Prospective comparison of three guideline development methods for treatment of actinic keratosis
Borgonjen, R.J.; van Everdingen, J.J.; Bik, C.M.; Tuut, M.K.; Spuls, P.I.; van de Kerkhof, P.C.

Published in:
BMJ quality & safety

DOI:
10.1136/bmjqs.2010.050443

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Prospective comparison of three guideline development methods for treatment of actinic keratosis

Rinke J Borgonjen,1 Jannes J van Everdingen,1 Cathelijne M Bik,1 Mariska K Tuut,2 Phyllis I Spuls,3 Peter C van de Kerkhof4

ABSTRACT

Objective: 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 vs 2: p = 0.003; method 1 vs 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 euro and 4 months) and method 1 was the most expensive and slowest method (€48 100 euro and 14 months).

Conclusions: 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 premalignant skin disease in people with skin types I–III and presents itself as solitary or multiple, variable erythematous and irregularly keratinised lesions in sun-exposed areas of the skin.1 2 Consequently, chronic exposure to ultraviolet radiation is an important aetiological 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 few 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.

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 (eg, 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. 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 2). 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. Additional literature brought in by working group members was allowed, but it was obligatory to document study details.

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 (box).

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]

Table 1 Characteristics of the three different methods

<table>
<thead>
<tr>
<th>Categories</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical participants</td>
<td>6 dermatologists, 1 pathologist</td>
<td>6 dermatologists, 1 pathologist</td>
<td>1 dermatologist</td>
</tr>
<tr>
<td>Ratio academic/non-academic dermatologists</td>
<td>3:2</td>
<td>3:2</td>
<td>NR</td>
</tr>
<tr>
<td>No. of participants with PhD in dermato-oncology</td>
<td>2</td>
<td>2</td>
<td>NR</td>
</tr>
<tr>
<td>Age (average)</td>
<td>35–49 years (43.4)</td>
<td>36–51 years (45)</td>
<td>NR</td>
</tr>
<tr>
<td>Ratio male/female</td>
<td>3:3</td>
<td>3:3</td>
<td>0</td>
</tr>
<tr>
<td>No. of participants with former experience in guideline development</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Conflicting interest reported</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

*An epidemiologist prepared the evidence report.
NR, not relevant.

[^71]:
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surgical techniques</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curettage (including supplementary electrodessication)</td>
<td>There is a place for treatment with electrodessication provided that there is enough knowledge and skill available (LOE 4)</td>
<td>Reasonable alternative for hyperkeratotic lesions (LOE 4)</td>
<td>Can be used for (multiple) solitary hyperkeratotic keratoses. At larger body surfaces non-surgical treatments are preferred (LOE 4)</td>
</tr>
<tr>
<td></td>
<td>Add. lit.: —</td>
<td>Add. lit.: —</td>
<td>Add. lit.: —</td>
</tr>
<tr>
<td></td>
<td>Grade: low positive</td>
<td>Grade: low positive</td>
<td>Grade: low positive</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Can be used (LOE 4)</td>
</tr>
<tr>
<td>Dermabrasion</td>
<td>There is a place for treatment with dermabrasion provided that there is enough knowledge and skill available (LOE 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Add. lit.: 1 C study</td>
<td>Add. lit.: —</td>
<td>Add. lit.: —</td>
</tr>
<tr>
<td></td>
<td>Grade: low positive</td>
<td>Grade: low positive</td>
<td>Grade: low positive</td>
</tr>
<tr>
<td>(Shave) excision</td>
<td>No place in the treatment of actinic keratosis, unless there is doubt about the diagnosis and histology is desirable (LOE 3)</td>
<td>Not recommended; unless histology is desirable (LOE 4)</td>
<td>Can be used for (multiple) solitary hyperkeratotic keratoses. At larger body surfaces non-surgical treatments are preferred (LOE 4)</td>
</tr>
<tr>
<td></td>
<td>Add. lit.: 1 C study</td>
<td>Add. lit.: —</td>
<td>Add. lit.: —</td>
</tr>
<tr>
<td></td>
<td>Grade: high/low negative†</td>
<td>Grade: moderate/low negative†</td>
<td>Grade: low positive</td>
</tr>
<tr>
<td><strong>Physical techniques</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cryotherapy</td>
<td>Cryotherapy is a suitable treatment for a few solitary actinic keratoses (LOE 2)</td>
<td>Cryotherapy with liquid nitrogen, preferably in 2 freeze-thaw sessions, is a good choice (LOE 2)</td>
<td>For multiple, solitary and superficial keratoses, cryotherapy can be chosen because of cost-effectiveness and cosmetic result (LOE 3)</td>
</tr>
<tr>
<td></td>
<td>Add. lit.: 1 B study</td>
<td>Add. lit.: —</td>
<td>Add. lit.: —</td>
</tr>
<tr>
<td></td>
<td>Grade: moderate positive</td>
<td>Grade: moderate positive</td>
<td>Grade: low positive</td>
</tr>
<tr>
<td>Photodynamic therapy (PDT)</td>
<td>Photodynamic therapy (PDT) can be used as a treatment for actinic keratosis in case of extensive skin defects over a large area. The added value compared with simpler and cheaper treatments for solitary abnormalities has not been demonstrated (LOE 2)</td>
<td>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)</td>
<td>Because of cost-effectiveness and practicability treatment of surfaces until 100 cm² is recommended (LOE 2/3)</td>
</tr>
<tr>
<td></td>
<td>Add. lit: 1 B study</td>
<td>Add. lit.: —</td>
<td>Add. lit.: —</td>
</tr>
<tr>
<td></td>
<td>Grade: low positive</td>
<td>Grade: moderate positive</td>
<td>Grade: moderate positive</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laser</td>
<td>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)</td>
<td>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)</td>
<td>There is yet insufficient place for laser in treating actinic keratosis because of side effects and insufficient evidence for effectiveness (LOE 3)</td>
</tr>
<tr>
<td>Chemical techniques</td>
<td>Jessner/tri-chloric acid/glycolic acid peel[34–36]</td>
<td>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-fluorouracil (5-FU) treatment, especially in patients who desire a single treatment and patients whose compliance is estimated to be low (LOE 3)</td>
<td>There are no benefits for chemical peeling compared with 5-FU. Therefore, chemical peeling is not recommended (LOE 3)</td>
</tr>
<tr>
<td>Drugs</td>
<td>5-FU[12 17 22 34–42]</td>
<td>5-FU is an effective therapeutic, particularly for the treatment of a greater number of actinic keratoses (LOE 2)</td>
<td>For large body surfaces treatment with 5-FU one time a day for 4 weeks is a useful choice of therapy (LOE 1/3)</td>
</tr>
<tr>
<td>Imiquimod</td>
<td>5-FU is an effective therapeutic, particularly for the treatment of a greater number of actinic keratoses (LOE 2)</td>
<td>Grade: moderate positive Treatment with imiquimod has additional value, only when other treatments have failed (LOE 1/3)</td>
<td>Grade: moderate positive 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)</td>
</tr>
<tr>
<td>Diclofenac</td>
<td>Diclofenac 3% in hyaluronacid 2.5% is not a first choice therapeutic.</td>
<td>Diclofenac is little effective and not recommended (LOE 2)</td>
<td>Grade: moderate positive Diclofenac gel is not recommended (LOE 1)</td>
</tr>
</tbody>
</table>

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

Grade: moderate positive Grade: moderate positive Grade: moderate positive

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderate negative

Grade: moderat...
Discrepancies were discussed until an agreement was reached. Factors attributing to possible differences were described. Secondary reported outcomes were total costs (e.g., retrieval of articles, working group expenses, salary, catering and so on) and time period necessary to make a draft guideline (including preparatory work like searching databases and retrieving full text articles).

### Table 2

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retinoids</td>
<td>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)</td>
<td>There is no place for local retinoids (LOE 2/3)</td>
<td>Primary treatment with retinoids is not recommended (LOE 3)</td>
</tr>
<tr>
<td>Colchicine</td>
<td>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)</td>
<td>Further research to compare local colchicine with the current standard therapies is desirable (LOE 3)</td>
<td>There is yet no place for colchicine cream (LOE 2/3)</td>
</tr>
<tr>
<td>Masoprocol</td>
<td>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)</td>
<td>Masoprocol has no place in the treatment of actinic keratosis, because it is not available in The Netherlands and there is only one study with a limited follow-up period (LOE 3)</td>
<td>Masoprocol has no place in the treatment of actinic keratosis, because it is not available in The Netherlands and there is only one study with a limited follow-up period (LOE 3)</td>
</tr>
<tr>
<td>Others</td>
<td>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)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

### 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
The members of method 1 used five additional references, method 2 used three and method 3 used one additional reference (table 2). 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 2). 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 4). 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 4 months) and method 1 was the most expensive and slowest (€48 100 and 14 months) (table 5).

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 4). 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

---

Box Classification of literature according to the level of evidence (levels according to those set by the Dutch Institute for Healthcare Improvement, CBO)

<table>
<thead>
<tr>
<th>Level of evidence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>One systematic review (A1) or at least two studies of level A2 which were carried out independently</td>
</tr>
<tr>
<td>2</td>
<td>At least two independent studies of level B</td>
</tr>
<tr>
<td>3</td>
<td>One study of level A2 or B or studies of level C</td>
</tr>
<tr>
<td>4</td>
<td>Expert opinion, for example, the members of the working group</td>
</tr>
</tbody>
</table>

For articles concerning intervention (prevention or therapy)

A1  Systematic reviews involving at least two studies at A2 level, of which the results of separate studies are consistent
A2  Randomised comparative clinical studies of good quality (randomised, double-blind controlled trials) 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 3 Grading of recommendation according to perceived level of obligation, including the Dutch CBO formulations used to make a weak, moderate or strong recommendation

<table>
<thead>
<tr>
<th>Low level of obligation</th>
<th>Moderate level of obligation</th>
<th>High level of obligation</th>
</tr>
</thead>
<tbody>
<tr>
<td>May</td>
<td>Should</td>
<td>Must</td>
</tr>
<tr>
<td>Including the positive formulations:</td>
<td>Including the positive formulations:</td>
<td>Including the positive formulations:</td>
</tr>
<tr>
<td>▶ can be considered</td>
<td>▶ is recommended</td>
<td>▶ is strongly recommended</td>
</tr>
<tr>
<td>▶ is an option</td>
<td>▶ is advised</td>
<td>▶ is first choice</td>
</tr>
<tr>
<td>▶ can be a place for</td>
<td>▶ is a preferred treatment</td>
<td>▶ is indicated</td>
</tr>
<tr>
<td>▶ can be useful</td>
<td>▶ strive for</td>
<td>▶ is standard therapy</td>
</tr>
<tr>
<td>May not</td>
<td>Should not</td>
<td>Must not</td>
</tr>
<tr>
<td>Including the negative formulations:</td>
<td>Including the negative formulations:</td>
<td>Including the negative formulations:</td>
</tr>
<tr>
<td>▶ maybe no place for</td>
<td>▶ is not recommended</td>
<td>▶ is strongly discouraged</td>
</tr>
<tr>
<td>▶ does not seem useful</td>
<td>▶ is discouraged</td>
<td>▶ is no choice</td>
</tr>
<tr>
<td>▶ is restraint necessary</td>
<td>▶ deserves no</td>
<td>▶ is contraindicated</td>
</tr>
<tr>
<td></td>
<td>recommendation</td>
<td>▶ is no place for</td>
</tr>
</tbody>
</table>
However, more sessions, time and costs were required (table 5) 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. 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. 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 (online supplementary table 1).

### 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.

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.

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. We graded terms in a recommendation according to the perceived level of obligation to follow that so-called

### Table 4 Correlation between the levels of evidence and recommendation grades of the three different guideline methods

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Intraclass correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of evidence</td>
<td></td>
</tr>
<tr>
<td>Method 1 vs method 2</td>
<td>.735 (95% CI 0.300 to 0.916) p=0.003†</td>
</tr>
<tr>
<td>Method 1 vs method 3</td>
<td>.826 (95% CI 0.534 to 0.943) p&lt;0.001†</td>
</tr>
<tr>
<td>Method 2 vs method 3</td>
<td>.859 (95% CI 0.585 to 0.957) p&lt;0.001†</td>
</tr>
<tr>
<td>Recommendations</td>
<td></td>
</tr>
<tr>
<td>Method 1 vs method 2</td>
<td>.755 (95% CI 0.372 to 0.922) p=0.001†</td>
</tr>
<tr>
<td>Method 1 vs method 3</td>
<td>.493 (95% CI 0.001 to 0.806) p=0.026*</td>
</tr>
<tr>
<td>Method 2 vs method 3</td>
<td>.673 (95% CI 0.199 to 0.893) p=0.007†</td>
</tr>
</tbody>
</table>

*Intraclass correlation is significant between the 0.05 and 0.01 level (two-way).
†Intraclass correlation is significant at least at the 0.01 level (two-way).

### Table 5 Costs and time necessary for the draft guidelines

<table>
<thead>
<tr>
<th></th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total costs*</td>
<td>€48100</td>
<td>€39630</td>
<td>€24770</td>
</tr>
<tr>
<td>Time period necessary for draft guideline†</td>
<td>14 months</td>
<td>5 months</td>
<td>4 months</td>
</tr>
</tbody>
</table>

*The costs for the preparatory work needed for all the guideline methods were added to each method.
†Includes the time for preparatory work; excludes the waiting time for the external approval to start making the guideline.
deontic term.\textsuperscript{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.\textsuperscript{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,\textsuperscript{81} the formation of (inter)national platforms (such as the Guideline International Network)\textsuperscript{82} and in promoting transparent guideline development and implementation methods (Grading of Recommendations Assessment, Development and Evaluation (GRADE), Guideline Implementability Appraisal (GLIA)).\textsuperscript{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.\textsuperscript{86} Matching deontic terms to grades of recommendation strength can help to further standardise the use of such expressions by guideline developers.\textsuperscript{70}

**CONCLUSION**

The therapeutic recommendations and appointed levels of evidence for the treatment of actinic keratoses differed on several occasions 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 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.

**Acknowledgements** We thank all the participating dermatologists and pathologists for making this study possible.

**Funding** Independent funding by Dutch Society of Dermatology and Venerology. The sponsors had no role in the design and conduct of the study; in the collection, analysis and interpretation of data; or in the preparation, review or approval of the manuscript.

**Competing interests** None.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**REFERENCES**


Jorizzo J, Stewart D, Bucko A, et al. Randomized trial evaluating a new 0.5% fluorouracil formulation demonstrates efficacy after 1–2, 4-week treatment in patients with actinic keratosis. [Cutis] 2002;70:335–9.


Loven K, Stein L, Forst 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 keratoses. [Cinther 2002;70:133–41.

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;70:22–9.


82. http://www.g-i-n.net/ (accessed 24 Nov 2010).


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

Rinke J Borgenjen, Jannes J van Everdingen, Cathelijne M Bik, et al.

BMJ Qual Saf 2011 20: 832-841 originally published online May 26, 2011
doi: 10.1136/bmjqs.2010.050443

Updated information and services can be found at:
http://qualitysafety.bmj.com/content/20/10/832.full.html

These include:

Data Supplement
"Web Only Data"
http://qualitysafety.bmj.com/content/suppl/2011/08/22/bmjqs.2010.050443.DC1.html

References
This article cites 81 articles, 7 of which can be accessed free at:
http://qualitysafety.bmj.com/content/20/10/832.full.html#ref-list-1

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/