DECIDE: medical testing

The DECIDE project included six research Work Packages (WPs), the first five of which aimed to develop and evaluate strategies for presenting evidence-based recommendations in guidelines to different types of user:

1. Health professionals.
2. Policymakers and managers.
3. General public.
4. **Users of diagnostic tests [covered by this summary].**
5. People developing health system policies.

The 6th Work Package was a toolkit that packaged much of the work coming from the first five Work Packages together. One of the key results of DECIDE was to deliver information in layers, most important first. So, in that spirit, the key findings of the whole DECIDE project are summarised in Figure 1. If you read no more, look at least at Figure 1.

Key findings:

- Guideline users – health professionals, patients and policymakers – want information delivered to them in layers, most important first.
- Guideline producers value structure when working through evidence to make recommendations and decisions.
- Numerical summaries of research findings can be understood by diverse audiences, including the public, but it is best if those summaries allowed users to interact with them so that they can choose the level of detail they require.
- Health professionals and their patients want materials that can be used in consultations to support their discussions.
- Guideline information about medical testing has to move beyond accuracy and precision and start talking about the effect on important patient outcomes.

Key tools:

- The Evidence to Decision framework to support guideline producers make evidence informed decisions.
- The interactive Summary of Findings tables to support interactive presentations of research findings to diverse types of user.
- The DECIDE/G-I-N public toolkit chapter for guideline producers on how to produce patient versions of guidelines.
- There are many ways information can be presented to users but we have not found a 'magic bullet' that always works for all users, especially members of the public. Guideline producers would be wise to do at least some testing of their materials with potential future users.
- The GRADEPro guideline development tool to package the bulk of DECIDE's work and to support guideline producers through the whole guideline process.

***Figure 1**: Key DECIDE findings and tools*

Presenting evidence-based recommendations about diagnostic tests

We planned our work in three phases. In a first phase we aimed at identifying current strategies to develop and communicate evidence-based recommendations about diagnostic tests. In a second phase we wanted to enhance these strategies and/or fill the identified gaps. In a third step, we moved to user testing for refining the strategies. Because the development of evidence-based

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 258583

recommendations of diagnostic tests is less well understood compared to recommendations for therapeutic interventions, we spent some time fine-tuning and elaborating the methods for arriving at diagnostic recommendations, to achieve effective presentation strategies.

Two systematic reviews were performed. In the first we compared grading systems for medical tests on how they use evidence in guideline development. Twelve grading systems were included in the review. They varied in the degree to which methodological and process characteristics were addressed. Five systems for grading evidence about medical tests in guideline development addressed - to differing degrees of explicitness - the need for and appraisal of different bodies of evidence, the linking of such evidence, and the translation into recommendations. At present, no one system addressed the full complexity of gathering, assessing and linking different bodies of evidence for making recommendations about tests. The review was published in 2013 in Implementation Science (http://www.implementationscience.com/content/8/1/78).

The second systematic review (submitted for publication) was aimed at methods used by organisations developing recommendations about diagnostic tests. We found 44 tools and modifications therefor to assess the quality of evidence supporting diagnostic tests and testing strategies. These tools used inconsistent terminology and the criteria for moving from evidence to recommendations were found incomplete for most guideline development frameworks that were evaluated.

One of the better known and developed grading systems is GRADE. Originally developed for evaluating and making recommendations around interventions, GRADE is now working towards a system that is also applicable for diagnostics. As part of our work in understanding current strategies in phase 1, we applied the GRADE for Diagnostics approach on three Cochrane diagnostic test accuracy reviews to further enhance understanding around the ‘‘practical’’ application of the approach and identify real-life challenges and considerations a user of this approach may encounter. By doing so, we aimed to provide suggestions on how the GRADE for diagnostic approach may be enhanced. This work was published in 2014 in the Journal of Clinical Epidemiology (http://www.sciencedirect.com/science/article/pii/S0895435614000444).

We then conducted two qualitative interview studies (submitted for publication) involving guideline developers and experts with a variety of backgrounds, formal training and experience in methods of evaluating evidence and making recommendations about diagnostic tests. In these we found that diagnostic test accuracy – based on comparisons between test results and the gold or reference standard – was the factor most commonly considered by organisations when formulating recommendations. The majority of experts pointed out that accuracy alone is not sufficient and that recommendations based on accuracy only may be misleading. From the analysis of the interviews we learned that the challenges guideline developers currently face are interlinked; these challenges can be found in methodological issues (e.g. how to link different types of evidence), resource limitations (e.g. the limited time and money for developing a guideline) and a lack of awareness regarding using patient important outcomes, instead of diagnostic accuracy, as the criterion for making recommendations.

The central and recurrent theme of the DECIDE approach with regard to diagnostic testing is the need to widen the focus of test evaluations, from relying on diagnostic test accuracy only to using the effect on patient important outcomes as the decisive factor. Similar to the development and presentation of recommendations about treatment interventions, the central question when building recommendations about diagnostic tests is the existence and magnitude of health benefits for the patients in whom testing is considered.

When evaluating potential benefits one should not only focus on the test, but also consider the clinical management that follows from the test result. Since studies that present direct effects from testing on health outcomes our rare, any methodology for evaluating tests or for building recommendations should consider using, assessing and possibly linking different types of evidence. This may include evidence about test performance, but also evidence about the effectiveness of downstream actions, guided by the test results.

This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no 258583

This complex process can be facilitated by adopting the test-treatment pathway approach and a diagnostic Evidence to Decision framework. The pathway starts with how and where patients disease present themselves, and runs via testing to management decisions and ultimately to patient outcomes. The pathway approach can be used to clarify differences between alternative and existing testing strategies. It can also be used to describe how the introduction of a new test change current management pathways.

At present, some guideline producers mention the identification of such pathways in their guidance. It can be an element of a scoping exercise, or a part of key question development. Explicit instructions on how to map the pathway are usually missing, unfortunately.

We used the Patient–Index test–Comparator–Outcome (PICO) elements, which are well known from questions about the effectiveness of therapeutic interventions. From these we developed a structured set of trigger questions that can be used as an initial starting point to identify the test-treatment pathway for a medical test. PICO is already a common feature in evidence-based medicine methodology. The Cochrane handbook for systematic reviews of diagnostic test accuracy studies also recommends the PICO system to define the pathway. PICO is also used in the GRADE approach for diagnosis.

The basic framework for starting to build test-treatment pathways and the corresponding PICO elements in it are shown in Figure 2.

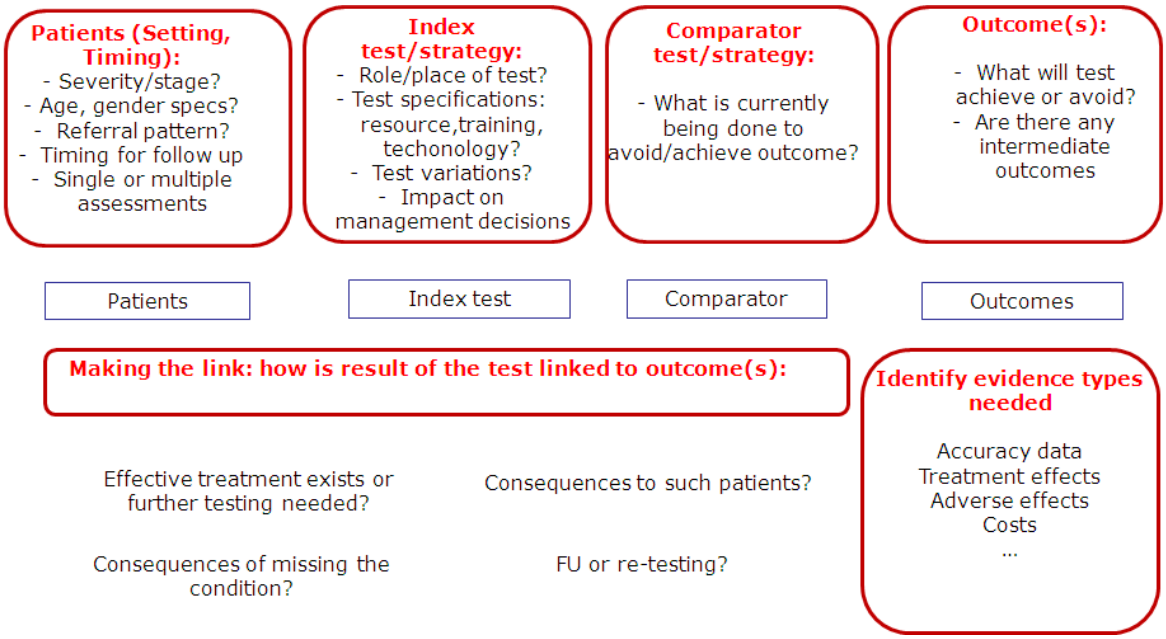

Figure 2: *Framework built on the PICO elements for identifying the test-treatment pathway*.

In a number of workshops and other applications we have done user testing of the PICO approach for identifying the test-treatment pathway. This led to a further refinement of the triggering questions to clarify the respective elements in the pathways. We also found out that graphical tools, borrowed from decision trees as used in decision analysis, help to clarify the structure and the possible alternatives in test-treatment pathways. In these tools, time runs from left to right, with patient presentation starting left and possible outcomes towards the right end. Branches in the trees indicate alternative lines of actions, guided by the results from testing.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 258583