Minutes from the 2018 Annual Cochrane Skin/CSG-COUSIN Meeting

Allard Pierson Museum, Oude Turfmarkt 127, 1012 GC Amsterdam, Nina van Leerzaal

Presentations from the meeting will be uploaded to the following webpages:
http://skin.cochrane.org/csg-annual-meeting-2018

https://www.uniklinikum-dresden.de/de/das-klinikum/universitaetscentren/zegv/cousin/meetings/meetings

15th January 2018

Concordance between primary and secondary outcomes in Cochrane systematic reviews and underlying clinical trials in Dermatology – a meta-epidemiologic study (Jochen Schmitt)

Objectives: (i) to assess concordance between Cochrane Review outcomes and those used in the clinical trials cited in the review; (ii) to see if meta-analysis differs depending whether data are concordant or discordant.

Ten Cochrane Reviews were assessed. Twenty-one primary effectiveness outcomes were used in the reviews: 5 were not assessed in a single trial, 6 outcomes were assessed in less than half of the trials. Adverse event outcomes, which are not dermatology specific, were not included.

All reviews noted outcome reporting as a barrier in the systematic review. Cochrane reviewers often did not define primary review outcomes precisely. Most primary studies didn't define their outcomes in terms of what was primary, secondary, etc. There was very low concordance between the Cochrane Review and included trial primary outcomes. There were not enough trials for meta-epidemiologic analysis. Good concordance was found at domain level (core

areas: life impact, resource use/economic impact/ pathophysiological manifestations). The results could be mapped to the three OMERACT domains to see what is missing in Skin reviews.

The SPIRIT guideline explains what outcomes should consist of. The hierarchy of outcome specification is as follows:

- What to measure: area – domain – outcome
- How to measure: time of measurement – measurement instrument – statistics to use

Limitations of the study include inconsistent definitions of outcomes, domains, and measurement instruments and the random sample.

Conclusions:

- Core outcome sets (COS) should include domains and measurement instrument – one alone is not enough
- Taxonomy for outcomes in dermatology is recommended
- Outcome assessment and reporting is a major problem

Discussion: patient-reported outcomes need to catch up more. To what extent is the agenda driven by the FDA? Had the team looked at the review protocols and seen if the outcomes listed as primary remained so in the review? Ben Goldacre has a good website that compares switched outcomes (http://compare-trials.org/)

Presentation of core outcomes methods advice and guidance (Jan Kottner)

Due to increased interest in COS and the challenge of keeping guidance up-to-date/heterogeneous guidance, CSG-COUSIN was founded to support COS in dermatology to help standardise trial outcomes to aid comparability in systematic reviews. The first guidance document was published in 2016 and is updated when new resources are available.

The COUSIN domain development process is as follows: complete a CSG-COUSIN proposal form, publish the protocol, produce a list of core outcome domains then outcome measurement instruments, and then select one instrument per domain using formal consensus methods. Selecting domains may be easier than selecting instruments.

The team requested feedback. In the first instance, they plan to ensure they're aligned with existing COUSIN project groups, review and develop processes and documents, and provide methods advice to COS groups. They also wish to further develop standards for instrument development, advance patient involvement methods, and keep abreast of methodological advances in COS.

Why should people work with CSG-COUSIN? They have methods/contents experts unique within dermatology; capacity for implementation directly in systematic reviews to inform research; and are well connected to other initiatives, like COMET.

Discussion: there's a need to expand outside of Cochrane and focus on dissemination to ensure implementation. People working on COS outside COUSIN can't register protocols retrospectively.

COS project updates – 7 presentations by project groups

1. Core Outcome Set for Incontinence - associated dermatitis (IAD) (CONSIDER): Three and nine item scales in Delphi studies lead to different conclusion (Jan Kottner)

This project was based on COS for IAD, which is at the domain identification stage, using Delphi methods. As there are some open questions in a Delphi study, scales are allocated thresholds, which leads to a loss of information. The

objective of the CONSIDER project was to compare a 3- or 9-item scale in scoring outcome domains. Two groups of participants were given different options: one group had three options, another group had nine options.

Disagreement was found in response to classifying as not important, important, or critical. There was much higher variation in responses especially in domains when between-group disagreement was present. The answer options (3 or 9) seem to affect the responses. Variability of the 9-item scale is related to disagreement. It would be helpful to reproduce this study in different COS projects.

Discussion: did authors check how participants interpreted critical? (Standard terminology was used.)

2. Involving Service-Users in the Outcome for Pressure Ulcer Trials (OUTPUTs) Project (Susanne Coleman)

The objective of OUTPUTs (**Ou**tcomes for **P**ressure **U**lcer **T**rial**s**) is to find the pressure ulcer outcomes of most importance to service users and preferred service user involvement, in collaboration with PURSUN (**P**ressure **U**lcer **R**esearch **S**ervice **U**ser **N**etwork).

To collect data, the researchers gave a study overview of the OUTPUTS project and explanation of trial outcomes to the service users; then they presented the scoping review outcomes and facilitated group work to help participants consider outcomes, which was fed back to the full group.

They found service users were confused about how outcomes might differ in clinical practice, treatment, and prevention trials. There was difficulty differentiating between treatment and outcomes, and users had differing opinions depending on their experience (those at risk of pressure ulcers may have different thoughts).

Future work will expand on the methods approach, possibly having one-to-one interviews, plus another meeting to further think about user involvement.

Discussion: covered the potential to extend this to other patient groups, and whether prevention outcomes need different guidance than what COMET provides.

3. A core outcome set for hidradenitis suppurativa trials (Linnea Thorlacius)

An overview of work done so far in this project: six groups of stakeholders were invited to a meeting; the researchers aimed for a 1:1 ratio of patients and healthcare (HC) professionals. Candidate items and potential core outcome domains were presented (complied from a systematic review of the literature, qualitative studies, healthcare (HC) professional item generation survey) and put through an E-delphi round, with new items also suggestable.

Results were categorised into domains/items that 1) reached 'consensus in' by both patients and HC professionals, 2) reached 'consensus in' by patients OR HC profs, and 3) items that did not reach 'consensus in' but was considered important.

Discussion: should the number of Delphi rounds be limited? Concentrate on a few domains at a time? (Perhaps better to not limit, but ensure the project has a co-ordinator).

4. Workshop for outcome measure for target lesion in vitiligo: results of patients workshop at the world vitiligo day (Khaled Ezzedine)

The vitiligo COS group identified re-pigmentation as a core outcome. Re-pigmentation was assessed by patients in only 4% of trials. 96% reported re-pigmentation generally, but used 48 different measures. Clinicians and patients were surveyed: patients wanted cosmetically acceptable re-pigmentation; they were not concerned with re-pigmentation measured by per cent. An e-delphi consensus found no agreement on the best outcome measure for re-pigmentation for target lesions.

The aim of the workshop was to garner patient feedback on how and when to measure re-pigmentation. Consensus was defined if at least 60% of patients agreed. The group looked at re- pigmentation and characteristics, such as 'what is successful re-pigmentation in patients' opinion?', 'What parts of the body are most important to patients in terms of gaining the colour back?' etc. Both noticeability scale and per cent re-pigmentation should be used in measuring treatment results rather than value-laden terms such as "cosmetically acceptable".

The next steps are to determine the best scale to use. Other domains will be assessed in the future.

Discussion: adverse events – ensuring all are caught and definitions of objective/subjective assessments. The patient voice is important for obtaining funding for COS research.

5. What are the best ways to involve patients in developing core outcome sets; experiences and reflections of the HOME initiative (Jo Chalmers)

HOME (**H**armonising **O**utcome **M**easures for **E**czema) has agreed four core domains (with instruments recommended for 2 of these). Patients were included in all aspects of consensus (e.g., e-delphi, consensus meetings). The team reflected on whether the patient input was meaningful.

Patients from HOME V consensus meeting had semi-structured interviews by phone/video call, by an independent researcher. They gave largely positive feedback, but highlighted areas for improvement. Factors that affected their perceived involvement were as follows: health at the meeting, perceived benefits of input, perception of "roles" at the meeting, sense of belonging, navigating the world of COS.

General principles for involvement include ensuring that patients are welcomed and supported, that they are given information With regard to expectations, it is important that patients are reassured that their lived experience is what's needed, and that they are provided with pre-reading material (hard copies at least 2 weeks before the meeting, with guidance on where to focus and lay summaries). It is also important to cater for the medical needs of patients (eg privacy to apply emollients) and ensure that facilitators encourage patient participation, allowing for input in a group setting (have different ways to input) and having a "jargon monitor".

HOME hold a pre-meeting, but patients said they want more time between the pre-meeting and main meeting to digest and reflect on what had been covered.

Discussion: concerns were raised that Pharma may try to influence patient. It is important that COS groups remain independent but still involve industry as stakeholders.

6. Development of a Core Outcome Set for Basal Cell Carcinoma: A Systematic Review and Delphi Survey (Amanda Maisel)

The group aim to use systematic literature review, stakeholder involvement, and consensus process to find the most important outcomes for the most common cancer. From over 500 outcomes, 60 outcomes have been retained. The next step is to present this in 2 rounds of Delphi survey followed by a consensus meeting.

Discussion: it was suggested that surgical interventions are not over-emphasised given other forms of therapy such as creams. It was also important to take into account how BCC in an aged population might be treated differently (i.e. not treated) with reference to the work of Eleni Linos (*Linos E, Chren MM, Stijacic Cenzer I, Covinsky KE. Skin Cancer in U.S. Elderly Adults: Does Life Expectancy Play a Role in Treatment Decisions? J Am Geriatr Soc. 2016 Aug;64(8):1610-5*)

7. Challenges encountered in the Acne Core Outcomes Research Network (ACORN) (Diane Thiboutot)

Acne core outcome domains have been established. The next steps include a systematic review of quality of life (QoL) in acne (how many studies use QoL instruments) and appraising existing measures for suitability.

Questions include: how many core outcome measures are needed? Does long-term control need a novel approach? Signs and symptoms – what to measure within these and who will assess? What would be included in a new HRQOL measure or PROM?

Challenges might be engaging patients effectively and resolving differences of opinion, as well as obtaining funding. Potential solutions include developing and sharing resources to inform and engage stakeholders in COS development process, web-based platforms to engage an international audience, and engaging patient advocacy groups to find novel ways to track long-term disease.

Discussion: some questioned whether the project was biting off too much with so many (6/7) domains? Having the EADV publish QoL results prior to this COS group was actually quite reassuring to the project team. There is less representation from teenage patients, which is a challenge.

The value of item banks and PROMIS for COS development (Caroline Terwee)

PROMIS (Patient-Reported Outcomes Measurement Information System) works towards the standardisation of patient-reported outcomes.
http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis

The challenges with existing PROMs (patient-reported outcome measures) include questionnaires too long to complete; patients are asked to complete all questions, but not all questions relevant to all patients; scores are difficult to interpret because of ordinal scales; scores are incomparable across different PROMs because they use their own scale; and the quality of PROMs is questionable.

Solution: use ITEM banks, which enable use of a Computerised Adaptive Test (CAT). ITEM banks encompass a lot of questions, which are categorised, and these are used to create population-specific short forms, which can be customised. They are sustainable – can be further improved as and when.

A CAT is used to select questions depending on the answer to the previous question, and results are ascribed confidence intervals. After 7 questions, a score is given that is usually more reliable than a 20-item questionnaire. Patients get less – but more relevant – questions.

ITEM banks start with systematic reviews to find all existing outcomes, patients (using focus groups/interviews) are consulted to ensure relevance and comprehensiveness, items are rewritten in a consistent format, and statistics are undertaken to ensure discrimination of questions.

Scores of different PROMs can be made comparable for meta-analysis. PROMIS is the largest example of an item-based bank. The generic configuration can be universally applied across patient populations – which is important because of multimorbidity (and this influences treatment effect and healthcare decision making). The universal nature of outcomes is underlined by the fact the same core outcomes appear in a lot of sets. Scores of PROMIS instruments are comparable within and between populations.

In future, common outcomes could be measured in all patients using generic item response theory-based item banks, which are more responsive (more precise) than traditional generic measures, with disease-specific outcomes added. Also, validation in every population of generic item banks isn't necessary. Differential item functioning can be evaluated across populations – items can be edited in response to the population.

IRT-based item banks offer solutions to problems presented by traditional PROMS. Applicable across patient populations

(Footage of Caroline's presentation can be found here: https://www.youtube.com/watch?v=mBUf2PmHYZ0)

Discussion: Computerised Adaptive Test technology can be harnessed through social media, but raises ethical issues of what groups do with the results. Working on applicability throughout the life span, rather than for children or adults. Translation of the tool? Consideration of anomalies is built into the computer.

Day 2

16th January 2018

Interventions for nail psoriasis – the challenges of multiple outcomes (Phyllis Spuls, Emma Mead)

Phyllis (author of the Cochrane Review 'Interventions for nail psoriasis') gave a description of the condition and discussed how there are many (unvalidated) outcome instruments to measure nail psoriasis. This presented challenges for the review: whether to pool through dosages, time point, or NAPSI used? (Footage of Phyllis' presentation can be found here: https://www.youtube.com/watch?v=IKoMVGso_v0&feature=youtu.be)

Emma, methodologist with Cochrane Skin, explained this was a common issue in skin reviews. Because skin reviews are often broad "all interventions for" reviews, they include lots of comparisons, interventions, and outcomes (with multiple measures/time points).

Solutions: prioritise the main comparisons at protocol stage, as this will help to construct the summary sections of the review.

Dealing with large reviews with multiple comparisons and outcomes: prioritise to enable the writing of the summary sections. Do this by deciding beforehand what will go in the 'Summary of findings' tables (SoFs): choose the important comparisons & up to seven outcomes and decide how to deal with multiple time points and scales. Prioritise further if a lot of SoFs. The **Plain language summary** and the **Summary of main results** should reflect the Abstract

Dealing with outcomes measured using different outcomes/scales: restrict to certain time points or measures. Will there be a hierarchy? Or for different measures use the standardised mean difference (SMD) for continuous data where many tools are used to measure outcomes.

Discussion: it was suggested that what is planned in a protocol is often too idealised - the data doesn't fit the plan. COUSIN will help move away from the need to use SMD. Colleagues were encouraged not to be scared to meta-analyse, but interpret results with caution.

Patients Included (Peter Smart, presented by Kim Thomas)

How to ensure research includes patients: make sure there is consumer input at the planning stage to ensure the relevance and scope of the research and to gain access to other patients. Later in the process, patients can help write better papers, e.g. by checking language issues, and post-publication, they can guarantee more patient-friendly presentation of the work and access to patient networks.

Cochrane Skin prioritisation project, and future plans (Bob Boyle)

There has been a boom in systematic reviews. Cochrane Reviews are now a small per cent of dermatology systematic reviews (SRs).

Systematic review threats: 40% have misleading conclusions due to not taking into account limitations of the studies. 67% of SRs have at least 1 overlapping review published within 3 years.

Going forward, Cochrane aims to support less reviews, but high-quality, high priority, speedy reviews. Part of the future strategy is to not insist on routine updating, and produce more targeted reviews. The editorial process will be tougher: reviews not ready for peer review are given only one opportunity to revise. There will be a faster editorial turnaround and publication times. More complex methods support will be offered by Cochrane, eg the Complex reviews support unit http://www.nihrcrsu.org/ .

The Skin Group's 2017 prioritisation project gathered suggested titles of importance from stakeholders, which were then summarised and ranked by the group's editors. The 7 highest scoring titles were shortlisted for priority support.

Further new title suggestions are welcome, but other than in exceptional circumstances, these will be reviewed annually, with a repeat of the formal prioritisation exercise in 2019, i.e., every 2 to 3 years.


Incorporating synthesis of non-randomised study designs (Ben Carter, Emma Mead)

The final session was a highly interactive and thought-provoking workshop on using non-randomised studies in systematic reviews. Up until now, Cochrane skin have only published one systematic review which includes non-RCTs as included studies: Safety of topical corticosteroids in pregnancy (Chi CC et al 2015). Eight published reviews/updates have searched for adverse events in non-RCTs but not as included studies. There is one protocol in development which plans to include non-RCT evidence: "*Interventions for treatment of Stevens-Johnson Syndrome (SJS), Toxic Epidermal Necrolysis (TEN) and SJS/TEN overlap syndrome"* (currently in the editorial process). We were encouraged to consider including non-randomised studies in Cochrane Reviews, where appropriate for answering the research question. Examples of this would include when RCTs are not feasible for the intervention of interest (e.g. surgery), when the outcomes are difficult to assess in RCTs (e.g. rare adverse events), and when the quality of RCT evidence is poor. There is more support available to authors who want to include non-RCTs in their Cochrane Reviews, such as from the NIHR complex review support unit. There are also a number of quality assessment tools for non-RCTs including ROBINS-I and the Newcastle-Ottawa Scale. GRADE can also be used to assess the quality of outcomes from non-RCTs.

Next meeting: Paris, France, 7th - 8th January 2019.