

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/105989/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Callander, J., Anstey, A.V., Ingram, John R. , Limpens, J., Flohr, C. and Spuls, P.I. 2017. How to write a Critically Appraised Topic: evidence to underpin routine clinical practice. *British Journal of Dermatology* 177 (4) , pp. 1007-1013. 10.1111/bjd.15873

Publishers page: <http://dx.doi.org/10.1111/bjd.15873>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



How to write a Critically Appraised Topic (CAT): evidence to underpin routine clinical practice

Callander J¹, Anstey AV¹, Flohr C², Ingram JR,³ Limpens J,⁴ Spuls P⁴

Department of Dermatology¹, Ysbyty Gwynedd, Penrhosgarnedd, Bangor, UK

Unit for Population-Based Dermatology Research², St John's Institute of Dermatology, Guy's & St Thomas' NHS Foundation Trust and King's College London, London, UK

Department of Dermatology & Academic Wound Healing,³ Division of Infection & Immunity, Cardiff University, Cardiff

Department of Dermatology⁴, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

Correspondence:

Phyllis I. Spuls, Department of Dermatology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

Tel.: +31 20 5664763

Fax: +31 20 5669152

E-mail: ph.i.spuls@amc.uva.nl

Funding sources: none

Conflict of interest: The authors state there are no other conflicts of interest for this study.

Word count text: 3131

Manuscript table count: 3

Manuscript figure count: 1

What are 'Critically Appraised Topics'?

Routine clinical practice often generates challenging questions about how to provide the best care for individual patients. For example, as dermatologists we may wonder whether ultrapotent topical corticosteroids are more effective in the treatment of alopecia areata than triamcinolone injections. Sometimes such questions might come directly from our patients: 'Doctor, I have read about acne treatments on the internet. Is it true that isotretinoin causes depression?' or 'How safe would maintenance ultraviolet B light therapy be to control my atopic eczema?'

Guidelines rarely provide the answers to these very specific questions. A systematic review is not feasible for every question that arises in the clinic, and quick internet searches are neither comprehensive nor reliable.

This is the gap that critically appraised topics (CATs) address. CATs are standardized summaries which draw together the best available evidence to answer questions based on real clinical scenarios often arising from specific patient encounters. CATs embrace the principles of evidence-based medicine (EBM) by following all four steps in EBM steps, but in a less rigorous and time consuming way. Since CATs focus on specific clinical scenarios, this creates a synthesis that generates new insights and knowledge.

Why are CATs useful?

CATs ensure that we use the highest quality, most up-to-date evidence help us address questions that arise organically and repeatedly in clinical practice. CATs help us to benefit from the rapidly increasing but potentially overwhelming volume of information now available.¹ Additionally, they help readers and authors develop skills in critically appraising literature.

CATs are not intended to replace rigorously designed and comprehensive systematic reviews, as they do not attempt to include all of the available evidence. Systematic reviews remain the gold standard for summarizing evidence but are often too time-consuming and costly to carry out for all clinical questions. In contrast, CATs are more focused, quicker to complete and can be done with a small team of authors at minimal cost. CATs, in comparison, are a more practical tool to address the many questions that are not considered by systematic reviews and guidelines. Like systematic reviews, CATs facilitate the integration of clinical research with clinical practice and avoid or minimise biases inherent with narrative reviews or internet searches.

CATs: background and future vision

CATs are well established in various specialities including radiology, neurology, emergency medicine and palliative care.²⁻⁵ They were introduced to dermatology in 2007⁶, and subsequently embraced by the *BJD* as 'short summaries of evidence on a topic of interest....focused around a clinical scenario'.⁷ Now, alongside 'Putting Papers into Practice' and Systematic Reviews, CATs form a key element of the Evidence-Based Dermatology section of the *BJD*. This article aims to help dermatologists and residents carry out their own CATs by describing the steps involved. Our vision is for CATs to become an increasingly important element of this section of the journal, creating an accessible resource of relevance and utility to daily clinical practice globally.

How to perform a CAT

There are four main steps to complete a CAT, identical to the steps in EBM. the first of which is to generate a structured question from the clinical encounter using the 'Population, Intervention, Comparator, Outcome' or 'PICO' format (table 1). Each of the four steps will now be described in more detail.

1. Formation of a focused and answerable question based on a clinical encounter
2. Search for the best available evidence
3. Critical appraisal of the evidence for validity and clinical relevance
4. Interpreting and applying the results to clinical practice

Table 1: Steps in performing a critically appraised topic (CAT).

Formation of a PICO question

Most questions arising from clinical practice relate to a specific treatment, diagnostic test or harmful agent. The systematic search for evidence about a topic requires a carefully formulated question

Population [P]: The PICO question requires a definition of the group of people with a disease or characteristic of interest. As the evidence must be applicable to this population, the P should be well specified . For example, age or gender, disease stage, disease severity, disease localisation or extent, related symptoms, co-morbidities, the effect of previous treatments, and current or future pregnancy could each be considered important in the scenario and should therefore be specified in the patient population component. An example of a specific population to use in a CAT is: a paediatric patient with severe psoriasis, unresponsive to topical treatments.

Intervention(s) [I]: Originally, when the acronym PICO was developed I related to treatments, but I can relate to a risk factor in questions about aetiology or prognosis and a diagnostic test in diagnosis. The I may include specific details such as treatment dose, duration and frequency of a therapy, mode of administration or a combination of therapies. Alternatively, it might concern the timing of a specific diagnostic test or a combination of tests. An example of an intervention in our case would be: treatment with narrow band UVB phototherapy three times per week for eight weeks

Comparator [C]: The comparator in a PICO question concerns the treatment or diagnostic test that the intervention treatment or test is compared to. For example, the specific alternative treatment might be no treatment or placebo or other possible treatments, such as oral methotrexate for our paediatric patient with psoriasis. When analysing the utility of a specific diagnostic test, a reference standard diagnostic test should be chosen as the comparator, such as procollagen III and liver function tests measured every three months compared to liver function tests alone.

Outcome: The outcome in a PICO question requires definition and might be short- or long-term physician assessed efficacy, specific patient-reported outcomes, adverse event rates or diagnostic accuracy (sensitivity and specificity). Outcomes in a PICO question may be multiple and can be divided into critically important, important and less important. An example of an outcome in a PICO question would be: reduction in PASI by 75% or greater after three months of treatment compared to baseline. In questions about harm the outcome is often a disease.

In summary, the careful construction of a clinical question using the PICO format ensures that the CAT remains linked to the original clinical scenario encountered and facilitates subsequent focused searching for the relevant evidence. Examples of PICO questions from CATs published in the *BJD* are shown in Table 2.⁸⁻¹²

Finding the best available evidence: where to start searching

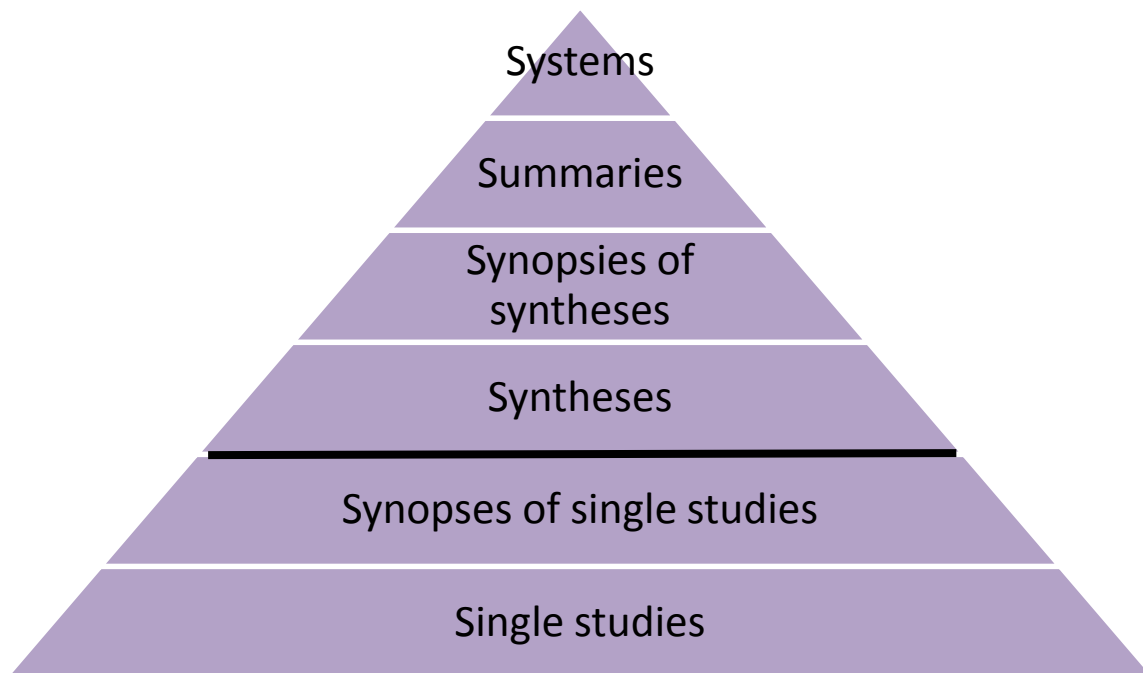


Figure 1. 6S Evidence pyramid adapted from DiCenso A, Bayley L and Haynes RB¹³ RCT = randomised controlled trial.

Evidence pyramids, like the 6S Pyramid (Fig. 1), provide structures for finding suitable evidence to answer PICO questions.¹⁴ Further explanation and details about where you can find evidence from each level are given in Table 3. The upper levels of the pyramid relate to “aggregated” evidence, such as up-to-date evidence-based guidelines and systematic reviews, whilst the lower levels relate to individual original studies. At the top of the pyramid, the pre-appraised, synthesised information is more desirable and convenient for answering a question in a limited time period. However, for very specific PICO questions, evidence is more likely to be found at the lower levels of the pyramid.

At the base of the pyramid are single studies. There are many databases which can be searched in order to find single studies (see Table 3). It is not mandatory to use more than one, thus a search in PubMed alone is sufficient. However, if you *are* searching more than one database, there are ways to prevent duplication of search results using citation managers such as ‘Refworks’ (www.refworks.com), ENDNOTE, or ‘Mendeley’ (www.mendeley.com).

It is helpful to remember that disease-specific databases also exist. These include the GREAT database for atopic eczema (www.greatdatabase.org.uk).

In practice, the best approach for a CAT is to check aggregate evidence first. If no evidence answering the PICO is found, the next step is to search for single studies. If a systematic review or other aggregated evidence is found it can be appraised and included in the CAT. If the aggregate evidence is older, single studies can be searched from the search date of this study.

Finding the best available evidence: the search terms

Based on the focused PICO question, the key search terms can then be identified. These are the terms to be entered into online databases and determine the relevance of the results. Often using the search terms for the P and I part of the PICO question is sufficient. For aggregated evidence fewer elements (just the P or the I) may do. In case of a large number of hits, additional elements can be added to refine the search results.

It is important to include alternative names for the disease, diagnostic test or treatment to avoid missing relevant evidence. It is also important to consider alternative spellings. The Boolean operator “OR” is used to combine search terms, for example ‘spider anima’ OR ‘spider naevus’ OR ‘spider nevus’. Some databases allow truncation (command: *) at the end of the term to allow for spelling variants.

In MEDLINE, PubMed, EMBASE and CENTRAL it is essential to make proper use of controlled terms (i.e. MeSH-terms in MEDLINE) and to use field codes, like [tiab] (title and abstract) to define the fields to be searched and to prevent automatic translation by PubMed. A PubMed search for atopic dermatitis should basically be as follows:

```
"Dermatitis, Atopic"[Mesh] OR atopic dermat*[tiab] OR atopic neurodermatit*[tiab] OR eczem*[tiab]
```

The final terms can then be combined using the Boolean operator “AND” to complete the search. Those who are inexperienced in performing this type of search should seek assistance and guidance from their local library Information Specialist. The sources the complete search strategy, key search terms, filters or limits applied and search date should all be documented and reported in the ‘methods’ section of a CAT. This also helps with future search updates.

Although most databases allow restriction to for instance language, study types (i.e. RCTs) or population (i.e. child), these limitations should be used with caution, since these may readily lead to missing relevant articles. It is more prudent to use broad search filters like the build-in clinical queries in PubMed, the Cochrane filters for identifying RCTs, paediatric search filters or to use no filter at all.

Critical appraisal of the evidence

Before performing your search, inclusion and exclusion criteria must be determined and documented. These criteria may include study type, minimum sample size and patient characteristics, treatment or test specifics, the comparator and outcomes of interest. If your search returns numerous results, these criteria will allow you to narrow them down to a small number of the most relevant articles to answer the clinical question. If the search relates to a commonly

encountered clinical question but has failed to find any relevant results, this suggests an area where research is needed (or the search is insufficient). It is important to bring this to the attention of other dermatologists and academics, but a meaningful CAT does require an evidence base that can be appraised. As a guide, the following section has been divided into three questions which should be addressed in relation to each piece of evidence:

1. Does this study address the PICO question?

For any piece of evidence, the important data should firstly be extracted and noted. This includes: author(s), year of publication, study design, patient characteristics, aspects of the disease such as duration and severity, details of the intervention(s) and their comparator(s) such as dose, frequency and duration of a treatment or timing of diagnostic tests. Finally the outcome(s) parameters, including instruments used to measure outcome and the timing of outcome measurement. By extracting these data it will become clear whether the study addresses the PICO question.

2. Were the study methods valid?

This question relates to internal validity. The internal validity of a study is the extent to which the study minimises bias and confounding factors. Some study designs are inherently better at minimising bias than others. For instance, well-performed randomised controlled trials (RCTs) are the gold standard to assess the effectiveness of a treatment. For questions about diagnostic accuracy, cross-sectional study designs are more appropriate.

In RCTs, the randomisation process removes the influence of known and unknown confounding factors. In addition, blinding of the physician and patient reduces performance and detection bias. Therefore, it is important to assess the methods of randomisation and blinding of patients, physicians and outcome assessors. Comparability of the study groups at baseline is also important to ensure they are not markedly different, despite the randomisation process and this can be improved by stratified randomisation. The proportion of participants lost from follow-up and use of intention to treat analysis should also be noted to assess for possible attrition bias. In addition to study design, the sample size, or number of participants, is important. If the study is not adequately powered then clinically important differences may not be detected.

Retrospective studies are inherently prone to selection bias and are also at risk of recall bias. At the base of the hierarchy of evidence are uncontrolled case series, case reports and expert opinions, all of which are at high risk of bias.

To aid the critical appraisal of different types of studies, a wide range of checklists are available, such as the Cochrane Risk of Bias tool for RCTs, Quality Assessment of Diagnostic Accuracy Studies (QUADAS) for diagnostic accuracy studies (<http://annals.org/aim/article/474994/quadas-2-revised-tool-quality-assessment-diagnostic-accuracy-studies>), the Newcastle-Ottawa Scale (NOS) for case-control and cohort studies related to aetiology and harm (http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp), and the Quality in Prognostic Studies (QUIPS) for cohort studies related to prognosis (<http://www.biomedcentral.com/content/supplementary/1546-0096-12-19-S1.pdf>), the AXIS tool for cross-sectional studies (<http://bmjopen.bmj.com/content/bmjopen/6/12/e011458.full.pdf>) and the CARE statement for case reports (<http://www.care-statement.org/>). In addition, websites such

as the Critical Appraisal Skills Programme (CASP) tools website, (www.casp-uk.net), and the Enhancing the Quality and Transparency Of health Research (EQUATOR) website (www.equator-network.org) provide links to checklists for critical appraisal of all main study types, which are listed in Table 4.

Study types
Randomised controlled trials
Observational studies <ul style="list-style-type: none"> • Cohort studies • Case control studies • Cross-sectional studies • Case series • Case reports

Table 4: A list of the main study types

When assessing synthesised evidence, such as systematic reviews, the reliability of the conclusion is based on the presence of all of the relevant evidence and absence of selection bias, as well as the validity of the included studies. It is essential to ensure that thorough search techniques, without inappropriate limits or exclusions, have been used. The search methods, as well as the author's critical appraisal of the included studies should be clearly documented. Reasons should be given for exclusion of any studies. Meta-analyses may have been used to combine trial results, however, meta-analyses should not be performed if there is substantial heterogeneity between the included studies. If results are pooled, appropriate weighting should be given to larger, more robust studies. Checklists are also available to aid critical appraisal of synthesised evidence, such as AMSTAR (https://amstar.ca/Amstar_Checklist.php) for systematic reviews and AGREE II (<http://www.agreetrust.org/agree-ii/>) for guidelines.

3. Are the results applicable to your patient?

If a study has minimal bias and used appropriate methods, it is then necessary to ensure that the results can be applied to the real life scenario that initiated the CAT. This means ensuring that the specific P, I, C, and O elements of the patient encounter are sufficiently comparable to the evidence provided by the studies. If the patient population in the studies is different from the patient in the PICO question then the results may not be relevant. Similarly, if the Intervention, Comparator or Outcomes are significantly different, then the results cannot be directly transferred to the original patient scenario.

A further resource to aid in appraising evidence and is the Centre for Evidence based Medicine's online worksheets (<http://www.cebm.net/critical-appraisal/>). These helpfully break down critical appraisal into the following parts:

- Does this study address a clearly focused question?
- Did the study use valid methods to address this question?
- Are the valid results of this study important?
- Are these valid, important results applicable to my patient or population?

Interpreting and applying the results

If the clinician is satisfied that the evidence is internally valid and applicable to their patient, the next step is to interpret and apply the results. For each piece of evidence, the following questions should be asked: What is the effect size of results? How significant and how precise are they? How will these results alter my patient care?

The specific results of interest relate to the outcome in the PICO question. In therapeutic studies, the absolute risk and the number needed to treat are usually the values of most relevance to clinical practice. In diagnostic studies, the most important values are the sensitivity and specificity of a test. Results of meta-analyses are often displayed in forest plots, showing the overall effect size (represented by the vertical axis of the black summary diamond) and associated confidence interval (the horizontal axis of the diamond) across all included studies.

A tool that can help to further assess the overall quality of the evidence and therefore the results is GRADE (Grading of Recommendations Assessment, Development and Evaluation). The GRADE system was developed to provide a transparent approach to grading quality of evidence and subsequent strength of recommendations. It uses a scoring system based on the study design, quality, consistency, directness and effect size to combine a number of studies, focusing on a specific outcome following a particular intervention, and determine how good the overall evidence is for a specific outcome. GRADE emphasises outcomes that are important to patients. The overall quality of evidence, and therefore certainty of the result, is rated as high, moderate, low or very low. Furthermore, information and tutorials can be found at www.gradeworkinggroup.org.¹⁵

Informed by the results of the CAT, a discussion with the patient can now take place to jointly agree the best management or diagnostic plan. It is this final stage which relates the evidence back to the patient and the clinical scenario from which the PICO question arose.

Summary

In summary, CATs are short structured summaries of critically appraised evidence based on patient encounters. They follow a systematic approach, can be produced in a relatively short time and should be easy to assimilate. However, CATs also have several disadvantages. Unlike systematic reviews, they only provide a selected synthesis of the overall evidence. In addition, the external generalisability of a CAT to any individual practice setting must be carefully considered. Nevertheless, a well-performed CAT, based on a question generated by a patient, provides evidence-based guidance to improve the care of patients in a similar situation. Performing the CAT process itself enhances our skills in building critical appraisal and our understanding of clinical topics. CATs can aid learning at a local level, via teaching sessions and journal clubs, and can impact nationally and internationally via publication in journals such as the *BJD*.

References

1. Sadeigh G, Parker R, Kelly A et al. How to write a critically appraised topic (CAT). *Acad Radiol* 2012;2012:19-872-888
2. Carroll A.G, Kavanagh R.G, Ni Leidjin C et al. Comparative Effectiveness of Imaging Modalities for the Diagnosis of Intestinal Obstruction in Neonates and Infants: A Critically Appraised Topic. *Academic Radiology*. 2016;23(5):559-568
3. Dhawan P.S, Goodman B.P, Harper C.M,et al. IVIG versus PLEX in the treatment of worsening myasthenia gravis: What is the evidence? A critically appraised topic. *Neurologist*.2015;19(5):145-148
4. Nishijima D.K., Zehtabchi S. Evidence-based Emergency Medicine/Critically Appraised Topic. The efficacy of recombinant activated factor VII in severe trauma. *Annals of emergency medicine*. 2009;54(5):737-744
5. Jones B.W. Critically Appraised Topic: The Effect of Hospice Care in Patients with Dementia. *American Journal of Hospice and Palliative Medicine*. 2013;30(6):590-593
6. Bigby M. Evidence-based dermatology section welcomes a new feature: critically appraised topic. *Arch Dermatol*. 2007;143(9):1185-1186
7. Batchelor JM, Langan SM. Introducing 'Critically Appraised Topics' *Br J Dermatol*. 2012;167(1):1
8. Kassim JM, Shipman AR, Szczecinska W et al. How effective is intralesional injection of triamcinolone acetonide compared with topical treatments in inducing and maintaining hair growth in patients with alopecia areata? A Critically Appraised Topic. *Br J Dermatol*. 2014;170(4):766-771
9. Patel AN, Simpson RC, Cohen SN. In a patient with an immunobullous disorder, is transportation of the skin biopsy in normal saline adequate for direct immunofluorescence analysis? A critically appraised topic. *Br J Dermatol*. 2013;169:6-10.
10. Garritsen FM, ter Haar NM, Spuls PI. House dust mite reduction in the management of atopic dermatitis. A critically appraised topic. *Br J Dermatol*. 2013;168(4):688-91
11. Matin RN, Acland KM, Williams HC. Is Mohs micrographic surgery more effective than wide local excision for treatment of dermatofibrosarcoma protuberans in reducing risk of local recurrence? A Critically Appraised Topic. *Br J Dermatol*. 2012;167:6-9.
12. [Garritsen F, van den Broek M, van Zuilen A et al.](#) Pregnancy and fetal outcomes after paternal exposure to azathioprine, methotrexate or mycophenolic acid: a critically appraised topic (CAT). *Br J Dermatol*. 2017. In press
13. Brian S Alper, R Brian Haynes. EBHC pyramid 5.0 for accessing preappraised evidence and guidance. *Evid Based Med* 2016;21:123-125.
14. DiCenso A, Bayley L, Haynes RB. Accessing pre-appraised evidence: fine tuning the 5S model into a 6S model. *Evidence-based Nursing*. 2009;12(4):99-101

15. Guyatt GH, Oxman AD, Vist GE et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal*. 2008 Apr 26;336(7650):924-6

Question	Population	Intervention	Comparison	Outcomes
How effective is intralesional injection of triamcinolone acetonide compared with topical treatments in inducing and maintaining hair growth in patients with alopecia areata? ⁸	Patients with alopecia areata	Intralesional triamcinolone acetonide	All topical treatments	Induction and maintenance of hair re-growth Frequency and reversibility of adverse events
In a patient with an immunobullous disorder, is transportation of the skin biopsy in normal saline adequate for direct immunofluorescence analysis? ⁹	Patients who have had a skin biopsy for an immunobullous disorder	Transport of the skin biopsy to the laboratory in saline	Transport of the skin biopsy to the laboratory in liquid nitrogen	Relative diagnostic sensitivity of samples transported in the different mediums
Dose House dust mite reduction improve symptom severity in patients with atopic dermatitis? ¹⁰	Patients with atopic dermatitis	House dust mite reduction	No house dust mite reduction	Patients symptom severity based on SCORAD or Leicester Sign Score
Is Mohs micrographic surgery more effective than wide local excision for treatment of dermatofibrosarcoma protuberans in reducing risk of local recurrence? ¹¹	Patients with dermatofibrosarcoma protruberans	Mohs micrographic surgery	Wide local excision	Rate of local recurrence at 3 years follow up
Does paternal exposure to azathioprine (AZA), methotrexate (MTX) or mycophenolic acid (MPA) during/before conception lead to adverse pregnancy and postnatal outcomes? ¹²	Male patients who have conceived a child whilst on immunosuppressive medication	Azathioprine	Methotrexate mycophenolic acid	pregnancy outcomes (risk of preterm birth, stillbirth, spontaneous abortion) fetal and neonatal outcomes (birth weight ,congenital malformations, morbidity)

Table 2: Examples of PICO questions from CATs published in the British Journal of Dermatology

6S pyramid level	Explanation	Examples of where the evidence can be found
Systems	Systems are databases that can integrate the best evidence based information with patients' individual circumstances	
Summaries	Summaries are up-to-date guidelines or online texts which use the best evidence based information related to a clinical topic	<p>General guidelines:</p> <ul style="list-style-type: none"> • Guidelines International Network (GIN) database • National Institute for Clinical Excellence (NICE) • Scottish Intercollegiate Guidelines Network (SIGN) • National Guideline Clearinghouse <p>Dermatology specific guidelines:</p> <ul style="list-style-type: none"> • European Dermatology Forum (EDF) • American Academy of Dermatology (AAD) guidelines • British Association of Dermatologists (BAD) • Nederlandse Vereniging voor Dermatologie en Venereologie (NVDV) databases.
Synopses of syntheses	Synopses of systematically reviewed evidence articles	Cochrane summaries
Syntheses	An article which systematically reviews the available evidence related to one clinical question	<p>Cochrane Database of Systematic Reviews</p> <p>Cochrane Skin Group (CSG) database</p> <p>Turning Research into Practice (TRIP) database</p> <p>Disease specific databases:</p> <ul style="list-style-type: none"> • GREAT database for atopic dermatitis (www.greatdatabase.org.uk) (SRs on eczema)
Synopses of single studies	Synopses and critical appraisals of robust original studies	<p>Turning Research into Practice (TRIP) database</p> <p>ACP journal club</p>
Single studies	The original research study investigating a particular clinical question e.g. a randomised controlled trial.	<p>Accessed through general databases:</p> <ul style="list-style-type: none"> • (EMBASE)MEDLINE i.e. via Pubmed or OVID (CENTRAL of the Cochrane Library (for RCTs <p>Disease specific databases:</p> <ul style="list-style-type: none"> • GREAT database for atopic dermatitis (www.greatdatabase.org.uk)

Table 3: Sources of evidence from different levels of the 6S pyramid