Less is more - the minimal dataset in reviews for guidelines

Maggie Westby
Clinical effectiveness lead
National Clinical Guideline Centre, UK
(NICE Collaborating Centre)
The Need

■ NICE clinical guidelines:
  ● developed by a guideline development group (GDG) - clinicians, patient representatives and a technical team (systematic reviewers, information specialists and health economists).
  ● produced in a relatively short space of time – 26 months between publication of guidelines in the same ‘slot’:
    => 6-8 months of scoping, 6-8 months validation, leaving around 12 months development time.

■ Typically, about 10-15 systematic reviews carried out in that 12 month development time (c.f. Cochrane reviews 6-12 months minimum each!)
Quality assessment

- 8 week public consultation period at the end of development
  - quality assessment of evidence reviews by NICE
  - peer review and comments from external stakeholders.
  - We must answer every comment and sometimes we have to revise reviews in the light of comments – time consuming and can be stressful
So...

- We need evidence reviews to be done both quickly and well.
Possible solutions to quickly and well

- More reviewers
- Fewer review questions
- Reviewers work harder / longer hours
- Reviews are done more efficiently
How do we do reviews efficiently…

…and still maintain the quality?

- Two main ideas
  - Cut down the amount of data abstracted
  - Abstract the data in a different way

- 3 different approaches:
  - ‘Simple’ selective data abstraction
  - Use a relational database
  - Be smart in the ways we do systematic reviews
Cut down the number of words

Our BlueMotion range combines lighter materials, enhanced aerodynamics, economical engines and tyres that create less friction, which saves you fuel and can reduce your tax, which means you will have more money.

Another example of Volkswagen efficiency.
Use a relational database

- For the user, linked tables containing pre-specified items appear as dropdown menus on user interface screens

- Advantages:
  - dropdown menus - quick and avoids typing errors
  - facilitate sorting of data
  - Outputs available in variety of formats
Be smart in how we do systematic reviews

- Good systematic review starts with a well defined review question

=> Protocol describing:

- PICO (population-intervention-comparison-outcome) inclusion criteria

- Combining and splitting (stratification) that will be done in a meta-analysis

- Subgroups and heterogeneity
Stratification at start

- Some GDGs split data into many strata (subgroups)
  - Negates advantages of meta-analysis and risks producing chance effects
  -Reviewer has much more work to do
  -Recommendations based on limited and potentially misleading data.
- GDGs should have a very good clinical reasons for separating into strata at the outset.
Subgroup analyses for heterogeneity

- Some reviewers /GDGs put in as many subgroup analyses as possible to cover all bases
- Many of these will not have good clinical reasons + risk of chance effects …
- Reviewer has to abstract data for each subgroup included

=> Often unnecessary
Minimal dataset as applied to subgroups

Therefore we proposed that:

1) only a few, highly relevant subgroup analyses should be specified, and

2) if a patient characteristic is not used for subgroup analyses, then no need to abstract the data.

i.e. Strict “minimal dataset” - only abstract data on what you need to do the review well.

Disadvantage is that may get it wrong and then what do we do?
Benefits

- Reviewer has to think carefully about the review protocol, so (as well as being quicker):
  - analysis and interpretation easier
  - GDG clinicians have confidence
  - no need for re-work at consultation

- Our aim of ‘becoming more efficient but still maintaining the quality’ → ‘efficient with improved quality’

- Decidedly ‘less is more’.
So far so good…

We proposed a database which:

- Was protocol based
  - which can only be changed exceptionally
- Had a strict minimal dataset (upper limit)
- Used dropdown menus where possible
- Cut back on words in text boxes
Projects on determining the minimal dataset

Study 1:
- Assessed a series of past reviews carried out in our merged centre
- Classified abstracted data into themes:
  - Items common to all guidelines
  - Items only for some guidelines
  - Items abstracted for guidelines but not used further in analysis/discussion
Project 2 – protocol for a Cochrane review

- 12 systematic reviewers from NCGC.
- Focus group work with author of a Cochrane review (MW) – objective to determine the protocol and the minimal dataset in a medical specialty unfamiliar to all of them.
  - MW presented background for subject area and review.
  - Then reviewers asked questions to help them determine the PICO items and any subgroups analyses they thought appropriate
  - Questions later classified as general or specific to the review
## Results for both studies

<table>
<thead>
<tr>
<th>Type of item</th>
<th>Data abstracted that was common to all reviews</th>
<th>Data abstracted not common to all reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study details</td>
<td>• Study design (e.g. RCT), country, funding, inclusion criteria, exclusion criteria</td>
<td>Unit of randomisation; setting</td>
</tr>
<tr>
<td></td>
<td>• Treatment duration and follow up time</td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>• Age, gender, ethnicity,</td>
<td>guideline- or review-specific items;</td>
</tr>
<tr>
<td></td>
<td>• Previous treatments/ line of therapy</td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>• description of what it is, dose/volume of intervention, number of participants receiving the intervention</td>
<td>concurrent medication, route of administration, guideline- or review-specific items</td>
</tr>
<tr>
<td></td>
<td>• Duration of treatment and/or follow up and distinguishing between them</td>
<td>Adjunctive treatments</td>
</tr>
<tr>
<td></td>
<td>• Treatment schedules</td>
<td></td>
</tr>
<tr>
<td>Comparisons</td>
<td>comparator, comparisons</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>what it was, results</td>
<td>Time outcome measured</td>
</tr>
<tr>
<td></td>
<td>Techniques for measuring outcomes</td>
<td></td>
</tr>
<tr>
<td>Quality assessment</td>
<td>Allocation concealment, blinding, attrition</td>
<td>The others!</td>
</tr>
</tbody>
</table>
Conclusions about generating dataset

We identified 3 groups of data items which we translated into our database for reviewing:

- common to all guidelines - dropdown menus which can be standardised (e.g. quality assessment, study design)
- common to all guidelines, but content varied by guideline – as text boxes (e.g. ethnicity)
- specific to a guideline, subgroups - dropdown menus generated by reviewers
Minimal dataset does it work?

We then planned to test this in study 2:

- produce two databases, one with a strict minimal dataset and one with blanket coverage
- Compare speed of data abstraction and effectiveness for the 2 processes using matched reviewers.
But there were barriers

- Reviewers were not confident that they could define a minimal dataset on the basis of a small amount of background information and their discussions with the Cochrane review author, when the subject matter was unfamiliar to them.

- Deciding what to leave out proved difficult
  => no substantive differences between a minimal dataset and a blanket dataset
More barriers

- Clinicians thought that a minimal dataset with strict limits might not be workable in practice
  - GDG members don’t understand reviewing at the beginning
  - Reviewers don’t know enough about the subject early on (mainly not clinicians)

- Concluded there is a need for some flexibility in the minimal dataset (but not too much)
Compromise, as implemented in the database

- Reviewers explain the reasons for choosing the subgroups, but no upper limit to the number of subgroups.

- Multiple stratification is discouraged, by requiring more than one split to be written out in full, e.g. people with diabetes and with dementia; people with diabetes but no dementia; people without diabetes but no dementia, people with neither diabetes nor dementia.

- Additional subgroups can be added later (but special procedure)

- Rescue text boxes
Getting information from GDGs

- Study 2 showed that it was possible to use a directed approach to obtain information from clinicians.
- Questions are used in the database to guide reviewing, for example, asking ‘are crossover studies appropriate for this review question’.
- Getting information is clearly an important topic and essential if one is to have a protocol-led guideline.
- Important for reviewers to work with clinicians as equal partners.
Future work

- About to test database with reviewers
- Further work on methods for getting information from clinicians (possibly with qualitative researchers)