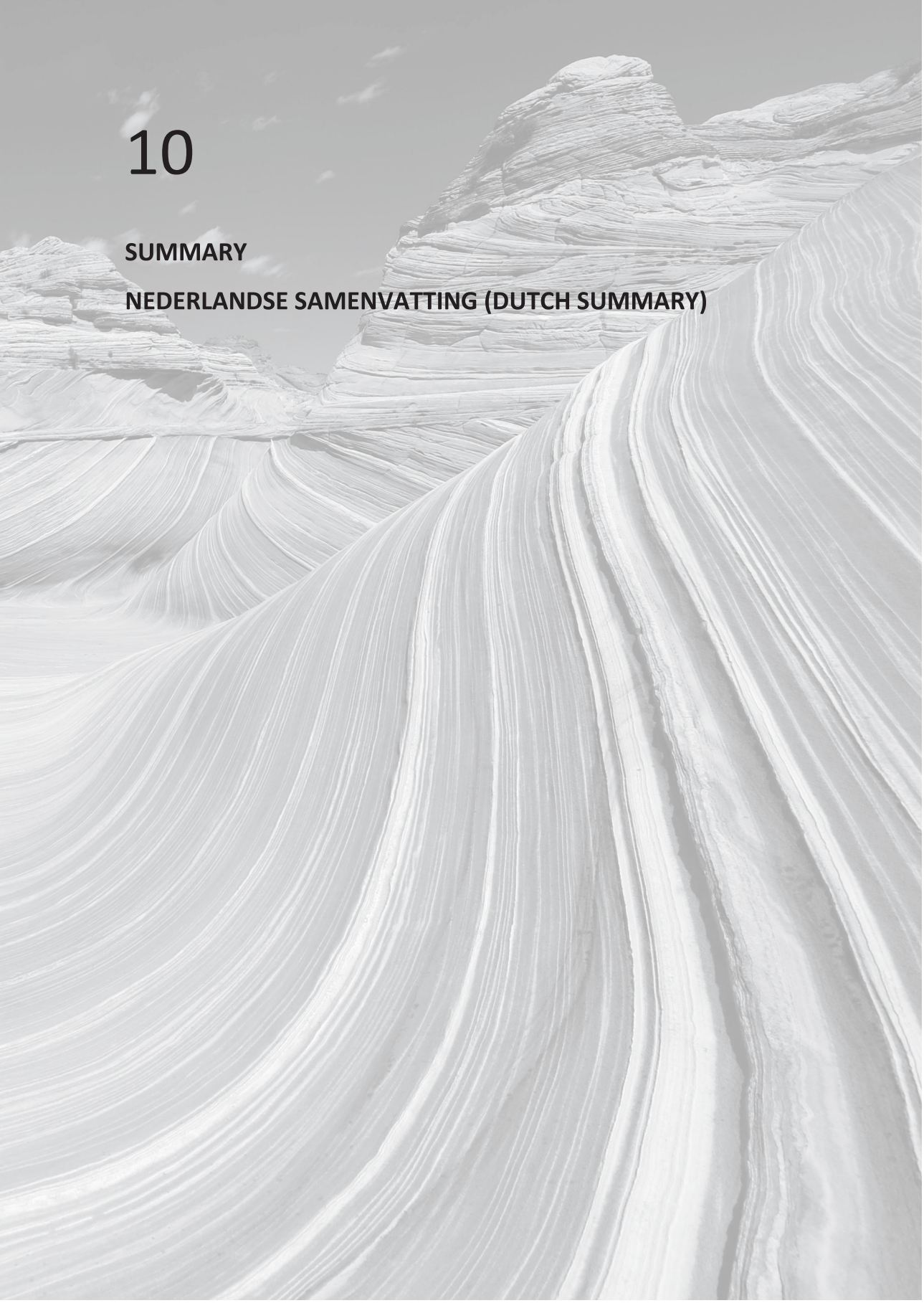
18 Handoll HH, Stott DJ, Elstub LJ, Elliott JC, Kavanagh AL, Madhok R. A framework for effective collaboration between specialist and broad-spectrum groups for delivering priority Cochrane reviews. J Clin Epidemiol. 2013;66(5):490-5.

19 Heinl D, Chalmers J, Nankervis H, Apfelbacher CJ. Eczema Trials: Quality of Life Instruments Used and Their Relation to Patient-reported Outcomes. A Systematic Review. Acta Derm Venereol. 2016;96(5):596-601.
20 Chalmers JR, Simpson E, Apfelbacher CJ, Thomas KS, von Kobyletzki L, Schmitt J et al. Report from the fourth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). Br J Dermatol. 2016;175(1):69-79.
21 Eleftheriadou V, Thomas K, van Geel N, Hamzavi I, Lim H, Suzuki T et al.; Vitiligo Global Issues Consensus Group. Developing core outcome set for vitiligo clinical trials: international e-Delphi consensus. Pigment Cell Melanoma Res. 2015;28(3):363-9.
22 Thorlacius L, Ingram JR, Garg A, Villumsen B, Esmann S, Kirby JS et al. Protocol for the development of a core domain set for hidradenitis suppurativa trial outcomes. BMJ Open. 2017;20;7(2):e014733.
23 Iyengar S, Williamson PR, Schmitt J, Johannsen L, Maher IA, Sobanko JF et al. Development of a core outcome set for clinical trials in rosacea: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. Trials. 2016;17(1):429.
24 Schmitt J, Deckert S, Alam M, Apfelbacher C, Barbaric J, Bauer A et al. Report from the kick-off meeting of the Cochrane Skin Group Core Outcome Set Initiative (CSG-COUSIN). Br J Dermatol. 2016;174(2):287-95.
25 Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336:924–926.
26 Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M et al.; GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ. 2016;353:i2016.
27 Burgers JS, Fervers B, Haugh M, Brouwers M, Browman G, Philip T et al. International Assessment of the Quality of Clinical Practice Guidelines in Oncology Using the Appraisal of Guidelines and Research and Evaluation Instrument. Med Care 2001;39(suppl 2):2-45.
28 Qaseem A, Forland F, Macbeth F, Ollenschlager G, Phillips S, van der Wees P. Guidelines International Network: toward international standards for clinical practice guidelines. Ann Intern Med 2012;156:525-531.
29 Ollenschlager G, Marshall C, Qureshi S, Rosenbrand K, Slutsky J, Burgers J, et al. Improving the quality of health care: using international collaboration to inform guideline programmes by founding the Guidelines International Network.
30 Reddi, A, Prescott, L, Doney, E, Delamere, F, Kollipara, R, Dellavalle, RP et al. The Cochrane Skin Group: a vanguard for developing and promoting evidence-based dermatology. Journal of Evidence-Based Medicine 2013;6:236-242.
31 Peter Fritsch, Walter Burgdorf (eds.) EDF White Book: Skin diseases in Europe. ABW Wissenschaftsverlag. Berlin, 2005 (2nd edition).
32 Baker R, Feder G. Clinical guidelines: where next? Int J Qual Health Care 1997;9:399-404.
33 Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M et al.; GRADE Working Group. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. BMJ. 2016;353:i2089.
34 Eisenberg JM. Globalize the evidence, localize the decision: evidence-based medicine and international diversity. Health Aff (Millwood) 2002;21:166–8.

35 Fervers B, Burgers JS, Voellinger R, et al. Guideline adaptation: an approach to enhance efficiency in guideline development and improve utilisation. BMJ Qual Saf 2011;20:228e36
36 Kristiansen A, Brandt L, Agoritsas T, Akl EA, Berge E, Bondi J et al. Adaptation of trustworthy guidelines developed using the GRADE methodology: a novel five-step process. Chest. 2014;146(3):727-34.
37 Simera I, Kirtley S, Altman DG. Reporting clinical research: Guidance to encourage accurate and transparent research reporting. Maturitas. 2012;72(1):84-87.
38 Chen Y, Yang K, Marušić A, Qaseem A, Meerpohl JJ, Flottorp S et al.; for the RIGHT (Reporting Items for Practice Guidelines in Healthcare) Working Group. A Reporting Tool for Practice Guidelines in Health Care: The RIGHT Statement. Ann Intern Med. 2017;166(2):128-132.
39 Dwan K, Gamble C, Williamson PR, Kirkham JJ; Reporting Bias Group. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One. 2013;8(7):e66844.
40 Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. J Clin Epidemiol. 2011;64(12):1277-82.
41 Neumann I, Santesso N, Akl EA, Rind DM, Vandvik PO, Alonso-Coello P et al. A guide for health professionals to interpret and use recommendations in guidelines developed with the GRADE approach. J Clin Epidemiol. 2016;72:45-55.
42 Raine R, Sanderson C, Black N. Developing clinical guidelines: a challenge to current methods. BMJ 2005;331:631e3.
43 Shekelle PG, Kravitz RL, Beart J, Marger M, Wang M, Lee M. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. Health Serv Res 2000;34:1429–48.
44 Grol R, Grimshaw J. Evidence-based implementation of evidence-based medicine. Jt Comm J Qual Improv 1999;25:503e10.
45 Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999;282:1458e65.
46 Lugtenberg M, Burgers JS, Westert GP. Effects of evidence-based clinical practice guidelines on quality of care: a systematic review. Qual Saf Health Care. 2009;18(5):385-92.
47 Grimshaw J, Eccles M, Thomas R, MacLennan G, Ramsay C, Fraser C et al. Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966–1998. Journal of General Internal Medicine 2006;21(suppl. 1):14–20.
48 Werkgroep Richtlijn voor Richtlijnen; in opdracht: Regieraad Kwaliteit van Zorg. Richtlijn voor Richtlijnen. Den Haag, 2010
49 Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, et al. Changing provider behaviour: an overview of systematic reviews of interventions. Med Care 2001;39(suppl 2):2-45.
50 Gagliardi et al. How can we improve guideline use? A conceptual framework of implementability. Implementation Science 2011;6:26.
51 Shahar Y, Miksch S, Johnson P. An intention-based language for representing clinical guidelines. Proc AMIA Annu Fall Symp. 1996:592-6.
52 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. The Lancet 2003;362:1225-1230.
53 Shekelle P, Eccles MP, Grimshaw JM, Woolf SH: When should clinical guidelines be updated? BMJ 2001, 323:155-157.

54 Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D: How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med 2007;147:224-233.
55 Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 2001;286:1461-1467.
56 Shalom E, Shahar Y, Parmet Y, Lunenfeld E. A multiple-scenario assessment of the effect of a continuous-care, guideline-based decision support system on clinicians' compliance to clinical guidelines. Int J Med Inform. 2015;84(4):248-62.
57 James BC. Making it easy to do it right. N Engl J Med 2001;345(13):991-3.
58 Rothman B, Leonard JC, Vigoda MM. Future of electronic health records: implications for decision support. Mt Sinai J Med. 2012;79(6):757-68.
59 Hsiao JL, Chen RF. Critical factors influencing physicians' intention to use computerized clinical practice guidelines: an integrative model of activity theory and the technology acceptance model. BMC Med Inform Decis Mak. 2016;16(1):3.
60 Burgers JS, van Everdingen JJE. Beyond the evidence in clinical guidelines. Lancet 2004;364(9432):392-3.
61 Lubeek SFK, Michielsens CAJ, Borgonjen RJ, Bronkhorst EM, van de Kerkhof PCM, Gerritsen MP. Impact of High Age and Comorbidity on Management Decisions and Adherence to Guidelines in Patients with Keratinocyte Skin Cancer. Acta Derm Venereol. 2017 [Epub ahead of print]
62 Garber AM. Evidence-based guidelines as a foundation for performance incentives. Health Aff (Millwood). 2005;24(1):174-9.
63 Claridge JA, Fabian TC. History and development of evidence-based medicine. World J Surg. 2005;29(5):547-53.
64 Nasser M, Welch V, Tugwell P, Ueffing E, Doyle J, Waters E. Ensuring relevance for Cochrane reviews: evaluating processes and methods for prioritizing topics for Cochrane reviews. J Clin Epidemiol. 2013;66(5):474-82
65 Utens CM, Dirksen CD, van der Weijden T, Joore MA. How to integrate research evidence on patient preferences in pharmaceutical coverage decisions and clinical practice guidelines: A qualitative study among Dutch stakeholders. Health Policy. 2016;120(1):120-8.
66 Bates AS, Davis CR, Takwale A, Knepil GJ. Patient-reported outcome measures in nonmelanoma skin cancer of the face: a systematic review. Br J Dermatol. 2013;168(6):1187-94.
67 Lugtenberg M, Weenink JW, van der Weijden T, Westert GP, Kool RB. Implementation of multiple-domain covering computerized decision support systems in primary care: a focus group study on perceived barriers. BMC Med Inform Decis Mak. 2015;15:82.
68 Trietsch J, van Steenkiste B, Grol R, Winkens B, Ulenkate H, Metsemakers J et al. Effect of audit and feedback with peer review on general practitioners' prescribing and test ordering performance: a cluster-randomized controlled trial. BMC Fam Pract. 2017;18(1):53.
69 Reddi, A, Prescott, L, Doney, E, Delamere, F, Kollipara, R, Dellavalle, RP et al. The Cochrane Skin Group: a vanguard for developing and promoting evidence-based dermatology. Journal of Evidence-Based Medicine 2013;6:236-242.
70 Agoritsas T, Heen AF, Brandt L, Alonso-Coello P, Kristiansen A, Akl EA et al. Decision aids that really promote shared decision making: the pace quickens. BMJ. 2015;350:g7624.





10

SUMMARY

NEDERLANDSE SAMENVATTING (DUTCH SUMMARY)

SUMMARY

Society increasingly requests transparency and accountability. In the field of medicine this led to the rise of evidence based medicine and an ever-growing amount of medical literature. In **chapter 1** clinical practice guidelines were introduced as a response to those developments. Guidelines aim to describe appropriate healthcare and to bridge the gap between science and practice. Due to a drive towards high quality and all-embracing guidelines, guideline development gradually became more complex and turned into an industry of its one. In **chapter 1** a number of associated problems and research questions were postulated. These questions addressed various difficulties related to dermatological guideline development at the moment and were further explored in chapter 2 until 8.

Chapter 2 introduced a standardized approach to select and prioritize dermatological topics for guideline development. Prioritization is necessary, because resources to develop a guideline are limited and not every topic needs a clinical practice guideline. A survey among Dutch dermatologists yielded a top 20 of most wanted dermatological guideline topics. Skin cancer, psoriasis and atopic dermatitis are statistically significant top 5 topics and a further 30 topics with overlapping confidence intervals apply for the top 20. Guidelines are available for all 20 topics most wanted by dermatologists. Responding Dutch dermatologists mostly agree with (inter)national guideline programmes and literature about the criteria important for selecting guideline topics. Widespread topics with significant morbidity and large variation in current care are often considered for a guideline. Respondents regarded mortality and healthcare costs as less important criteria in topic selection. The shortlist of most wanted topics can serve as a basis for quantitative prioritization by other stakeholders, further pan-European research and translation into specific clinical questions.

In **Chapter 3** the current methodological quality of Dutch dermatological guidelines was investigated with the AGREE II instrument. The 6 AGREE II domain scores highlighted existing good quality guideline areas as well as domains with room for future methodological improvements. In general Dutch dermatological guidelines are methodological sound and meet the current international quality criteria described in the AGREE II instrument. Dutch dermatological guidelines achieved high scores in the domains scope and purpose, stakeholder involvement, rigour of development and clarity of presentation. The domains applicability and editorial independence received lower scores and awareness and improvements in those domains remain necessary.

How to translate the evidence from scientific literature to a clear recommendation fit for practice? In **chapter 4** GRADE (Grades of Recommendation, Assessment, Development and Evaluation) was used to summarize the available literature on the off-label use of cyclosporin in dermatology. GRADE is a method to appraise literature and explicate considerations in formulating recommendations. Evidence for off-label cyclosporin use exists in 22 different dermatological conditions. Overall, the quality of the included literature was low; downgrading due to sparse data or indirectness was common. Prospective long-term registries and studies should add useful evidence and provide legal support for off-label drug use in dermatology. Harmonising outcomes by defining core outcome sets is necessary.

Chapter 5 focussed on the most efficient way to develop a high quality guideline. Three different methods of guideline development were tested, to see if variation occurred in the recommendations for treating actinic keratosis. Correlations between the recommendations formed by the group that followed a traditional evidence-based approach with multiple sessions, the one that followed a two-day pressure cook method and the third method consisting of a systematic evidence report and recommendations made by a single dermatologist, were varying but significant. The methods all had pros and cons. The standard multiple meetings method offered time to adapt and reconsider texts. The pressure cook method presented a possibility for direct discussion and was more time- and cost-efficient. The single dermatologist method is even quicker and cheaper but is more susceptible to bias and hampered acceptance.

Clinical practice guidelines are intended to improve quality of care. However, implementation of guidelines in daily practice is challenging. In **Chapter 6** an experiment was carried out with incorporation of guideline recommendations in the electronic health record system to increase guideline adherence and achieve a reduction in clinical practice variation and costs. The experiment tested the impact of active and passive guideline based clinical decision support on laboratory monitoring of acne patients receiving isotretinoin. Simple guideline based active support incorporated into daily workflow could diminish overtesting. Still, active support was far more effective than passive support. Passive support remained subject to barriers preventing implementation.

Balancing treatment decisions in non-melanoma skin cancer care can be challenging, especially in frail older adults. Clinical practice guidelines could provide assistance.

Chapter 7 explores to what extent bottlenecks experienced in everyday practice are incorporated in guidelines. Potential important items were selected and ranked. Limited life expectancy, treatment goals other than curation, comorbidities and tumour characteristics are items that should be integrated in non-melanoma skin cancer guidelines. Integration of frailty-related items is currently limited. More integration might stimulate a more holistic, personalized and patient-centered care and might make physicians feel more empowered to deviate from a guideline in the best interest of a patient.

Chapter 8 described the development of a tool for obtaining and giving feedback about guideline adherence. Feedback about adherence is essential to monitor the effect that a guideline has in clinical practice. A 32-item questionnaire was used as audit tool to measure adherence to recommendations in the Dutch basal cell carcinoma guideline. The adherence percentages varied considerably due to many possible adherence and implementation barriers. Nonadherence to recommended actions proved to be related to, whether or not a dermatologist was directly involved. The data provides direct feedback to dermatologists and dermatological centers. But the described tool could also be used by developers in a guideline update, to identify and anticipate barriers to successful implementation.

Nowadays, all the different parts of guideline development eventually form a Deming plan-do-check-act cycle. In **chapter 9** the cyclic character of this thesis is followed to highlight several aspects important in guideline development and connect these aspects to the postulated research questions. Thereby, the results from chapters 2 until 8 are put in a broader perspective. **Chapter 9** concludes with thesis recommendations, limitations, implications and suggestions for further research**.**

10

This thesis aimed to explore the quality of dermatological guidelines and how to further improve that quality. Aspects of quality were reviewed in order to get more insight into the necessary steps that are required to reach a high-quality guideline. Thereby it is important to remember that guidelines and its spinoff products are designed as instruments to improve quality of care.

10

NEDERLANDSE SAMENVATTING (DUTCH SUMMARY)

De maatschappij vraagt in toenemende mate om transparantie en verantwoording. In de geneeskunde heeft dit geleid tot de opkomst van evidence based medicine en een nog immer groeiende hoeveelheid aan medische literatuur. In **hoofdstuk 1** werden richtlijnen geïntroduceerd als antwoord op bovenstaande ontwikkelingen. Richtlijnen hebben als doel verantwoorde zorg te omschrijven en als brug tussen wetenschap en praktijk te fungeren. Door het streven naar hoge kwaliteit en allesomvattende richtlijnen werd richtlijnontwikkeling langzamerhand steeds complexer en verwerd het een eigen onderdeel binnen de zorgsector. In **hoofdstuk 1** werden een aantal daar mee samenhangende problemen en onderzoeksvragen opgeworpen. Deze onderzoeksvragen gaan in op verschillende kwesties die momenteel rondom dermatologische richtlijnontwikkeling spelen en werden verder onderzocht in de hoofdstukken 2 tot en met 8.

Hoofdstuk 2 introduceerde een gestandaardiseerde aanpak om dermatologische onderwerpen voor een richtlijn te selecteren en prioriteren. Prioriteren is noodzakelijk omdat de middelen om een richtlijn te ontwikkelen begrensd zijn en niet elk onderwerp een richtlijn nodig heeft. Een enquête onder Nederlandse dermatologen leverde een top 20 van meest gewilde onderwerpen op. Huidkanker, psoriasis en constitutioneel eczeem zijn statistisch significante top 5 onderwerpen en verder kwamen 30 andere onderwerpen, vanwege overlappende betrouwbaarheidsintervallen, in aanmerking voor de top 20. Voor alle 20 meest gevraagde onderwerpen zijn richtlijnen beschikbaar. De Nederlandse dermatologen die reageerden op de enquête, waren het grotendeels eens met de criteria die (inter)nationale richtlijnprogramma's en de literatuur belangrijk vinden om een richtlijnonderwerp te selecteren. Veelvoorkomende onderwerpen met een hoge mortaliteit en variatie in de huidige zorg worden vaak overwogen voor een richtlijn. Mortaliteit en kosten voor de gezondheidszorg werden door dermatologen die reageerden op de enquête beoordeeld als minder belangrijke criteria voor richtlijnonderwerp selectie. De lijst van meest gewilde onderwerpen kan fungeren als basis voor verdere (kwantitatieve) prioritering door andere belanghebbenden, voor verder pan-Europees onderzoek en voor vertaling naar specifieke klinische vragen.

In **hoofdstuk 3** werd de huidige methodologische kwaliteit van de Nederlandse dermatologische richtlijnen onderzocht met behulp van het AGREE II instrument. De 6 AGREE II domeinscores lieten binnen richtlijnen zowel domeinen zien met goede kwaliteit als domeinen met ruimte voor verdere methodologische verbeteringen. Over

het algemeen zijn de Nederlandse dermatologische richtlijnen methodologisch in orde en voldoen ze aan de huidige internationale kwaliteitscriteria zoals omvat in het AGREE II instrument. De Nederlandse dermatologische richtlijnen behaalden hoge scores op de domeinen 'onderwerp en doel', 'betrokkenheid van belanghebbenden', 'methodologie' en 'helderheid en presentatie'. De domeinen 'toepassing' en 'onafhankelijkheid van de opstellers' haalden minder hoge scores en blijvende aandacht en verbetering in deze domeinen is onontbeerlijk.

Hoe vertaal je bewijs uit de wetenschappelijke literatuur naar een heldere aanbeveling voor de dagelijkse praktijk? In **hoofdstuk 4** werd GRADE (Grades of Recommendation, Assessment, Development and Evaluation) gebruikt om de beschikbare literatuur over het off-label gebruik van cyclosporine binnen de dermatologie samen te vatten. GRADE is een methode om literatuur te beoordelen en de overwegingen in het maken van aanbevelingen expliciet weer te geven. Er is bewijs voor het off-label gebruik van cyclosporine bij 22 verschillende dermatologische aandoeningen. Over het algemeen was de kwaliteit van de geïncludeerde literatuur laag en werd het naar beneden bijstellen van de kwaliteit door indirecte of gebrekkige data schering en inslag. Prospectieve lange termijn registers en studies zouden bruikbaar bewijs toe voegen en zorgdragen voor juridische ondersteuning in het off-label geneesmiddelengebruik binnen de dermatologie. Het harmoniseren van heterogene uitkomsten door de definitie van een set van fundamentele uitkomst(mat)en is noodzakelijk.

Hoofdstuk 5 richt zich op de meest efficiënte manier om een hoge kwaliteit richtlijn te maken. Drie verschillende methoden van richtlijnontwikkeling werden getest om te zien of er variatie optrad in de aanbevelingen voor behandeling van actinische keratose. Er was een groep die de traditionele evidence-based methode met meerdere sessies volgde, een groep met een tweedaagse snelkookpan methode en een derde methode met aanbevelingen door één dermatoloog aan de hand van een systematisch rapport van de literatuur. Het verband tussen de onderlinge aanbevelingen varieerde maar was significant. De methodes hadden allemaal voor- en nadelen. De standaard methode met meerdere bijeenkomsten gaf tijd voor aanpassingen en heroverweging van de teksten. De snelkookpan methode bood de mogelijkheid tot directe discussie, spaarde tijd en was kosteneffectief. De methode met één dermatoloog was nog sneller en goedkoper maar is kwetsbaarder voor bias en verminderde acceptatie. Richtlijnen zijn bedoeld om de kwaliteit van zorg te

verbeteren. De toepassing van richtlijnen in de dagelijkse praktijk vormt echter een uitdaging.

In **hoofdstuk 6** werd een experiment uitgevoerd waarin richtlijnaanbevelingen in het elektronisch patiëntendossier werden opgenomen met als doel de naleving van richtlijnen te vergroten en een reductie in variatie van zorg en kosten te bewerkstelligen. Het experiment testte de invloed van actieve en passieve ondersteuning, gebaseerd op de richtlijn, op de laboratorium monitoring van acne patiënten met isotretinoine. Eenvoudige op de richtlijn gebaseerde actieve ondersteuning opgenomen in de dagelijkse routine kan bovenmatig testen verminderen. Echter, actieve ondersteuning was vele maten effectiever dan passieve ondersteuning. Passieve ondersteuning bleef blootgesteld aan belemmeringen die toepassing verhinderen.

Afgewogen beslissingen omtrent behandeling bij non-melanoma huidkanker kunnen uitdagend zijn, vooral in het geval van kwetsbare oudere volwassenen. Richtlijnen kunnen hierbij helpen. **Hoofdstuk 7** exploreert in welke mate knelpunten uit de dagelijkse praktijk zijn opgenomen in richtlijnen. Items die mogelijk van belang waren werden geselecteerd en geprioriteerd. Beperkte levensverwachting, behandeldoelen anders dan curatie, comorbiditeit en tumorkenmerken zijn items die in non-melanoma richtlijnen opgenomen moeten zijn. De integratie van items gerelateerd aan kwetsbaarheid is momenteel beperkt. Verdere integratie van deze items zou een meer holistische, gepersonaliseerde zorg stimuleren waarin de patiënt centraal staat. Ook zou het medici meer vertrouwen kunnen geven om van de richtlijn af te wijken in het belang van de patiënt.

Hoofdstuk 8 beschrijft de ontwikkeling van een instrument voor het krijgen en geven van terugkoppeling over het volgen van de richtlijn. Terugkoppeling over naleving van richtlijnen is essentieel om het effect van een richtlijn in de dagelijkse praktijk na te gaan. Een vragenlijst bestaande uit 32 items werd als audit gebruikt om het volgen van de aanbevelingen in de Nederlandse richtlijn over basaalcelcarcinoom te meten. Het percentage waarin de aanbevelingen werden gevolgd varieerde sterk door de vele mogelijke belemmeringen in zowel naleving als toepassing. Het niet volgen van de aanbevolen acties bleek gerelateerd aan het al dan niet direct betrokken zijn van een dermatoloog. De data biedt directe terugkoppeling aan dermatologen en dermatologische centra. Ook kan het beschreven instrument gebruikt worden door

richtlijnontwikkelaars bij een herziening van de richtlijn, om belemmeringen voor een succesvolle toepassing te identificeren en anticiperen.

Tegenwoordig vormen al de verschillende onderdelen van richtlijnontwikkeling uiteindelijk een zogenaamde plan-do-check-act cyclus volgens Deming. In **hoofdstuk 9** wordt het cyclische karakter van dit proefschrift gevolgd om diverse belangrijke aspecten in richtlijnontwikkeling te onderstrepen en om deze aspecten te verbinden met de opgeworpen onderzoeksvragen. Daardoor worden de resultaten van de hoofdstukken 2 tot en met 8 in een breder perspectief geplaatst. **Hoofdstuk 9** eindigt met aanbevelingen, beperkingen, implicaties en suggesties voor verder onderzoek; allen voortkomend uit dit proefschrift.

Dit proefschrift heeft als doelstelling het nagaan van de kwaliteit van dermatologische richtlijnen en mogelijke verdere verbeteringen in deze kwaliteit. Aspecten van kwaliteit werden belicht teneinde meer inzicht te verkrijgen in de stappen die noodzakelijk zijn om een hoge kwaliteit richtlijn te bereiken. Daarbij is het belangrijk om te herinneren dat richtlijnen en afgeleide producten bedoeld zijn als instrumenten om de kwaliteit van zorg te verbeteren.

11

LIST OF PUBLICATIONS

CURRICULUM VITAE

ACKNOWLEDGEMENTS

LIST OF PUBLICATIONS

Borgonjen RJ, Lubeek SFK, Bik CM, van Everdingen JJE. Quality of Dutch dermatological guidelines according to AGREE II. J Eur Acad Dermatol Venereol. 2017 Oct 6.

Lubeek SFK, Michielsens CAJ, **Borgonjen RJ**, Bronkhorst EM, van de Kerkhof PCM, Gerritsen MP. Impact of High Age and Comorbidity on Management Decisions and Adherence to Guidelines in Patients with Keratinocyte Skin Cancer. Acta Derm Venereol. 2017;97(7):825-829.

Borgonjen RJ, de Lange JA, van de Kerkhof PCM. Guideline-based clinical decision support in acne patients receiving isotretinoin; improving adherence and cost-effectiveness. J Eur Acad Dermatol Venereol. 2017;31(10):e440-e442.

Lubeek SF, **Borgonjen RJ**, Van Vugt LJ, Olde Rikkert MG, Van de Kerkhof PCM, Gerritsen MJP. Improving the applicability of guidelines on nonmelanoma skin cancer in frail older adults: a multidisciplinary expert consensus and systematic review of current guidelines. Br J Dermatol. 2016;175;5:1003-1010.

Borgonjen RJ, van Everdingen JJE, Bruijnzeel-Koomen CA, van de Kerkhof PC, Spuls PI. A national study on adherence to a basal cell carcinoma guideline; development of a tool to assess guideline adherence. Br J Dermatol. 2015;172;4:1008-13.

Borgonjen RJ, van Everdingen JJE, van de Kerkhof PC, Spuls PI. Prioritizing dermatoses; rationally selecting guideline topics. J Eur Acad Dermatol Venereol. 2015;29;8:1636-40.

Zweegers J, de Jong EM, Nijsten TE, de Bes J, te Booij M, **Borgonjen RJ** et al. Summary of the Dutch S3-guidelines on the treatment of psoriasis 2011. Dutch Society of Dermatology and Venereology. Dermatol Online J. 2014;17;20(3).

Tuut M, van Everdingen J, **Borgonjen R**, Oudejans J. Evidence-based guideline development using pressure-cook sessions: study and experience. GIN Conference Berlin 2012.

Borgonjen RJ, Van Everdingen JJ, Bik CM, Tuut MK, Spuls PI, Van De Kerkhof PC. Prospective comparison of three guideline development methods for treatment of actinic keratosis. BMJ Qual. Saf. 2011;20(10):832-841.

Schram ME, **Borgonjen RJ**, Bik CMJM, Van Der Schroeff JG, Van Everdingen JJE, Spuls PI. Off-label use of azathioprine in dermatology: A systematic review. Arch. Dermatol. 2011;147(4):474-488.

RJ Borgonjen, JJE van Everdingen. Dermatological guideline summaries from the Netherlands 2010. DCHG, medische uitgeverij, Haarlem 2010 (ISBN 978-94-90826-05-5)

RJ Borgonjen, JJE van Everdingen, PhI Spuls, PCM van de Kerkhof. Prospective comparison of three different methods of guideline development for treatment of actinic keratosis. GIN Conference Lissabon 2009. Acta. Med. Port. 2009;(22)5:694.

Borgonjen RJ, Kanis J, Van Der Vleuten CJM. Large port-wine stain on the scalp: Suspicious or not?. Ned. Tijdschr. Dermatol. Venereol. 2014;24;10:684-687.

Otters EFM, **Borgonjen RJ**, Hoppenreijs EP, Blokx WAM, Seyger MMB. Pityriasis rubra pilaris: Think of dermatomyositis. Ned. Tijdschr. Dermatol. Venereol. 2013;23;11:658-660.

van der Wees PJ, Kamphuis M, **Borgonjen RJ**. Hoofdstuk 9: Project management. In: Handboek evidence-based richtlijnontwikkeling: Een leidraad voor de praktijk. Bohn Stafleu van Loghum. 2013; 2e editie. (ISBN: 978-90-36802-66-6)

Samenvatting richtlijn actinische keratose, plaveiselcelcarcinoom, hidradenitis, acne en rosacea. Allen in: Bik CMJM, Van Everdingen JJE. Samenvatting richtlijnen Dermatologie 2012. DCHG medische uitgeverij, Haarlem 2012. (ISBN:978-94-90826-12-3)

Borgonjen RJ, Bik CMJM. Establishment and use of the acneïform dermatoses guideline: Acne, rosacea and suppurative hidradenitis. Ned. Tijdschr. Dermatol. Venereol. 2011;21(2):118-120.

JJE van Everdingen, **RJ Borgonjen**. Samenvatting Richtlijnen Dermatologie 2009; DCHG, medische uitgeverij, Haarlem 2009 (ISBN:978-90-811646-6-5)

CURRICULUM VITAE

Rinke Joachim Borgonjen zag het levenslicht op 24 juni 1983 in Eerbeek en groeide op in Brummen. Hij volgde voorbereidend wetenschappelijk onderwijs aan het Isendoorn College te Warnsveld en behaalde zijn diploma in 2001. Tijdens de studie Geneeskunde aan de toen nog Katholieke Universiteit Nijmegen werd zijn interesse voor het vakgebied dermatologie gewekt tijdens het blok Zintuigen en Huid en bevestigd tijdens de co-schappen. Het arts- en doctoraalexamen behaalde hij in 2008; nominaal met een tussenjaar als bestuurder bij de Medische Faculteits Vereniging Nijmegen. Na een korte tijd als arts-assistent chirurgie op de SpoedEisende Hulp ging hij een nieuwe uitdaging als richtlijnontwikkelaar bij de Nederlandse Vereniging voor Dermatologie en Venereologie (NVDV) aan. Gedurende de twee jaar die volgden bekwaamde Rinke zich in de evidence based richtlijnontwikkeling en volgde klinisch epidemiologisch onderwijs aan de Universiteit Utrecht. Tijdens deze twee jaar startte ook een promotietraject onder leiding van prof. P.C.M. van de Kerkhof en prof. P.I. Spuls, waarbij de directe begeleiding door dr. JJE van Everdingen plaats vond. Het proefschrift getiteld 'The perfect guideline, a utopia?' is de vrucht van dit traject. De opleiding tot dermatoloog (2011-2016) doorliep Rinke aan het Radboudumc te Nijmegen met deelstages in het Jeroen Bosch ziekenhuis te Den Bosch en het Canisius Wilhelmina ziekenhuis te Nijmegen. Reizen gecombineerd met stages in Australië rondden de vorming tot veelzijdig dermatoloog af. Als waarnemend dermatoloog was hij vervolgens werkzaam in het Bernhoven te Uden, het Slingeland ziekenhuis te Doetinchem en het Rijnstate te Arnhem. Rinke is nu als dermatoloog verbonden aan het Ziekenhuis Gelderse Vallei en de Padberg kliniek te Ede.

11

ACKNOWLEDGEMENTS

'Beter een brandblaar dan een doorligplek'
-anonieme graffiti Wedren Nijmegen-

Na ettelijke jaren aan dit proefschrift te hebben gewerkt, rest mij nog iedereen te bedanken die op enigerlei wijze betrokken is geweest bij de totstandkoming ervan. Bij dezen. Toch wil ik een aantal mensen eveneens in het bijzonder bedanken.

Jannes bedankt voor de kennisoverdracht omtrent richtlijnontwikkeling en de vele inspirerende ideeën. Mooi dat ik 30 jaar na het verschijnen van jouw proefschrift over 'Consensusontwikkeling in de geneeskunde' een klein beetje in je voetsporen mag treden. De eerste wijsheid van jou op mijn allereerste werkdag in de Domus Medica blijft nog altijd een geldig adagium: het ontbreken van evidence om iets te doen, is niet gelijk aan evidence om iets niet te doen.

Dank Phyllis voor de scherpte, je hartelijkheid en het altijd op zoek blijven naar de klinische relevantie binnen de evidence based dermatologie. Of zoals epidemioloog John Ioannidis zei: 'geen patiënt is ooit gered dankzij een p-waarde.'

Prof. van de Kerkhof bedankt voor het optimisme en het enthousiasme. Als ik twijfelde of dit proefschrift er ooit zou komen dan waren vijf minuten genoeg om weer het tegenovergestelde te denken. Die energie is een gave.

Pieter van der Valk en Peter Arnold dank voor het vertrouwen als wetenschappelijke stage student en het bieden van een opstapje richting de dermatologie.

Ronald Houwing, Peter Arnold, Vigfus Sigurdsson, Ellen de Haas, Nicole Kelleners en Marijke van Dijk bedankt voor het meewerken aan het meest spraakmakende onderzoek van dit proefschrift. De tweedaagse in een hotel in Schoorl behoort tot de leukste herinneringen aan wetenschappelijk onderzoek doen.

Shirley Kwee en Heleen de Vries. Jullie maakten dat ik me vanaf de eerste minuut welkom voelde als richtlijnontwikkelaar bij de Nederlandse Vereniging voor Dermatologie en Venereologie. Ook al zijn we alle drie andere wegen gegaan; wat hebben we soms gelachen. Totdat de buren van het Capaciteitsorgaan kwamen vragen of het niet iets stiller kon.

Een jaar lang constant zicht te hebben op mijn voorhoofd en dan nog steeds samen willen werken wil wat zeggen. Cathelijne Bik, dank voor je vriendschap. Al raakt deze soms helaas wat ondergesneeuwd door onderzoeks- en werk gerelateerde zaken. Hopelijk volgt jouw proefschrift snel.

Mandy Schram, dank voor je samenwerking bij de totstandkoming van richtlijnen voor off-label geneesmiddelengebruik. Leuk ook dat we elkaar blijven tegenkomen als dermatoloog.

Mariska Tuut bedankt voor je professionele begeleiding, de prettige samenwerking en je epidemiologische kennis. Ik hou me nog steeds aanbevolen voor een nieuw gezamenlijk project.

Alle stafleden en arts-assistenten tijdens mijn opleidingstijd in het Radboudumc, het Jeroen Bosch ziekenhuis en het Canisius Wilhelmina ziekenhuis. Bedankt voor de vorming tot dermatoloog, de interesse in dit proefschrift en het meedenken.
De teams van de afdelingen dermatologie in het Bernhoven, in het Slingeland ziekenhuis en in het Rijnstate; bedankt voor de ondersteuning als dermatoloog en de belangstelling.

Collega's in de Gelderse Vallei en de Padberg Kliniek; we kennen elkaar ten tijde van het schrijven van het dankwoord nog niet heel goed, maar het voelt in ieder geval prettig collegiaal.

Aan mijn C40 kamer- en wintersportgenoten dank voor de nodige ontspanning tijdens en buiten werktijd en het bij tijd en wijle fungeren als klaagmuur.
Satish Lubeek, dank voor je inzet en vooral je heldere inbreng. Als je op iemand kan vertrouwen dat iets naar behoren gebeurd, dan ben jij het.
Jonah de Lange, bedankt voor je EPIC gerelateerde hulp. Als je ooit over een carrière switch nadenkt dan is de ICT-helpdesk wellicht het overwegen waard.

Meine deutschsprachige Familie und Freunde. Vieles in diesem Buch werdet ihr vielleicht nicht verstehen, aber für eine deutsche Zusammenfassung stehe ich jederzeit zur Verfügung. Danke für euer Interesse und eure Unterstützung.
Franziskus, dein Spruch "Es genügt nicht, das Leben zu ertragen, man muss es erkennen und lieben" werde ich hoffentlich immer erfüllen. Schade, dass wir einander nie kennen gelernt haben und über meine Doktorarbeit diskutieren konnten.

Nijmeegse vrienden en vriendinnen. Sinds de studietijd inmiddels groot geworden, gezinnen gesticht en uitgewaaierd over de globe, ben ik nog steeds dankbaar voor de tijd die we vinden om nieuwe herinneringen te maken. En voor de interesse in het proefschrift natuurlijk.

Bevers en Beverinnen; het is altijd fijn om ongecompliceerd samen te zijn. Hoewel we inmiddels bijna allemaal artsen of een soort van arts zijn gaat het gelukkig bijna nooit over geneeskunde. Dat er nog maar veel onconventionele plannen uitgevoerd mogen worden.

Nooit gedacht dat hardlopen gezellig kon zijn. Hoewel de groep is uitgedund tot de harde kern, is het wekelijkse rondje meer dan alleen sporten. Dank voor jullie luisterend oor.

Koen en Quirina, Otto en Imke en Guus, Koen en Ilma. Vriendschap vanaf dat we enigszins abstract konden redeneren. Bijzonder dat dit bestaat. De afstand maakt dat de stap soms wat groter is om elkaar te zien, maar deze afstand is in een mum van tijd verdwenen als we elkaar spreken.

Jasper en Maartje; bedankt dat jullie mij willen ondersteunen als paranimf. Het is mooi om te zien Jasper, hoe we ons ontwikkeld hebben maar de vriendschap is gebleven. Maartje, we zijn verschillend van karakter maar onlosmakelijk als broer en zus met elkaar verbonden.

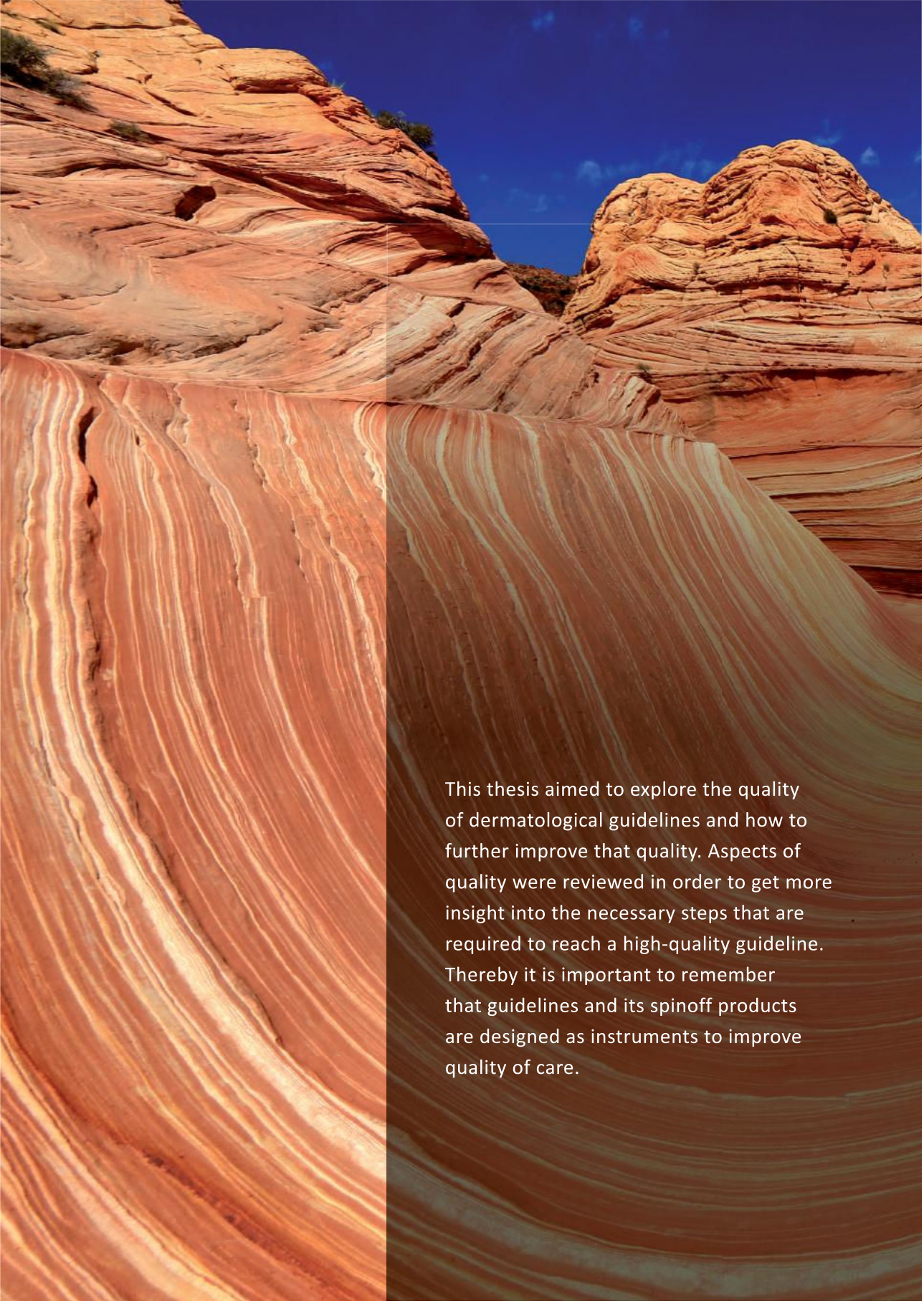
Tonnie en Liesbeth bedankt voor de basis en het laten zien wat liefde is.

Aileen het laatste dankwoord is natuurlijk gericht aan jou inclusief onze ongeboren zoon. Dank voor je liefde. Je laat mij immer mijzelf zijn maar ondersteunt mij tegelijkertijd;

11

Ist meine Erde eine Scheibe, machst du sie wieder rund.
Zeigst mir auf leise Art und Weise was Weitsicht heißt.
Will ich mal wieder mit dem Kopf, durch die Wand.
Legst du mir Helm und Hammer in die Hand.

Dixi. Rinke



This thesis aimed to explore the quality of dermatological guidelines and how to further improve that quality. Aspects of quality were reviewed in order to get more insight into the necessary steps that are required to reach a high-quality guideline. Thereby it is important to remember that guidelines and its spinoff products are designed as instruments to improve quality of care.