SUPPLEMENTARY INFORMATION / WEB APPENDIX

Conflicts of interest (COIs)

The following mandatory conditions and exclusions from the policy are highlighted:

a) Individuals with the following types of personal financial interests shall not take part in guideline development:
   - any shareholdings in the commercial sector held by the guideline development group (GDG) member or their family member
   - funds which include investments in the commercial sector that are held in a portfolio where the GDG member has the ability to instruct the fund manager as to the composition of the fund

b) GDG members with any ***specific*** interests during and 36 months prior to the commencement of the guideline work shall not chair a GDG meeting or be the first or last author of the published guideline.

c) A GDG should consist of a minimum of 51% of members with no ***specific*** personal or non-personal financial interests.

d) In formulating recommendations for guidelines the chair should consider whether members with COIs should take part in the proceedings – the actions required at GDG meetings are outlined in the table below:

Type of interest		**Action**
Personal financial	*Specific*	Declare and leave the meeting. In exceptional circumstances* the chair may rule that they can attend to answer specific questions
	Non-specific	Declare and participate unless, exceptionally, the chair rules otherwise
Personal non-financial	*Specific*	Declare and action is then at the discretion of the chair
	Non-specific	Declare and participate unless, exceptionally, the chair rules otherwise
Non-personal financial	*Specific*	Declare and participate unless, exceptionally, the chair rules otherwise
	Non-specific	Declare and participate unless, exceptionally, the chair rules otherwise

*For example, if the GDG member is the only expert in the country on a particular topic being discussed

e) All GDG members should stop accruing COIs during the development of the guideline so as to be able to participate fully in meetings. Any new COIs acquired during the guideline development should be declared.

The Clinical Standards Unit (CStU) will provide assistance to GDG members with their declaration of interests.

Meetings

Three meetings of the GDG will be offered at the British Association of Dermatologists (BAD) offices in Willan House, London:

1st: At commencement, going through the process involved, determining clinical questions, outcomes and structure of the guidelines, as well as dividing responsibilities for sections between GDG members.
2nd: Review of the evidence, formulation of recommendations and any discussion of drafts produced for the relevant sections.
3rd: Towards the end of the drafting process for pulling the document together.

Additional meetings may be appropriate, e.g. in reviewing the evidence and formulating recommendations where one meeting may be inadequate; discussing feedbacks received during the consultation period with the BAD membership and relevant stakeholders; or ratifying any amendments and actions.

All minutes will be prepared by the CStU and approved by the GDG lead before circulation.

GDG members' travel expenses will be reimbursed, with lunch and refreshments provided at the meetings as appropriate.

Selection of identified literature and retrieval of papers

The titles (only) of the pooled results will be divided equally by the number of GDG members (including relevant BAD members of staff in the CStU, if appropriate) who should delete any publications that are obviously irrelevant; publications with ambiguous titles should be retained at this stage. The shortlisted titles will then be returned to the Information Scientist.

The Information Scientist will collate the shortlists and produce a second list with titles and abstracts, dividing it equally for *paired* GDG members to carry out positive selections of publications they would like to read for further assessment. The systematic review protocol should be used to guide and remind the GDG of the inclusion and exclusion criteria. Any disagreements in the selection within each pair will be resolved by the GDG lead, with the shortlisted abstracts returned to the Information Scientist. During this second-round selection stage GDG members should also pigeon-hole their selections into the appropriate sections of the guideline.

It may be necessary to carry out a third selection round, e.g. in sections where there may be too many studies of lower quality shortlisted in the presence of those of higher quality. GDG members (including relevant BAD members of staff in the CStU, if appropriate) should review the papers to determine if they meet the criteria set in the systematic review protocol, especially in reporting the outcomes as agreed by the GDG patients.

GRADE-ing the quality of evidence

As the common P, I, C, O strands from included studies will need to be extracted, pooled and synthesised, the quality of evidence is assessed *per comparison* (intervention vs. comparator/control), *per outcome* and then *across all the studies* (i.e. for the *body of evidence*) which report that comparison and that outcome.

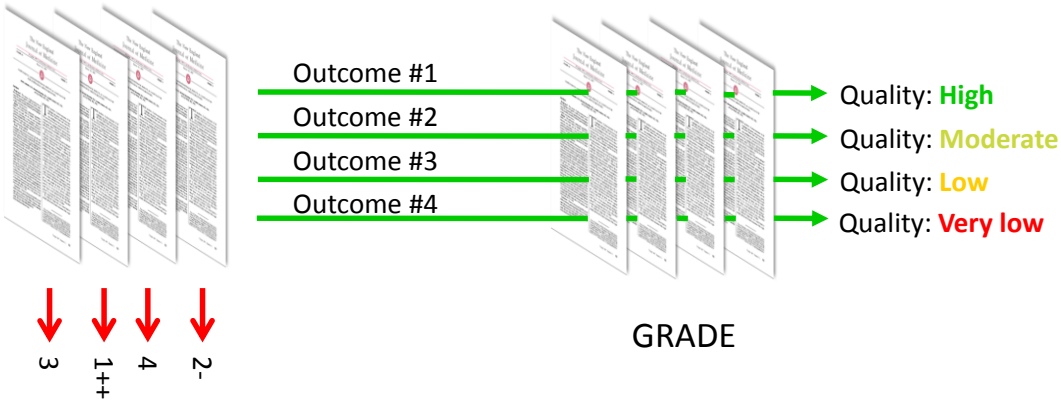

Figure S1. Comparison of the evidence grading approaches for four studies evaluating the same intervention and comparator/control.

Evidence based on RCTs will be graded initially as high quality, but confidence in the trials may be decreased due to five factors:

1. Study limitation/risk of bias, with the domains for this factor involving:
   a. selection bias (e.g. lack of allocation concealment or sequence generation)
   b. performance bias (e.g. lack of blinding)
   c. attrition bias (e.g. loss of participants – dropouts; non-responders; stopping early for benefit; protocol deviators and failure to adhere to an intention-to-treat (ITT) analysis)
   d. measurement bias (e.g. inaccuracy in the measurement instrument; bias in the expectations of study participants, carers or researchers)
   e. outcome reporting bias (e.g. selective reporting of outcomes or failure to report outcomes)

2. Inconsistency of the pooled results, where:
   a. widely differing estimates of the treatment effect across included studies would suggest true differences in the underlying treatment effect; variability in the results may arise from differences in the P (e.g. disease severities), I (e.g. treatment doses), C or O (e.g. time points at which treatment effects are measured), in which case subgroup analyses may be appropriate
   b. heterogeneity exists without plausible explanation

3. Indirectness of the evidence, where:
   a. indirect comparisons are made between studies with differing P, I, C or O, e.g. making comparisons of the magnitude of effect of drug A vs. drug B when the trials investigated drug A vs. placebo and drug B vs. placebo, or combining data from inappropriate patient population

4. Imprecision of the pooled results:
   a. studies with relatively few patients and few events, often denoted by a wide confidence interval (CI)

5. Reporting/publication bias:

a. failure to report studies or funding issues exist

The GDG should be vigilant of these potential negative factors, i.e. those that would reduce their confidence in the quality of the evidence. For each of the factors the GDG will need to determine 1) if there is a problem, 2) whether the problem matters and 3) if so, by how much does it matter.

For each of the five factors above, a quality rating will be given across all the studies for each outcome in each comparison as being:

- with no serious problem
- with a serious problem
- with a very serious problem

An overall quality rating of the evidence will then be obtained for each outcome. Problems with the ratings for the five factors will lead to downgrading of the evidence. Using RCTs as an example, for each outcome:

- 0 if there are no problems with any of the five factors (**high**)
- -1 if there is a problem with one factor (**high** → **moderate**)
- -2 if there are problems with two factors (**high** → **low**)
- -3 if there are problems with three factors (**high** → **very low**)

Observational studies will start with "low" quality ratings due to inherent selection bias. However, *upgrading* may be possible in unusual circumstances, but only if the evidence has not been downgraded already, due to study limitation/risk of bias, inconsistency and imprecision. Factors that may upgrade observational studies are in cases where:

- the magnitude of the treatment effect is large (RR>2 or RR<0.5) or very large (RR>5 or RR<0.2)
- there is evidence of a dose-response relationship
- all plausible confounders or biases behave in the opposite direction to what might be expected

Therefore, for example, observational studies that have not been downgraded, e.g. due to risk of bias, may be upgraded to:

- +1 when RR>2 or RR<0.5 (**low** → **moderate**)
- +2 when RR>5 or RR<0.2 (**low** → **high**)

For decision-making, the GDG will need to establish:

1. outcomes that are critical and important to patients
2. the magnitude of the treatment effect as absolute differences *per outcome*
3. the overall quality of evidence *per outcome*

The GDG will need to look at the clinical importance of each outcome, and not the statistical significance, i.e. consideration of the absolute changes (magnitude of treatment effect) and direction (benefit or harm). In terms of judging clinical importance, which is only based on the magnitude of treatment effect, GDGs will need to decide if the effect estimate represents a clinically important difference; if the difference is clinically important, then they will need to indicate the direction.

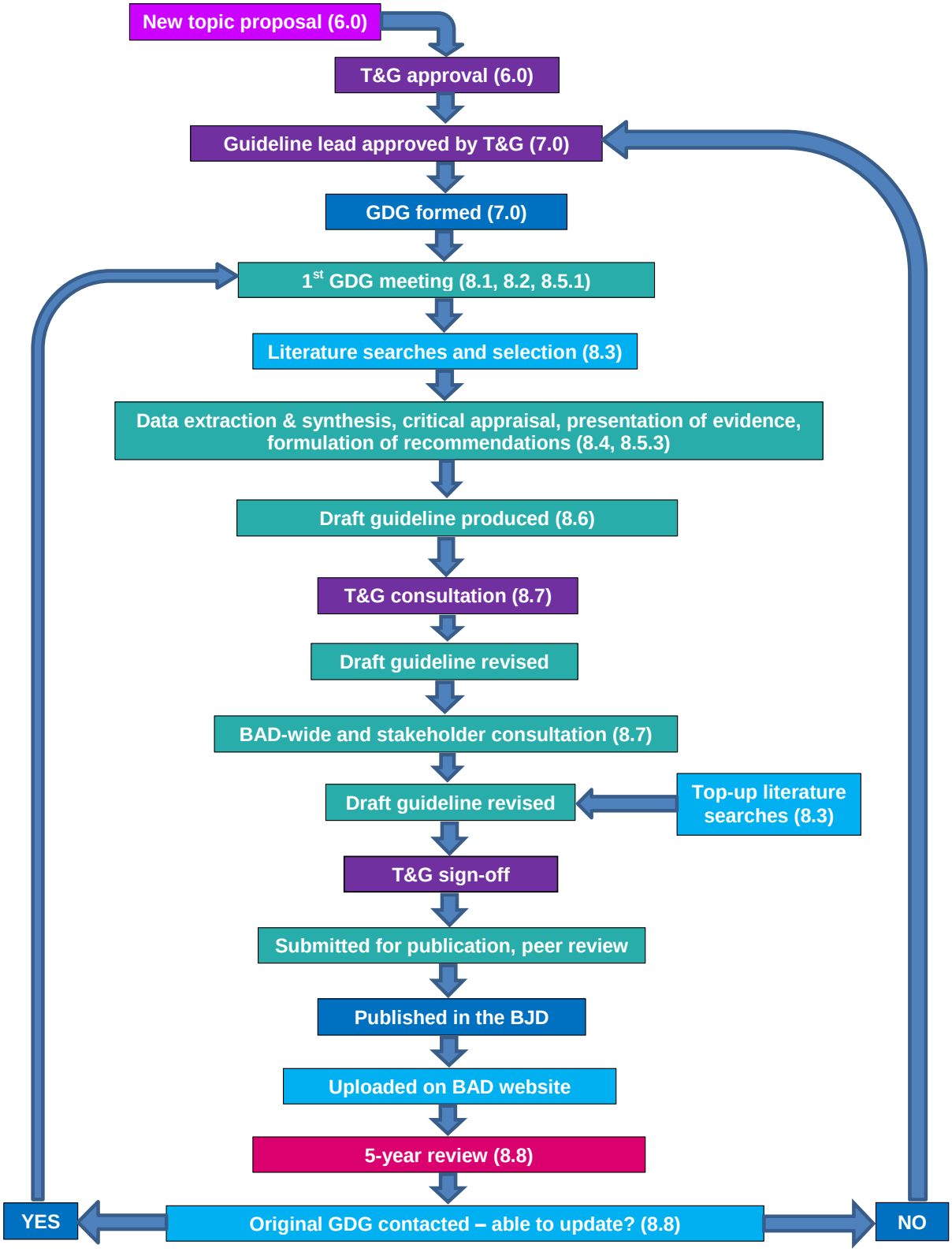

Figure S2. Flow diagram covering the major stages of a BAD guideline development (reference to the relevant sections in the main article in brackets)