DOI: 10.1111/jdv.16854 JEADV

ORIGINAL ARTICLE

Outcome assessment in dermatology clinical trials and cochrane reviews: call for a dermatology-specific outcome taxonomy

T. Lange,[1] J. Kottner,[2] T. Weberschock,[3,4] E. Hahnel,[2] C. Apfelbacher,[5,6] S. Brandstetter,[5] A. Dreher,[4,7] T. Datzmann,[1] E. Burden-Teh,[8] N.K. Rogers,[8] P. Spuls,[9] M.J. Grainge,[10] L. Jacobi,[1] H.C. Williams,[6] J. Schmitt[1,*]

[1]Center for Evidence-based Healthcare, University Hospital and Medical Faculty Carl Gustav Carus Dresden, TU Dresden, Dresden, Germany
[2]Department of Dermatology and Allergy, Charité-Universitätsmedizin Berlin, Berlin, Germany
[3]Department of Dermatology, Venereology and Allergology, University Hospital Frankfurt, Frankfurt/Main, Germany
[4]Working Group Evidence-Based Medicine Frankfurt, Institute for General Practice, Goethe University Frankfurt, Frankfurt/Main, Germany
[5]Medical Sociology, Institute of Epidemiology and Preventive Medicine, University of Regensburg, Regensburg, Germany
[6]Institute of Social Medicine and Health Economics, Otto von Guericke University Magdeburg, Magdeburg, Germany
[7]Department of General Internal Medicine and Psychosomatics, Universität Heidelberg, Heidelberg, Germany
[8]Centre of Evidence Based Dermatology, University of Nottingham, Nottingham, UK
[9]Department of Dermatology, Amsterdam Public Health, Infection and Immunity, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
[10]Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Nottingham, UK
*Correspondence: J. Schmitt. E-mail: Jochen.Schmitt@uniklinikum-dresden.de

Abstract

Background Standardized outcome reporting is crucial for trial evidence synthesis and translation of findings into clinical decision-making. The OMERACT 2.0 Filter and COMET outcome domain taxonomy propose frameworks for consistent reporting of outcomes. There is an absence of a uniform dermatology-specific reporting strategy that uses precise and consistent outcome definitions.

Objectives Our aim was to map efficacy/effectiveness outcomes assessed in dermatological trials to the OMERACT 2.0 Filter as a starting point for developing an outcome taxonomy in dermatology.

Methods We critically appraised 10 Cochrane Skin Reviews randomly selected from all 69 Cochrane Skin Reviews published until 01/2015 and the 220 trials included covering a broad spectrum of dermatological conditions and interventions. Efficacy/effectiveness outcomes were mapped to core areas and domains according to the OMERACT 2.0 Filter. The extracted trial outcomes were used for critical appraisal of outcome reporting in dermatology trials and for the preliminary development of a dermatology-specific outcome taxonomy.

Results The allocation of 1086 extracted efficacy/effectiveness outcomes to the OMERACT 2.0 Filter resulted in a hierarchically structured dermatology-specific outcome classification. In 506 outcomes (47%), the outcome concept to be measured was insufficiently described, hindering meaningful evidence synthesis. Although the core areas assessed in different dermatology trials of the same condition overlap considerably, quantitative evidence synthesis usually failed due to imprecise outcome definitions, non-comparable outcome measurement instruments, metrics and reporting.

Conclusions We present an efficacy/effectiveness outcome classification as a starting point for a dermatology-specific taxonomy to provide trialists and reviewers with the opportunity to better synthesize and compare evidence.

Received: 20 December 2019; Accepted: 23 July 2020

Conflicts of interest

Christian Apfelbacher reports grants and personal fees from Dr Wolff GmbH, personal fees from Sanofi Genzyme, outside the submitted work. Jochen Schmitt reports grants from Novartis, Sanofi, Pfizer, ALK outside the submitted work. Hywel Williams is a former co-ordinating editor of Cochrane Skin. Phyllis Spuls reports consultancies in the past for Sanofi 111017 and AbbVie 041217 (unpaid); independent research grants in the past >5 years ago; contract support: Phyllis Spuls is involved in performing clinical trials with many pharmaceutical industries including Novartis, Abbvie, Lilly,

[Correction added on 18 September 2020, after first online publication: Projekt DEAL funding statement has been added.]

Regeneron, Sanofi, Boehringer, Leopharma, UCB, Pfizer, Amgen, Celgene that manufacture drugs used for the treatment of, e.g. psoriasis and atopic dermatitis for which we get financial compensation paid to the department/hospital. The other authors declare that they have no competing interests.

Funding source

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Introduction

In the rapidly developing field of clinical research, it is necessary for clinicians and health policy decision makers to have access to synthesized evidence, since clinical and health policy decisions should always be based on the best current available evidence.[1] Over the last decade, it has become increasingly clear that lack of comparability and the failure to consistently assess all relevant outcomes in clinical trials constitutes a major obstacle in synthesizing evidence. Thus, limiting the use of trial information for clinical decision-making.[2,3] At a primary study level, it is recognized that trialists frequently use very different terminologies for the same underlying outcome.[4]

One approach to solving these problems is the development and implementation of a *Core Outcome Set* (*COS*). A *COS* is an internationally agreed, minimum set of outcomes that are consistently measured and reported in a standardized way for all trials within a health condition.[5] Ideally, a *COS* defines both the *outcome domains* (i.e. WHAT to measure) and the outcome measurement instruments most suitable to measure these *outcome domains* (i.e. HOW to measure).

The *Cochrane Skin-Core OUtcome Set INitiative* (*CS-COUSIN*) is a multi-professional international initiative that supports *COS* groups working to improve and standardize outcome assessment and reporting in dermatology clinical trials and reviews.[6,7] Currently, 17 *COSs* are in development within *CS-COUSIN* for a broad range of dermatological conditions such as atopic dermatitis,[8,9] acne, nail psoriasis,[10] vitiligo,[11] different types of skin cancer,[12] vascular malformations,[13] congenital melanocytic naevi,[14] hidradenitis suppurativa,[15] and incontinence associated dermatitis[16] (www.cs-cousin.org/). A major problem is that across different dermatological conditions, various expressions and definitions for identical or similar *outcome domains* are used (e.g. clinical signs of atopic eczema,[17] progression-free survival in melanoma[6]). Precise and consistently used outcome definitions as well as a sufficiently detailed hierarchy of sub-domains within the *outcome domains* are still missing for dermatology. This creates a significant barrier not only towards the development of *COSs* in dermatology, but also for indexing, referencing and identifying relevant clinical trials and reviews.

In 2014, the international initiative *Outcome Measures in Rheumatology* (*OMERACT*) introduced a conceptual framework for outcome classification and suggested that all *COS* should include at least one outcome from each of the four *core areas* 'death', 'life impact', 'pathophysiological manifestations' and 'resource use/economical impact'.[18] Building on the *OMERACT 2.0 Filter* and other frameworks, the international *Core Outcome Measures in Effectiveness Trials* (*COMET*) group recently introduced a generic taxonomy for *outcome domain* classification in medicine.[19] One of its 29 *outcome domains* is 'skin and subcutaneous tissue outcomes' within the *core area* 'physiological/clinical', but further *sub-domains are not specified* within this relatively broad spectrum of potential *outcome domains* in dermatology. Therefore, the authors pointed out that the *COMET* taxonomy needs to be extended, with a higher level of detail for different specific areas of medicine (e.g. skin and subcutaneous tissue).

A common dermatology-specific outcome taxonomy provides a resource to standardize outcome definitions, classifies outcomes and underlying concepts of outcomes and suggests a systematic hierarchy of outcome domains, sub-domains and outcome concepts. The combination of a taxonomy and the corresponding structured terminology for outcomes/outcomes domains forms the basis for a kind of reference book. The development and expansion of the taxonomy can lead to an improvement in the consistent use of results with corresponding definitions in the future. Furthermore, such a taxonomy of outcomes is important to allow dermatologists and all other research users to better identify relevant outcomes in trials and reviews and thus make evidence identification and synthesis possible and much more efficient.

To move forward in the development of an urgently needed dermatological taxonomy, we sought to map the broad spectrum of outcomes assessed in dermatological trials, investigating a wide range of conditions, to the *OMERACT 2.0 Filter*. Furthermore, we aimed to quantify the utilization of specific *outcome domains/sub-domains* within the corresponding *core areas* to identify their distribution and highlight dermatology outcome domain priorities.

Methods

We critically appraised 10 Cochrane Skin systematic reviews randomly selected from all 69 Cochrane Skin Reviews, published until January 2015, and the underlying clinical trials included within these reviews. The simple random sample was drawn using the statistical software R. The primary focus of the review was to investigate the (i) quality (sufficient and clear outcome reporting) and (ii) degree of consistency and completeness in outcome reporting in clinical trials. In particular, we collected the level of detail in outcome reporting and whether all components of outcome reporting were provided.

Taking the outcome frameworks by Zarin *et al*.,[20] Chan *et al*.[21] and the OMERACT Filter 2.0[18] into account, an outcome within this work was described by the following eight components (Fig. 1):

1 *core area* e.g.: 'life impact';
2 *outcome domain/sub-domain* e.g.: 'perception of health' > 'single symptom' > 'pain';
3 *outcome concept to measure* e.g.: 'reduction', 'progression', 'response';
4 *measurement instrument* e.g.: score: 'visual analogue scale 1–10';
   - including the setting 'who' measured/assessed the outcome (patient self-/external (relative/representative) reported, clinical or laboratory assessment
5 *time of measurement* e.g.: '2 weeks after treatment';
6 *analysis metric* e.g.: 'change from baseline';
7 *method of aggregation* e.g.: 'mean difference', 'frequency of changes', 'odds ratio';
8 *statistical test procedure* e.g.:
   - statistical test: '*t* test', 'Mann–Whitney test'
   - modelling approach: 'logistic regression', 'cox regression' (non-adjusted, adjusted for, e.g. age and sex)

The review protocol was registered in the international *Prospective Register of Systematic Reviews* (*PROSPERO*; CRD42015025005). Detailed description of data sources and data extraction procedures have been reported elsewhere.[4]

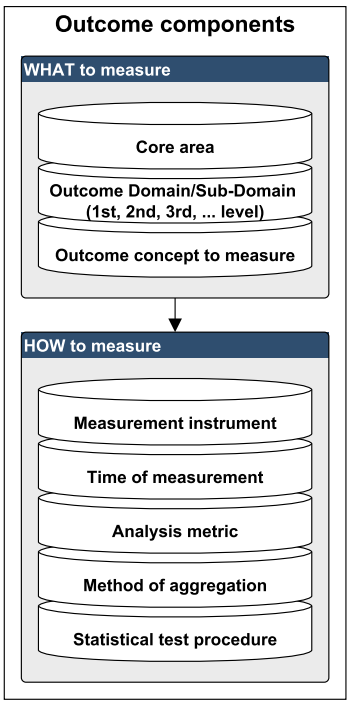

Figure 1 Hierarchical schema of outcome specification based on Zarin *et al*.[20] and Chan *et al*.[21]

Data source and extraction

We used different data sources for this project. First, a random sample of 10 systematic reviews from 69 *Cochrane Skin Reviews*, published until January 2015, were drawn. Predefined inclusion criteria were set to include only trials published in English, German, Italian or Spanish and reported efficacy/effectiveness outcomes other than biomarkers.

Information on trial characteristics, interventions, study populations and the efficacy/effectiveness outcomes relevant for the *Cochrane* reviews was retrieved from the *Archie* database provided by *Cochrane Skin* (https://archie.cochrane.org/). The *Archie* database contains data originally extracted by the *Cochrane Review* teams during their review process. This data source was supplemented by data extraction from the original trial publications for relevant information not covered in *Archie*, such as information on the trial outcomes not considered as relevant from the *Cochrane Reviewers*. These outcomes will be termed as 'additional outcomes' subsequently. Two reviewers independently extracted all 'additional outcomes' and other trial/ outcome information into a standardized electronic template form in MS Access using the *Archie* database and the original trial publications.

Allocation process of outcome domains and sub-domains

The framework of the *OMERACT 2.0 Filter* was used to map all extracted efficacy/effectiveness outcomes to *outcome domains* such as 'acceptance of care', 'daily life impact' or 'perception of health' in an iterative process performed by two reviewers with dermatological and methodological expertise (JS, TL). We first listed all outcomes and mapped them into preliminary outcome domains. Then, we reviewed the resulting lists to determine whether we could combine these outcome domains and re-categorized them if considered appropriate. After completeness and consistency checks, a classification list of *outcome domains* was obtained. In order to take different levels of complexity into account, we used up to two levels of *sub-domains* (1st, 2nd). This allowed the development of a more detailed hierarchical structure on the *outcome domains* as shown in Fig. 2. Figure 2 also provides definitions for the different outcome components. The final list of *outcome domains* was mapped to the *OMERACT core areas* ('death', 'life impact', 'pathophysiological manifestations', 'resource use/economic impact')[18] (Fig. S1). Using the linkage of trial outcomes to *outcome domains/sub-domains*, the frequencies of the domains and *core areas* overall and within each review were calculated.

Results

Study and outcome selection

A total of 220 out of 242 trials (91%) met the inclusion criteria and could be retrieved for analysis (Fig. S2). Overall, these trials covered 1086 efficacy/effectiveness outcomes. 395 of these outcomes (36%) were considered as relevant outcomes for the

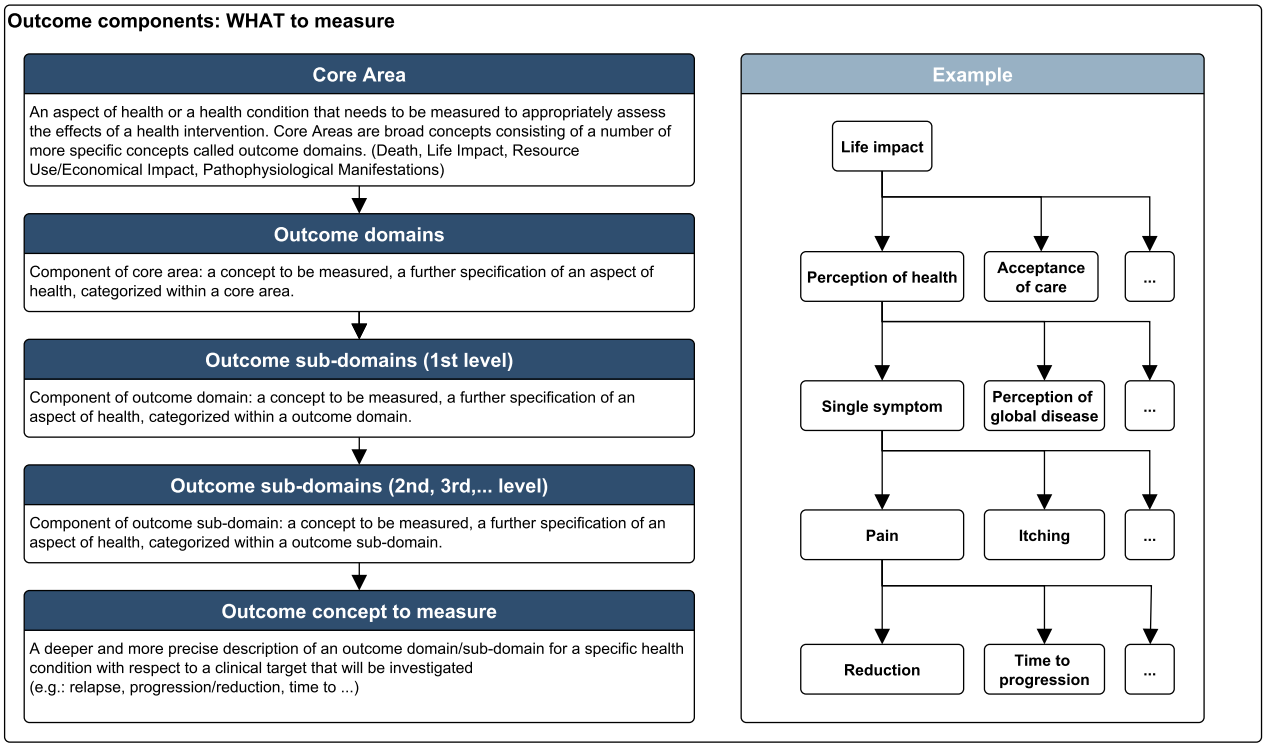

Figure 2 Hierarchical schema of outcome components: WHAT to measure based on Boers *et al*.[18]

corresponding *Cochrane* reviews. The remaining 691 additional outcomes (64%) were not included in the *Cochrane* reviews and were extracted from the trial publications. In total, an average of 4.9 (range 1–34) outcomes per trial were assessed with an average of 2.8 (range 1–14) different measurement instruments.

Allocation of domains to the OMERACT 2.0 Filter

In an iterative process, all 1086 outcomes identified were mapped to *outcome domains and core areas* (Table 1). This procedure led to the grouping of outcomes to *domains* and up to two levels of *sub-domains* (1st level >2nd level). All resulting *outcome domains/sub-domains* could be mapped to one of the four *core areas* as defined by the OMERACT 2.0 filter. The frequencies in Table 1 show the relevance of the different *core areas* in the different sets of dermatology trials. 'Pathophysiological manifestations' such as physical signs of skin disease were considered in all 10 included systematic reviews. Table 2 shows the resulting list of assessed *outcome domains/sub-domains* and their corresponding frequency distributions.

Level of detail in study outcome reporting

The detail of the efficacy/effectiveness outcome description varied considerably between trials. Unfortunately, synthesizable descriptions of outcomes for data extraction (Fig. 1) were not provided in the majority of the included trial publications. Frequently, very broad concepts such as 'response', 'clearance' or 'progression' were used, but with either divergent or unclear/not reported underlying definitions, even within a group of trials for a specific condition. In summary, in 506 out of 1086 cases, we were unable to extract the *outcome domain* (WHAT to measure) in such depth that an evidence synthesis would be feasible. The *outcome concept to measure* therefore needs to be defined more clearly within the outcome domain. Table 3 shows the most frequently investigated *outcome domains* in the most prominent *core areas* 'life impact' and 'pathophysiological manifestations'. It is clear that the mapped outcomes overlapped between trials from a moderate to a high extent at *domain* level. However, for quantitative evidence synthesis, it is also crucial that all other components of the outcome definition are reported, too (Fig. 1). Table 4 illustrates such divergence. Even if trial investigators measured the same outcome (according to an outcome definition that is reduced to the first five components), an incomparability of this outcome across the trials remains. The reason is that this simplification of outcome definition to the described five components is not enough to enable a quantitative evidence synthesis: using this simplified outcome definition about 48% of all potentially synthesizable outcomes were not comparable. If the missing components were included, the extent of

Table 1 Summary OMERACT 2.0 Core areas of underlying trials of included Cochrane reviews (1086 outcomes)

Reference	Condition	*N* trials	*N* outcomes	Core area: death		Core area: life impact		Core area: pathophysiological manifestations		Core area: resource use/economical impact	
				N trials (%)	*N* outcomes (%)	*N* trials (%)	*N* outcomes (%)	*N* trials (%)	*N* outcomes (%)	*N* trials (%)	*N* outcomes (%)
Bamford[27]	Atopic eczema	26	192	0 (0%)	0 (0%)	20 (77%)	74 (39%)	25 (96%)	110 (57%)	7 (27%)	8 (4%)
Bath-Hextall[28]	Atopic eczema	11	99	0 (0%)	0 (0%)	6 (55%)	28 (28%)	11 (100%)	62 (63%)	3 (27%)	9 (9%)
Chen[29]	Psoriasis	13	78	0 (0%)	0 (0%)	2 (15%)	2 (3%)	13 (100%)	76 (97%)	0 (0%)	0 (0%)
Chi[30]	Lichen sclerosus (genital)	7	26	0 (0%)	0 (0%)	6 (86%)	13 (50%)	6 (86%)	13 (50%)	0 (0%)	0 (0%)
Eekhof[31]	Ingrowing toenails	21	93	0 (0%)	0 (0%)	11 (52%)	28 (30%)	21 (100%)	57 (61%)	5 (24%)	8 (9%)
Ersser[32]	Atopic eczema	9	40	0 (0%)	0 (0%)	7 (78%)	23 (57%)	7 (78%)	16 (40%)	1 (11%)	1 (2%)
Kastarinen[33]	Seborrhoeic dermatitis (face and scalp)	35	210	0 (0%)	0 (0%)	32 (91%)	80 (38%)	35 (100%)	130 (62%)	0 (0%)	0 (0%)
Kwok[34]	Common cutaneous warts (non-genital)	70	150	0 (0%)	0 (0%)	2 (3%)	3 (2%)	70 (100%)	145 (97%)	2 (3%)	2 (1%)
Martin[35]	Pemphigus vulgaris and pemphigus foliaceus	10	59	0 (0%)	0 (0%)	1 (10%)	1 (2%)	10 (100%)	44 (75%)	6 (60%)	14 (24%)
Sasse[36]	Metastatic malignant melanoma	18	139	16 (89%)	44 (32%)	3 (17%)	3 (2%)	18 (100%)	92 (66%)	0 (0%)	0 (0%)

N, Number.

incomparability for quantitative evidence synthesis would increase further. The following example illustrates the aforementioned problem of data synthesis. The condition 'common cutaneous warts' was analysed with 150 different outcomes. Within the investigated trials, 143 outcomes were mapped to the *outcome domain* 'global disease severity' (most frequent *domain* here). Of these 143 outcomes, nine were assessed with the same *measurement instrument*, but only five outcomes reported the same *time of measurement*. Thus, of initially 143 'global disease severity' outcomes, only five outcomes have the potential for a quantitative evidence synthesis.

Discussion

Main findings

This review extends previous research on dermatology outcomes in several aspects. This study shows that all identified efficacy/effectiveness outcomes of dermatology trials included in a randomly selected set of dermatological *Cochrane* Reviews could be successfully mapped to the four *core areas* as defined by the *OMERACT 2.0 Filter*. The *OMERACT 2.0 Filter* therefore appears to be a valid starting point for a dermatology-specific taxonomy. In 2018, OMERACT updated this framework (OMERACT Filter 2.1) in order to improve the precision and accuracy. The four core areas have been renamed but have a similar meaning compared to the previous version.[22]

In terms of outcome content, our review highlights the vast predominance of physician-reported 'pathophysiological manifestations' in all trial sets included. The patient-relevant *core area* 'life impact', where most of the patient-reported outcomes like disease symptoms and quality of life are included, were not considered in the majority of trials.

The main contribution of our study is providing a starting point of a dermatology-specific outcome taxonomy mapping *outcome domains/sub-domains* to the four *core areas* 'death', 'life impact', 'pathophysiological manifestations', 'resource use/economic impact'[18] as the highest hierarchical structure. Such an outcome taxonomy is an important resource of COS developers in dermatology, as well as for trialists, guideline developers and everyone else who develops or uses dermatology trials and reviews.

Another important finding of our study is that the common level of outcome description is insufficient in most of the trial publications to be used for quantitative evidence synthesis. In almost half of the outcomes examined, we could not clearly break them down into the underlying outcome components. This lack of detail severely restricts the potential of quantitative evidence synthesis.

The use of a list of components to describe an outcome, which was proposed by Zarin *et al.*,[20] Chan *et al.*[21] and the OMERACT Filter 2.0,[18] is very helpful to define and classify outcomes. However, in our view, we need a more detailed description of the WHAT to measure. In addition to the *outcome domains*, it is also

Table 2 Proposed starting point for a dermatology outcome taxonomy based on the underlying trials of included Cochrane reviews

Core area (OMERACT 2.0)	Domain			Sub-domain (1st level)			Sub-domain (2nd level)		
		Trials†	Outcomes‡		Trials†	Outcomes‡		Trials†	Outcomes‡
Death (7/4)	Overall survival	7%	4%	Overall survival (all cause death)	6%	2%			
	Treatment related mortality	1%	<1%						
Life impact (41/23)	Acceptance of care	4%	1%	Compliance/ adherence	2%	1%			
				Coping	2%	1%	Itching behaviour (scratching)	<1%	<1%
							Skin (care) behaviour	1%	<1%
	Activity participation	1%	<1%	Sports activity participation	<1%	<1%			
	Daily life impact	3%	<1%						
	Family impact	2%	1%	Burden for mother	<1%	<1%			
				Mother-child interaction	1%	<1%			
				Parental QOL	1%	<1%			
	Patient satisfaction	3%	1%	Treatment satisfaction	2%	<1%	Satisfaction with cosmetic outcome	1%	<1%
							Satisfaction with treatment effectiveness	2%	<1%
	Perception of health	34%	19%	Perception of cosmetic outcome	<1%	<1%			
				Perception of general symptoms	3%	1%	Appetite/thirst	1%	<1%
							Sleeping problems	1%	<1%
				Perception of global disease	2%	<1%	Perception of eczema severity	1%	<1%
							Disease related sleeping problems	1%	<1%
				Perception of global health	1%	<1%			
				Single symptoms	39%	20%	Burning	1%	<1%
							Dandruff	<1%	<1%
							Dryness	3%	1%
							Dyspareunia	1%	<1%
							Erythema/ redness	23%	8%
							Flakiness	<1%	<1%
							Itching	7%	2%
							Lichenification	2%	1%
							Pain	4%	1%
							Papules	<1%	<1%
							Pruritus	15%	4%
							Scaling	4%	1%
							Seborrhea	<1%	<1%
							Soreness	<1%	<1%
							Stinging	<1%	<1%
							Unspecified dermatological symptoms	5%	2%
				Visibility appearance	1%	<1%			
	Quality of life	6%	1%	Quality of life, generic	2%	<1%			
				Quality of life, dermatologic	4%	1%			
				Quality of life, oncologic	<1%	<1%			

Table 2 *Continued*

Core area (OMERACT 2.0)	Domain			Sub-domain (1st level)			Sub-domain (2nd level)		
		Trials†	Outcomes‡		Trials†	Outcomes‡		Trials†	Outcomes‡
Pathophysiological manifestations (98/69)	Clinical assessment	98%	67%	Global disease severity	90%	40%	Appearance of nail	10%	5%
							Global severity of eczema	<1%	<1%
							Presence of wart	32%	13%
							Visibility of disease signs	<1%	<1%
				Single signs	26%	17%	Bleeding	<1%	<1%
							Crusting	3%	1%
							Dandruff	<1%	<1%
							Desquamation	1%	<1%
							Dryness	4%	1%
							Erythema/ redness	19%	7%
							Excoriation	3%	1%
							Exudation	1%	<1%
							Fissures	1%	<1%
							Greasiness	1%	<1%
							Infiltration	2%	<1%
							Inflammation	1%	<1%
							Lichenification	3%	1%
							Maceration	<1%	<1%
							Oedema	1%	<1%
							Oozing	<1%	<1%
							Papulation	1%	<1%
							Peeling	<1%	<1%
							Roughness	<1%	<1%
							Scaling	12%	3%
							Sebum secretion	<1%	<1%
							Skin texture	<1%	<1%
							Surface damage	<1%	<1%
							Ulceration	1%	<1%
							Vesiculation	<1%	<1%
							Weeping	<1%	<1%
				Combinations of two or more signs	19%	10%	Erythema, oedema, excoriation, lichenification		
							Erythema, oedema, excoriation, lichenification, crusting, dryness		
							(. . .)		
	Laboratory assessment	4%	1%	Biomarkers	<1%	<1%			
				Colonization	2%	<1%			
				Histo-pathology	2%	1%			
	Physiological skin assessment	1%	<1%						
Resource use/ economical impact (11/4)	Direct costs	<1%	<1%	Financial and temporal burden	<1%	<1%			
	Healthcare	10%	4%	Treatment visits	<1%	<1%			
				Treatment utilization	10%	3%	Indication for concomitant treatment	1%	<1%
							Utilization of specific treatment	9%	3%
	Sick leave	<1%	1%						

†Indicates the percentage of all 220 trials that investigated the core area, domain or sub-domain.
‡Indicates the percentage of all 1,086 outcomes that investigated the core area, domain or sub-domain.

Table 3 Most frequently assessed outcome domains in the core areas life impact and pathophysiological manifestations in dermatological trials described in 10 Cochrane reviews

Cochrane reviews				Individual trials					
Reference	**Condition**	***N* trials**	***N* outcomes**	***N* trials measuring Life impact**	***N* outcomes measuring Life impact**	**Life impact Outcome domain: *N* trials (%); *N* outcomes (%)**	***N* trials measuring Pathophysiological Manifestations**	***N* outcomes measuring Pathophysiological Manifestations**	**Pathophysiological manifestations Outcome domain: *N* trials (%); *N* outcomes (%)**
Bamford[27]	Atopic eczema	26	192	20	74	1 Perception of health: 20 (100%); 69 (93%) 2 Quality of life: 2 (10%); 3 (4%) 3 Acceptance of care: 1 (5%); 1 (1%)	25	110	1 Clinical assessment: 25 (100%); 108 (98%) 2 Physiological skin assessment: 2 (8%); 2 (2%)
Bath-Hextall[28]	Atopic eczema	11	99	6	28	1 Perception of health: 6 (100%); 24 (86%) 2 Acceptance of care: 2 (33%); 2 (7%) 3 Daily life impact: 2 (33%); 2 (7%)	11	62	1 Clinical assessment: 10 (91%); 59 (95%) 2 Laboratory assessment: 1 (9%); 2 (3%) 3 Physiological skin assessment: 1 (9%); 1 (2%)
Chen[29]	Psoriasis	13	78	2	2	1 Perception of health: 1 (50%); 1 (50%) 2 Quality of life: 1 (50%); 1 (50%)	13	76	1 Clinical assessment: 13 (100%); 76 (100%)
Chi[30]	Lichen sclerosus (genital)	7	26	6	13	1 Perception of health: 6 (100%); 13 (100%)	6	13	1 Clinical assessment: 6 (100%); 9 (69%) 2 Laboratory assessment: 4 (67%); 4 (31%)
Eekhof[31]	Ingrowing toenails	21	93	11	28	1 Perception of health: 8 (73%); 15 (54%) 2 Daily life impact: 4 (36%); 7 (25%) 3 Patient satisfaction: 4 (36%); 5 (18%)	21	57	1 Clinical assessment: 21 (100%); 57 (100%)
Ersser[32]	Atopic eczema	9	40	7	23	1 Quality of life: 6 (86%); 8 (35%) 2 Famility impact: 5 (71%); 6 (26%) 3 Acceptance of care: 3 (43%); 7 (30%)	7	16	1 Clinical assessment: 7 (100%); 16 (100%)
Kastarinen[33]	Seborrhoeic dermatitis (face and scalp)	35	210	32	81	1 Perception of health: 31 (97%); 79 (98%) 2 Acceptance of care: 1 (3%); 1 (1%)	35	130	1 Clinical assessment: 35 (100%); 125 (96%) 2 Laboratory assessment: 4 (11%); 4 (3%)

Table 3 *Continued*

Cochrane reviews				Individual trials						
Reference	**Condition**	***N* trials**	***N* outcomes**	***N* trials measuring Life impact**	***N* outcomes measuring Life impact**	**Life impact Outcome domain: *N* trials (%); *N* outcomes (%)**	***N* trials measuring Pathophysiological Manifestations**	***N* outcomes measuring Pathophysiological Manifestations**	**Pathophysiological manifestations Outcome domain: *N* trials (%); *N* outcomes (%)**
Kwok[34]	Common cutaneous warts (non-genital)	70	150	2	3	1 Acceptance of care: 1 (50%); 1 (33%) 2 Perception of health: 1 (50%); 1 (33%) 3 Quality of life: 1 (50%); 1 (33%)	66	145	1 Clinical assessment: 66 (100%); 145 (100%)
Martin[35]	Pemphigus vulgaris and pemphigus foliaceus	10	59	1	1	1 Patient satisfaction: 1 (100%); 1 (100%)	10	44	1 Clinical assessment: 10 (100%); 44 (100%)
Sasse[36]	Metastatic malignant melanoma	18	139	3	3	1 Quality of life: 3 (100%); 3 (100%)	18	92	1 Clinical assessment: 18 (100%); 92 (100%)

N, Number.

crucial to specify the *outcome concept measured* when describing the outcome (Fig. 1). We understand this part of the component (WHAT to measure) as a deeper and more precise description of an *outcome domain* for a specific health condition with respect to a clinical target that will be investigated (e.g.: 'occurrence', 'relapse', 'progression/reduction'). Considering the single symptom itching (within the core area of 'life impact'), different concepts of WHAT to measure are feasible (e.g. the occurrence of itching after 2 weeks intervention, time of progression-free itching, the itching status, itch severity, itch frequency, itch intensity or the fully remission of itching).

In many dermatological studies, only the *outcome domain* or its concept was reported. It makes a big difference, whether – for example – erythema is considered as a *sub-domain* of the *outcome domain* clinical signs to describe the 'intensity of skin lesions', or if the degree of erythema is used as an indicator for 'disease remission' or 'progression' as an outcome on its own.[23] In summary, even if the same *outcome domain* is investigated in different trials, outcome definitions are still too broad to be able to summarize these outcomes across the trials.

Quantitative evidence synthesis will only be fully possible if all of the described components (Fig. 1) were reported and have key similarities across trials. This comparability is the rationale for developing *COSs. COS* developers aim to standardize relevant outcomes for trials to enable an effective evidence synthesis. Despite these efforts, developing a *COS* is an important step forward; however, a *COS* is still a too broad concept, without aligning all components of the outcome definition in a *COS*. Therefore, we recommend the consideration and standardized reporting of all proposed components (including statistical tests to be applied) of an outcome for every efficacy/effectiveness trial to harmonize across-trial outcome synthesis.

Future taxonomy development

Hierarchy is required for different medical fields. Therefore, we used a bottom-up approach to provide a basis of mapped *outcome domains* for further development of a dermatology-specific extension of this taxonomy. This approach was based on the common procedure for identification of potential *outcome domains* in the context of *COS* development.[24] In the future, using the taxonomy other dermatology-specific conditions might require the inclusion of further levels of *sub-domains* in order to minimize the loss of information or to describe newly introduced *domains*.

Limitations

The source of data was restricted to a random sample of 10 *Cochrane Skin Reviews,* and therefore, only a selection of certain conditions could be considered. Due to the time restriction, new developments in standardized reporting in dermatology trials after 2014 could not be considered. However, we believe that the random sample lead to a saturation of themes and covers an appropriate range of dermatological topics and diseases.

Table 4 Application of hierarchical schema to the most frequently chosen outcome constructs in the 10 sets of dermatology trials

Review	Condition	*N* trials	*N* outcomes	Most frequent outcome construct (domain, sub-domain: 1st level 2nd level)	Most frequent instrument†	Most frequent time of measurement‡, *N* outcomes (%)	Most frequent statistic *N* outcomes§ (%)
Bamford[27]	Atopic eczema	26	192	Life Impact; • Perception of health; • Single symptoms: 59 outcomes in 20 trials	Visual analogue scale (0 = no symptoms, 100 = worst ever): 17 outcomes in 6 trials	16 weeks: 16 outcomes in 5 trials	Difference of mean change from baseline to post-treatment: 16 outcomes in 5 trials
Bath-Hextall[28]	Atopic eczema	11	99	Pathophysiological Manifestations; • Clinical assessment; • Single signs: 28 outcomes in 3 trials	SCORAD: 19 outcomes in 1 trial	8 weeks: 19 outcomes in 1 trial	Difference of mean change from baseline to post-treatment: 19 outcomes in 1 trial
Chen[29]	Psoriasis	13	78	Pathophysiological Manifestations; • Clinical assessment; • Global disease severity: 43 outcomes in 11 trials	Patients were reviewed once weekly during the study and monthly after clearance for 12 months: 3 outcomes in 1 trial	13 weeks: 1 outcome in 1 trial	Comparison of event rate: 1 outcome in 1 trial
Chi[30]	Lichen sclerosus (genital)	7	26	Life Impact; • Perception of health; • Single symptoms: 13 outcomes in 6 trials	Classified based on treatment response according to persistent/unchanged (scores remaining the same), improved (final scores lower than initial), in remission (no symptoms/ lesions): 4 outcomes in 1 trial	13 weeks: 4 outcomes in 1 trial	Difference of mean change from baseline to post-treatment: 4 outcomes in 1 trial
Eekhof[31]	Ingrowing toenails	21	93	Pathophysiological Manifestations; • Clinical assessment; • Global disease severity: 54 outcomes in 21 trials	Recurrence was defined as evidence of ingrowth of the nail edge or spicule formation: 4 outcomes in 2 trials	52 weeks: 2 outcomes in 2 trails	Comparison of event rate: 2 outcomes in 2 trails
Ersser[32]	Atopic eczema	9	40	Pathophysiological Manifestations; • Clinical assessment; • Combinations of two or more signs: 8 outcomes in 6 trials	SCORAD: 6 outcomes in 6 trials	12 weeks: 2 outcomes in 2 trials	Mean difference between 2 groups post-treatment: 1 outcome in 1 trial

Table 4 *Continued*

Review	Condition	*N* trials	*N* outcomes	Most frequent outcome construct (domain, sub-domain: 1st level 2nd level)	Most frequent instrument†	Most frequent time of measurement‡, *N* outcomes (%)	Most frequent statistic *N* outcomes§ (%)
Kastarinen[33]	Seborrhoeic dermatitis (face and scalp)	35	210	Life Impact; • Perception of health; • Single symptoms: 79 outcomes in 31 trials	4 point score (0 = absent, 1 = slight, 2 = moderate, 3 = severe): 7 outcomes in 6 trials	4 weeks: 2 outcomes in 1 trials	Comparison of event rate: 2 outcomes in 1 trials
Kwok[34]	Common cutaneous warts (non-genital)	70	150	Pathophysiological Manifestations; • Clinical assessment; • Global disease severity: 143 outcomes in 66 trials	Clinical examination: 9 outcomes in 5 trials	8 weeks: 5 outcomes in 2 trials	Comparison of event rate: 5 outcomes in 2 trials
Martin[35]	Pemphigus vulgaris and pemphigus foliaceus	10	59	Pathophysiological Manifestations; • Clinical assessment; • Single signs: 26 outcomes in 8 trials	Complete response was defined as lesion-free state, while the patient was receiving a minimum dose of steroid according to the treatment protocol: 6 outcomes in 1 trial	52 weeks: 6 outcomes in 1 trial	Comparison of event rate: 5 outcomes in 1 trial
Sasse[36]	Metastatic malignant melanoma	18	139	Pathophysiological Manifestations; • Clinical assessment; • Combinations of two or more signs: 52 outcomes in 15 trials	WHO Criteria: 31 outcomes in 8 trials	9 weeks: 6 outcomes in 2 trials	Comparison of event rate: 6 outcomes in 2 trials

NA, Not available.
†In most frequent outcome construct.
‡In most frequent outcome construct and their most frequent measurement instrument.
§In most frequent outcome construct and their most frequent measurement instrument at the most frequent time of measurement.

In this investigation, only efficacy and effectiveness outcomes were included. Therefore, *outcome domains*, which typically also refer to safety outcomes, were not represented in the proposed taxonomy and should be included at a later stage of development.

When reviewing the proposed taxonomy, it must be acknowledged that *sub-domain* differentiation is subjective. For example, the HOME initiative considered clinical signs and disease symptoms to be core *outcome domains*[25] whereas the recently published update of the OMERACT 2.1 filter considers both *sub-domains* as a part of clinical manifestations without explicit reflection of who determined this clinical manifestation and whether it is visible (clinical signs of skin diseases) or not (symptoms).[26]

Additionally, only two raters performed the iterative allocation process, although all authors reviewed the results and contributed to the conceptual framework. Despite these limitations, this endeavour provides the first crucial steps towards providing a dermatology-specific taxonomy.

Implications for further research

To the best of our knowledge, this study represents the first attempt to provide a framework and resource for an extension of the general taxonomy of Dodd *et al*.[19] to a specific medical area: dermatology. *CS-COUSIN* plans to further expand the current state of development to support and accelerate the extension of a dermatology-specific taxonomy as a resource for study trialists, reviewers and *COS* developers (www.cs-cousin.org/). The next step of this development process will be an extension and possible revision of the *outcome domain* mapping by considering further conditions with respect to efficacy/effectiveness, safety, and biomarker outcomes to complement the dermatology-specific taxonomy. Outcome lists of *COS* groups or systematic reviews of outcome research in dermatology will be a resource for this.

In order to reduce the diversity of terminologies used by different *COS* initiatives and groups in dermatology, it is of crucial importance to provide a taxonomy with definitions at different *outcome domain* levels. Therefore, we propose a consensus process with relevant clinical experts, methodologists and patient representatives to develop a consensus based dermatology-specific extension of the general taxonomy proposed by the *COMET* group.

After this comprehensive process of taxonomy development, it is necessary to support the translation of the research results, assess the extent of feasibility and support capability in dermatology research. Furthermore, it would be meaningful to examine the transferability of the development process of the aforementioned suggestions to other medical areas.

Conclusion

In this study, we mapped outcomes to the *OMERACT 2.0 Filter* and provide a resource of potentially relevant *outcome domains* for dermatological trials. This hierarchically structured *outcome domain* list may be used as a starting point for a dermatology-specific extension of the general taxonomy proposed by the *COMET* group to provide a framework for trialists, reviewers and *COS* developers to facilitate reporting and evidence synthesis of trial outcomes. Even if similar outcome constructs are measured in trials, there is substantial variation regarding use of *measurement instruments*, *time of measurements* and *methods of aggregation*. These variations do not allow firm meta-analysis or qualitative comparison of evidence from different trials, indicating that only defining *core outcome domains* and measurement instruments are not sufficient for meaningful evidence synthesis. *COS* should go further, and at least standardize the *time of measurements*, the specific *analysis metric* and the *method of aggregation*. For dermatology, there is a clear need for outcome taxonomy and a corresponding standardized reporting terminology.

Acknowledgements

We thank the Cochrane Skin management team for providing access to *Archie* data and Sanna Prinsen for her participation in the data extraction. Open access funding enabled and organized by Projekt DEAL.

Author contribution

JS and TL jointly conceived the study, TL performed the analysis, TL and TD drafted the paper, and JK, CP, TW, EH, CA, SB, AD, EBT, NKR, PS, MG, LJ and JS extracted the outcome data. JS and TL contributed to development of the dermatology-specific taxonomy, and all authors commented on the manuscript and approved the final version.

Availability of data and material

The data sets used and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1 Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't – it's about integrating individual clinical expertise and the best external evidence. *Br Med J* 1996; **312**: 71–72.
2 Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007; **8**: 39.
3 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; **374**: 86–89.
4 Schmitt J, Lange T, Kottner J *et al*. Cochrane reviews and dermatological trials outcome concordance: why core outcome sets could make trial results more usable. *J Invest Dermatol* 2019; **139**: 1045–1053.
5 Williamson PR, Altman DG, Blazeby JM *et al*. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012; **13**: 132.
6 Schmitt J, Deckert S, Alam M *et al*. Report from the kick-off meeting of the Cochrane Skin Group Core Outcome Set Initiative (CSG-COUSIN). *Br J Dermatol* 2016; **174**: 287–295.
7 Prinsen CA, Spuls PI, Kottner J *et al*. Navigating the landscape of core outcome set development in dermatology. *J Am Acad Dermatol* 2019; **81**: 297–305.

8 Schmitt J, Apfelbacher C, Spuls PI *et al*. The Harmonizing Outcome Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. *J Invest Dermatol* 2015; **135**: 24–30.
9 Chalmers JR, Thomas KS, Apfelbacher C *et al*. Report from the fifth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *Br J Dermatol* 2018; **178**: e332–e341.
10 Busard CI, Nolte JYC, Pasch MC, Spuls PI. Reporting of outcomes in randomized controlled trials on nail psoriasis: a systematic review. *Br J Dermatol* 2018; **178**: 640–649.
11 Eleftheriadou V, Thomas K, van Geel N *et al*. Developing core outcome set for vitiligo clinical trials: international e-Delphi consensus. *Pigment Cell Melanoma Res* 2015; **28**: 363–369.
12 Schlessinger DI, Iyengar S, Yanes AF *et al*. Development of a core outcome set for clinical trials in basal cell carcinoma: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. *Trials* 2017; **18**: 490.
13 Horbach SER, van der Horst C, Blei F *et al*. Development of an international core outcome set for peripheral vascular malformations: the OVAMA project. *Br J Dermatol* 2018; **178**: 473–481.
14 Oei W, Fledderus AC, Korfage I *et al*. Protocol for the development of core set of domains of the core outcome set for patients with congenital melanocytic naevi (OCOMEN project). *J Eur Acad Dermatol Venereol* 2019; **34**: 267–273.
15 Ingram JR, Hadjieconomou S, Piguet V. Development of core outcome sets in hidradenitis suppurativa: systematic review of outcome measure instruments to inform the process. *Br J Dermatol* 2016; **175**: 263–272.
16 Van den Bussche K, De Meyer D, Van Damme N, Kottner J, Beeckman D. CONSIDER – Core Outcome Set in IAD Research: study protocol for establishing a core set of outcomes and measurements in incontinence-associated dermatitis research. *J Adv Nurs* 2017; **73**: 2473–2483.
17 Schmitt J, Langan S, Williams HC, European Dermato-Epidemiology N. What are the best outcome measurements for atopic eczema? A systematic review. *J Allergy Clin Immunol* 2007; **120**: 1389–1398.
18 Boers M, Kirwan JR, Wells G *et al*. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *J Clin Epidemiol* 2014; **67**: 745–753.
19 Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol* 2018; **96**: 84–92.
20 Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database–update and key issues. *N Engl J Med* 2011; **364**: 852–860.
21 Chan AW, Tetzlaff JM, Gotzsche PC *et al*. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ* 2013; **346**: e7586.
22 Boers M, Beaton DE, Shea BJ *et al*. OMERACT Filter 2.1: elaboration of the conceptual framework for outcome measurement in health intervention studies. *J Rheumatol* 2019; **46**: 1021–1027.
23 Kottner J, Schmitt J. Core outcome sets in dermatology: next steps. *Brit J Dermatol* 2018; **179**: 549–550.
24 Williamson PR, Altman DG, Bagley H *et al*. The COMET handbook: version 1.0. *Trials* 2017; **18**: 280.
25 Kottner J, Jacobi L, Hahnel E *et al*. Core outcome sets in dermatology: report from the second meeting of the International Cochrane Skin Group Core Outcome Set Initiative. *Br J Dermatol* 2018; **178**: e279–e285.
26 Maxwell LJ, Beaton DE, Shea BJ *et al*. Core domain set selection according to OMERACT Filter 2.1: the OMERACT methodology. *J Rheumatol* 2019; **46**: 1014–1020.
27 Bamford JTM, Ray S, Musekiwa A, van Gool C, Humphreys R, Ernst E. Oral evening primrose oil and borage oil for eczema. *Cochrane Database Syst Rev* 2013; **4**: CD004416.
28 Bath-Hextall FJ, Jenkinson C, Humphreys R, Williams HC. Dietary supplements for established atopic eczema. *Cochrane Database Syst Rev* 2012; **2**: CD005205.
29 Chen X, Yang M, Cheng Y, Liu GJ, Zhang M. Narrow-band ultraviolet B phototherapy versus broad-band ultraviolet B or psoralen-ultraviolet A photochemotherapy for psoriasis. *Cochrane Database Syst Rev* 2013; **10**: CD009481.
30 Chi C-C, Kirtschig G, Baldo M, Brackenbury F, Lewis F, Wojnarowska F. Topical interventions for genital lichen sclerosus. *Cochrane Database Syst Rev* 2011; **12**: CD008240.
31 Eekhof JAH, Van Wijk B, Knuistingh Neven A, van der Wouden JC. Interventions for ingrowing toenails. *Cochrane Database Syst Rev* 2012; **4**: CD001541.
32 Ersser SJC, Latter F, Gardiner S *et al*. Psychological and educational interventions for atopic eczema in children. *Cochrane Database Syst Rev.* 2014; **3**: CD004054.
33 Kastarinen HO, Okokon T, Kiviniemi EO *et al*. Topical anti-inflammatory agents for seborrhoeic dermatitis of the face or scalp. *Cochrane Database Syst Rev* 2014; **5**: CD009446.
34 Kwok CS, Gibbs S, Bennett C, Holland R, Abbott R. Topical treatments for cutaneous warts. *Cochrane Database Syst Rev* 2012; **9**: CD001781.
35 Martin LK, Agero AL, Werth V, Villanueva E, Segall J, Murrell DF. Interventions for pemphigus vulgaris and pemphigus foliaceus. *Cochrane Database Syst Rev* 2009; **1**: CD006263.
36 Sasse AD, Sasse EC, Clark LGO, Ulloa L, Clark OAC. Chemoimmunotherapy versus chemotherapy for metastatic malignant melanoma. *Cochrane Database Syst Rev* 2007; **1**: CD005413.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Iterative classification process – performed by two reviewer.
Figure S2. Flow chart.