The BEME Collaboration

Best Evidence Medical Education

Proposal for Systematic Review

A Systematic Review of the Psychometric and Edumetric Properties of Assessment Tools for Communication and Consultation in Medicine for Undergraduates and Postgraduates

07.09.09
BEME review title:
A Systematic Review of the Psychometric and Edumetric Properties of Assessment Tools for Communication and Consultation in Medicine for Undergraduates and Postgraduates

Introduction

We propose to develop a suite of systematic reviews of assessment tools for the medical consultation our ultimate goals being to.

- Enable comparison in terms of their psychometric reliability and validity,
- Support an expansion of the classical concept of validity by incorporating an assessment of edumetric approaches to evaluation
- Evaluate the suitability of assessment tools for their selective and formative qualities, as the presence of the latter quality may enhance the potential of lifelong learning when the tool is being used.
- Distinguish between assessment tools directed at essential generic communication skills and those directed at wider consultation skills (i.e. complex skills applied to a specific professional context).
- Distinguish between applicability for an undergraduate or postgraduate context.

The relevant domains to be considered are mapped out below:

Approach to Systematic Review of Assessment Tools for Medical Consultation

- Undergraduates
  - Communication skills
  - Consultation skills
  - Psychometrics
  - Edumetrics

- Postgraduates
  - Communication skills
  - Consultation skills
  - Psychometrics
  - Edumetrics

Medical Doctors.
Communication skills are bound up with consultation skills and in many ways can be considered a sub-set. The purpose of the diagram is to clarify our approach to the review and hence each relevant element under consideration has been represented as a box. We accept that there may be less rigid boundaries and some overlap between the domains to be covered ultimately in the suite of reviews, however we will be able to map this out further as we proceed with the actual review and we will re-visit these domains and refine our approach as we proceed.

As we will need to develop a robust methodological approach to support these systematic reviews we have decided to start with a systematic review focussing solely on the assessment tools for communication skills in the medical consultation in the context of undergraduates. The reason for this being that we hypothesize that in the undergraduate context the focus is more on the training and assessment of generic communication skills. The remainder of this proposal will therefore specifically focus on this first review alone. We will then build on our experience with this particular narrow focus to address the other areas outlined above which are relevant to educators working in this field. So the first of this suite of systematic reviews is the area coloured yellow below.

Names and Affiliations of Key Topic Review Group (TRG) Members:

We have assembled a review group from 3 institutions across Europe to ensure we develop a broad perspective, and a multi-lingual one which will enable us to search relevant non-English language literature effectively. We have included those who
teach and research communication and consultation skills: doctors and psychologists. Further these institutions all use different assessment tools and the review team comprises of academics with diverse but relevant skills in terms of education and research.

**Reviewers:**

Teresa Pawlikowska BSc (Hons) MB BS, M.Sc. (Hons), MRCP, DRCOG.
Dr Pawlikowska is Associate Clinical Professor at Warwick Medical School where her special interest is input into both undergraduate and postgraduate communication and consultation skills education. In her previous post as Head of the International Unit in the Department of General Practice at University College London she developed an interest in international comparisons as she was responsible for the roll-out of general practice in former Eastern Europe under several EU Phare and World Bank programmes and input into countries as diverse as Albania and Japan. This included curriculum development and support, establishing and mentoring new university departments of academic general practice. She devised and taught on a wide variety of these courses, with a particular focus on consultation skills training and delivery. Her special interest is in communication and consultation skills both teaching and research, and student learning. Her PhD, at the University of Maastricht (prom. Prof Cees van der Vleuten, Jan van Dalen) has explored the use of patient enablement as an outcome measure for consultation quality using a multi-method approach, and she has published on the use of the Patient Enablement Instrument in the context of educational reform programmes. She has over fifteen years experience of the development of educational training programmes at both undergraduate and post-graduate levels and is a member of the Academy of Medical Educators in the UK. She chairs the Education Committee at The European General Practice Research Network, as well as being an elected Board member.

Michael Davies MB ChB DRCOG FRCGP.
Mike Davies is currently a part-time General Medical Practitioner working in a mixed suburban practice in Leicester, England. He has worked there for over 20 years and developed the educational side of the practice such that at any one time there are three GP registrars and a medical student in training at the medical centre. For many years he has been involved in undergraduate medical education at Leicester University Medical School in various roles including as a part time clinical tutor specialising in consultation and communication training and assessment. He is a postgraduate trainer in medical education and a Programme Director for the East Midlands Healthcare Workforce Deanery with responsibility for the training and assessment of approximately sixty GP registrars. He has recently been appointed as an Associate Postgraduate Dean for East Midlands Healthcare Workforce Deanery and he is currently working on a dissertation to examine the reliability and validity of communication assessment tools as the final part of a master’s degree in medical education (MMEd) at Warwick University supervised by Teresa Pawlikowska.

Angelique Timmerman, PhD
Angelique Timmerman works as a communication skills trainer and researcher specialising in the area of education in doctor-patient communication at the Department of Vocational Training in General Practice, University of Maastricht. Her
current research area is in the development of a competence based national curriculum of doctor-patient communication for all institutes for vocational training in general practice in the Netherlands. She has experience with conducting systematic reviews e.g. in 2007 she conducted a systematic review into the psychometric qualities of questionnaires for the assessment of functional health status in otitis media. (*Clinical Otolaryngology.* 2007; 32, 429-439) and is currently completing the write up of a systematic review with a speech pathologist about the psychometric qualities of health-related quality of life questionnaires for patients with dysphagia.

Paul Ram, MD, PhD

Paul Ram is general practitioner and director of the department of vocational training of Maastricht University. His academic work is directed at teaching and research on performance assessment. In 1998 he defended his thesis “Comprehensive assessment of general practitioners” successfully. He is an expert in video-assessment of medical performance and doctor-patient communication, assessment on medical knowledge and practice-organisation. He studied the relation between the different levels of assessment such as the predictive value of knowledge tests, OSCE and assessment of organization and collaboration for the performance of doctors in daily consultations. As co-promoter he supervises different projects concerning assessment of doctor-patient communication. He is (co)author of many articles in both national and international peer-reviewed journals. He has combined these tasks with the presentation of medical education courses on radio and television (TELEAC). He worked part-time as a member of staff at the Dutch College for General Practitioners (NHG) in Utrecht, principally concerning implementation of comprehensive assessment in quality improvement activities in general practice. Currently he is part-time project manager assessment (PMA) for all the (eight) departments of vocational training in the Netherlands, joined in Huisartsopleiding Nederland, in Utrecht as well.

Paul Van Royen, MD, PhD, is Professor at the University of Antwerp, Belgium and head of the Department of Primary and Interdisciplinary Care. His academic work is directed at teaching and research in primary care. He is co-ordinator of several teaching modules and programs, at undergraduate levels as well as graduate and postgraduate levels. At his department, he has developed a research centre with a strong expertise in qualitative research, systematic reviews and several research themes, including research on medical education (patient-centredness, medical decision making, Skillslab). He is/was involved as co-coordinator/researcher in 9 systematic reviews, including one Cochrane review and one medical education review on clerkships. He is author of more than 150 articles in peer-reviewed journals and reviewer of different scientific journals. At this moment he is President of EGPRN (the European General Practice Research Network), the WONCA network for Primary Care research in Europe. Since 1996 he is co-ordinator of the Flemish Clinical Guidelines project.

Katrien Bombeke (GP) is currently working as a researcher and teacher at the Department of Primary and Interdisciplinary Care, University of Antwerp. Since 2005, she has been working on a PhD-project on the topic of patient-centredness in medical education (promoter: Prof. Dr. Paul Van Royen). One of the core research questions of her thesis concerns the development and validation of a new measurement scale for patient-centredness in medical students. Being a
communication skills trainer too, she is familiar with this proposal’s topic not only from a scientific viewpoint (reliability and validity of measurement scales), but also from daily educational practice (Calgary-Cambridge Observation Guide, OSCE, ALOBA).

Rationale for BEME review protocol

The importance of good communication in medical practice has been widely recognised by professional bodies in Europe [1] and North America [2]. Despite the knowledge that physician-patient communication is pivotal in determining positive outcomes from consultations, the commonest mistakes reported to the UK Health Service Ombudsman are still ones of communication [3]. Effective communication has been shown to increase diagnostic accuracy [4] and studies both in Europe and in the US have demonstrated improved patient outcomes [5] [6] [7-9] [10](e.g. patient satisfaction, enablement, compliance).

There is good quality evidence to support the assertion that the consultation process is important and having the skills to consult effectively with patients is an essential prerequisite for competent medical practice. In its statement “Good Medical Practice” the UK General Medical Council (GMC) has underlined the importance of communication in commenting specifically on the necessity for graduates from UK medical schools to be able to communicate effectively with patients [1]. Increasing awareness of the importance of communication skills extends to the post-graduate sector. In the UK the examination for membership of the Royal College of General Practitioners formerly included a video assessment of consultation performance and more recently this has changed to a form of OSCE or “Clinical Skills Assessment” (nMRCGP 2009 [11]. Other membership examinations including the Royal College of Physicians also now include an assessment of communication skills as part of the process (MRCPUK 2009 [12]) and also the Royal College of Psychiatrists [13].

In Europe there is a similar emphasis on good communication as an integral part of good practice. All Flemish Universities have integrated communication skills training across the continuum of medical education. For example, at the University of Antwerp, students’ communication skills are assessed by OSCE’s, written multiple choice and open question exams, and by portfolio evaluation. Communication skills also are an important part of the Flemish entrance examination students need to pass before starting medical school, and of the Flemish end examination for Master in General Practice (OSCE). In the Netherlands there is a new ‘Raamplan artsopleiding’ (2009) [14] which formulates the competencies for the 7 professional roles of a competent doctor, resembling the model of CanMEDS-2005. The physician role of communicator is included in this profile. The development of an efficient working alliance with patients, families and co-workers in health care is a prerequisite for appropriate application of communication skills in the consultation. In the postgraduate medical training for general practitioner, competencies and final qualification requirements have been described in the ‘Competentieprofiel en eindtermen van de huisarts’ (2009) [15]. From 2001, the assessment of medical communication skills by using videotaped consultations, rated with the Maas-Global scoring list by trained observers, [16] has been implemented in all 8 schools for vocational training in the Netherlands.
In parallel with the growth of communication skills models and teaching a number of assessment tools have been developed [17] [16] all over the world (MAAS –Global [16], Calgary-Cambridge [18], Pendleton et al [19], Four Habits [20], SEGUE [21]) to name a few that are in common use. There is a great deal of congruency in underlying concepts between them although parameters may differ or overlap. Consequently there is over-arching agreement on important elements of what constitutes good communication in medical consultation [22].

These assessment tools are important as they enable us to discriminate between medical students and doctors who have those skills and those that do not in terms of classical validity and reliability criteria. However, medical training qualifies for a profession, and in this respect competence-based education has become increasingly important over the past years using complex clinical problems as a resource [23, 24]. The assessment of competences requires other assessment criteria, besides classical psychometrics because consistency is very difficult in repeated measurements for these real life problems. New assessment criteria e.g. edumetric properties not only focus on the validity of the tasks or tool, but also on the validity or fairness of the judgment, generalisability or authenticity of the assessment to other tasks measuring the same competence and consequential validity. An important extension then, is the evaluation of the consequences of assessment for educational purposes and the learning process of the student (edumetric properties [25, 26]).

With this proliferation of interest and activity there are now a large number of assessment tools currently in use. They have evolved in different ways and in different educational (e.g. undergraduate and/or postgraduate) contexts. They have been adapted over time as both education and societal needs have changed, and also to ensure local relevance. This complex process impacts on the reliability and validity of tools used which, as they have been in routine use often for long periods may have gravitated significantly from their original provenance, e.g. being used for different medical specialties in postgraduate training, while these professions may require different communication skills to be used in their practice. This review will highlight not only the similarities in approach between the assessment tools, but also the differences and in doing so will seek to optimise the practical application of these tools due to the comparison of their evaluative properties and signpost areas for future development.

**Approach**

We have taken a pragmatic approach to this multi-layered area and propose a suite of systematic reviews (see diagram). All reviews will cover the psychometric and edumetric aspects of assessment tools as we believe these properties to be important from an educational perspective: competency based education is directed at life long learning, while regular education assesses the achieved training level at one point in time. For the evolution of medical education edumetrics is the most important, due to the integration of entrusted professional activities into competency based education.

Undergraduates and postgraduates, although linked through life long learning, have differing training needs and function in different contexts. We will focus therefore the
reviews on assessment tools used at either undergraduate or postgraduate level. Although competencies may be intertwined in the medical consultation there is a need to distinguish communication skills from wider consultation skills to be employed.

We therefore propose a staged approach with a suite of systematic reviews all concerned with the psychometric and edumetric properties of assessment tools used for the medical consultation. Namely those:

1. used for at least communication skills at undergraduate level
2. used for at least communication skills at postgraduate level
3. used for consultation skills at undergraduate level
4. used for consultation skills at postgraduate level

We will seek to relate our findings to Miller’s pyramid, not at the level of passive knowledge but at the level of “shows” and “does” (the former may be more relevant for undergraduates, the latter for postgraduates). In line with this, we will endeavour to anchor our findings in terms of case-specificity and generalisability [27] (I completed this reference, see list) and where appropriate relate them to contemporary frameworks defining the profile of the medical professional, e.g. the increasingly popular Can MEDS core competencies. We expect the outcome not only to provide a comprehensive systematic review of the area but also to define important gaps, e.g. regarding necessary assessment themes, the quality of assessment tools and the direction of further development to ensure the fit between educational goals and later clinical practice.

Definitions

Psychometric assessment
If the primary purpose of a test is to measure individual differences over time or the fulfilment of licensing criteria at one point in time compared against a standardised norm statistically referred from scores in a reference group, then the test should be primarily evaluated using classical psychometric principles. The focus is ‘receiving information, memorising and reproducing it’. In tests this translates to the level of knowledge present and how to apply it, in terms of Miller’s pyramid the levels ‘knows’ and ‘knows how’, which is only part of the necessary development in competency-based education programmes. However, a test may be evaluated from an edumetric standpoint as well, when included in a larger range of tools in the context of an assessment of developing complex competences in the context of life long learning [28].

Edumetric assessment
Edumetric assessments are designed to measure the growth of an individual’s knowledge skills and abilities and support generalisation about a domain of tasks. Aspects such as professional and problem solving skills and being able to function in authentic, i.e., real-life contexts are important from an edumetric perspective. Edumetric validity is not dependent on variances in test scores, but on the formative role of assessment, stimulating the integration of instruction and assessment of complex or higher-order meta cognitive skills, including attitude, knowledge and skills. Assessment then may serve an important role in the constructing rich and
authentic learning environments, enabling the development of reflective practitioners [18][29, 30].

Reliability and validity:
An operational consensus will be reached by the review team on the definition of reliability and validity based on current work (MD Masters in Medical Education, Supervisor TP, Warwick University), the team’s experience[31] and current studies, such as Baartman and colleagues developing quality criteria for competence assessment programmes [19]. The intention is to assess issues such as reliability, generalisability and comparability of performance assessments. These issues need reconciliation when using assessment tools from an edumetric perspective in competency-based educational settings: an appropriate set of criteria and standards to simultaneously support the validity of an assessment based interpretation and of its impact on the education system [32-34].

Population:
As described above in the first instance we will focus on medical under-graduates, and then progress to post-graduates.

Outcomes:
- A suite of Best Evidence in Medical Education systematic reviews of the current state of available assessment tools for the medical consultation as detailed above.
- An assessment of both the psychometric and the edumetric quality of these tools for different educational purposes.
- Pragmatic reference guide for medical professions to increase awareness of derivation of tools and so assist in the choice of the appropriate assessment for new educational contexts and initiatives.
- Identification of areas which need further research or updating, e.g. development areas for assessment tools.
- Workshop is available to disseminate the process of systematic review on assessment tools for communication skills in the medical consultation elaborated in this project.

Process:
We have initial approval for funding which will enable us to meet at one of the universities involved at crucial stages of this review: we met on 9/4/10 initially to debate, discuss and agree a framework with scoring criteria, in- and exclusion criteria and procedure for enacting this systematic review. We will build on our experience of critical appraisal, education and research in communication and consultation skills. We will carry out an initial literature search and appraisal individually using the agreed framework (modified BEME template) on an agreed selection of papers and meet after this first pass (see Chart) to exchange experience and refine our criteria and methodology as required. We will then perform the final literature search, allocate the final selection of papers for appraisal using the agreed final template. We will then meet after this process is complete (see Chart) to agree on an approach to synthesising our work and also to agree on a write up procedure. We will discuss, debate and refine the final review at another meeting and agree the write up Procedure. The final review
will then be written up and agreed by all participants by email (see Chart) and submitted to BEME.

Search strategy:

Searches will be conducted by our clinical librarian together with the reviewers. Building on our experience we will target several databases: ERIC, PubMed, Web of Science, Web of Knowledge, EMBASE, PsycInfo, Dissertation and Sociological Abstracts, CINAHL, Topics in Medical Education (TIMELIT), British Education Index and the Cochrane Library. In these databases we will search on relevant topics in the field, such as ‘education’, ‘communication’ and ‘training’ and then combine them to trace papers describing the psychometric and/or edumetric qualities of relevant assessment tools from January 1960- January 2010. We will also define MeSH terms for each focussed review, and document any changes needed for each different database.

We will hand search relevant journals (Academic Medicine, Medical Education, Medical Teacher, Patient Education and Counseling, etc.).

The reference sections of relevant articles found by using these search strategies will be scrutinised for further relevant publications.

When we have developed a list of assessment tools we will then conduct a search on them by name to ensure that no relevant material pertaining to them has been missed.

We will also include internet and library searches to try to capture books which may have been written concerning this area, but which may be difficult to find listed in the above databases.

We will ensure we are not limited to English language articles as the research team also have a working knowledge of European languages to the required level.

We will also build on our experience with searching the literature for a related, but more limited review of this issue for a Master’s Dissertation (MD, TP).

Inclusion and exclusion criteria for the first systematic review

We will include studies involving:
Medical undergraduate students
At least communication skills assessment tools
At least observational assessment tools for both patients and simulated patients
We will not exclude non-English language publications, (i.e. include English, Dutch, German, French, Spanish and Scandinavian languages)
Both quantitative and qualitative data.
We will include studies that describe the development, validation or use of assessment tools for communication in the medical consultation.

Exclusion:
We will concentrate on medical doctors; we will exclude other health care professions, dental practitioners, vets, etc.
We will exclude:
- Articles describing constructing an tool, but not tested and validated in an educational context
- Articles describing a tool only used for selection purposes
- Tools used exclusively for postgraduate medical specialist training
- No full article (i.e. not an opinion-based work)

Data selection and extraction:

We will develop a template on the basis of BEME experience and our “dry run” with selected papers during summer 2010. We will enlist the advice of the Warwick Medical School Librarian to ensure our search strategy is inclusive focussed and efficient. The final search for this review will be performed in late autumn after our project meeting at Warwick. Two reviewers will independently evaluate the retrieved articles to determine their suitability for inclusion in the review against the eligibility criteria defined. The two reviewers will make decisions regarding inclusion based on discussion and consensus. If consensus cannot be established then a third author from a different country will be consulted to reach a final decision, however we will ensure that this group consists of a representation from both a physician and psychologist to ensure that representation of different perspectives is covered.

The final selection of full text articles that meet the criteria and are to be included in the analysis will be apportioned between countries for data extraction with two reviewers from each country performing the data extraction. Discrepancies will be resolved by discussion.

Data handling
We will use Endnote and EXCEL to build this data-base. We expect heterogeneity and will work to develop and later report ways to synthesise data meaningfully and display our results so as to achieve the pragmatic aims discussed above and support medical educationalists in their teaching and assessment. Quality assessments of quantitative and qualitative data will be based on our own past work[31, 35, 36]and the expertise available to us at Maastricht, Antwerp and Warwick.

Project timetable:

Meetings:
1. To finalise proposal in terms of research question/s and approach. Conducted 9/4/10 Dept.GP Maastricht
2. Local in country meeting of team members to discuss experience with inclusion/exclusion criteria, template and pilot study (2010 summer) to underpin next meeting.
3. To exchange experience of dry run. To finalise agreed operationalisation framework, criteria and template for data collection. Date 5thNovember to be arranged at Warwick .
4. Final search November 2010, review and data extraction November-February 2011
5. To review extracted data in database and main messages, agree on a write up procedure February 2011 meeting Antwerp.
6. To review and refine final report and dissemination strategy.
Summary:
Final protocol June 2010
Literature search Preliminary 3 months from final protocol, definitive search after Warwick meeting 5th November 2010.
Review active November 2010 –February 2011

<table>
<thead>
<tr>
<th>Activity</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finalise proposal</td>
<td>9/4/10 Maastricht</td>
</tr>
<tr>
<td>Pilot study: in country team meeting to discuss experience (criteria, template, analysis, synthesis, etc).</td>
<td>July/August 2010</td>
</tr>
<tr>
<td>Finalise operational template for 1st review</td>
<td>Oct 2010 Warwick</td>
</tr>
<tr>
<td>Search performed, papers apportioned to team members</td>
<td>Nov 2010 Warwick</td>
</tr>
<tr>
<td>Team performs systematic review</td>
<td>Nov-Jan /Feb 2011</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>Feb/March 2011 Antwerp</td>
</tr>
<tr>
<td>Write up review produced</td>
<td>March-April 2011</td>
</tr>
<tr>
<td>Final report &amp; dissemination</td>
<td>By July 2011</td>
</tr>
</tbody>
</table>

Dissemination
Publication of first BEME review.
Conference presentation own national and international.
Workshop for European General Practice Research Network (EGPRN) on “How to develop an approach to systematic review”. This would also be suitable to submit as a workshop to the above conferences with appropriate re-focussing.

Conflict of interest: none

Update: within 5 years of publication

References


