CPG 2002 – Oral Presentations – Saturday, June 8th

Track 2: CPGs – Development and Implementation

CPG Development – an international comparison

GENE FEDER[1], JAKO BURGERS[2]. *[1] Barts and The London, Queen Mary's School of Medicine and Dentistry, University of London, UK. [2] Centre for Quality of Care Research (WOK), University of Nijmegen, Netherlands.*

Introduction
Professional and bureaucratic organisations are busy developing clinical guidelines across the globe. Historically, methods of guideline development have varied between and within countries. As part of the AGREE programme that aimed to advise the European Commission on guideline development, dissemination and implementation, we investigated variation (or uniformity) between guideline development models and guideline content.

Guideline development models
In most of the countries studied, guidelines are increasingly produced within structured national guidelines programmes funded by central or state governments, with active participation of professional bodies and a wide group of stakeholders, including patient groups. A survey of 19 national guideline programmes from the North America, the Antipodes and nine European countries showed that all guideline organisations claim to develop evidence-based guidelines. Most use electronic databases to collect evidence, and systematic reviews to analyse the evidence. Consensus about the essential features of guideline programmes is growing. A recent examination of the websites of national guideline development organisations showed continuing harmonisation of methods The international collaboration of guideline organisations and agreement of criteria for guideline quality may bring about further convergence of guideline development methods and an international guideline ideology.

Guideline content
With increasing availability of the same body of research evidence (e.g. MEDLINE), one might expect similar clinical recommendations in guidelines. In another AGREE study we compared 15 guidelines on type 2 diabetes mellitus from 13 countries using qualitative and bibliometric methods. We found a high degree of international consensus in recommendations made for the clinical care of type 2 diabetes despite the variation in cited evidence and preferential citation of evidence from a guideline's country of origin. Globalisation of recommended management of diabetes is not a simple consequence of the globalisation of research evidence or the use of robust methods of guideline development. Professional bodies and cultural factors, such as different health care systems, expectations of health care or socio-economic factors will influence the process of selection, analysis and interpretation of evidence as well as the formulation of recommendations.

Conclusion
The internationalisation of guideline development methods is striking. But this is not a sufficient condition for the convergence of recommendations in guidelines between countries. Convergence of recommendations many not even be a desirable goal. International collaboration should focus on improving guideline development methods and globalising the collection of evidence needed for guideline development rather than the development of international guidelines.

Quality and critical appraisal of CPGs – a relevant topic for health care?

Cluzeau, Françoise, Dr., Department of Public Health Sciences, St Georges Hospital Medical School, London

Objective:

This presentation will describe the development and validation of an international instrument for appraising clinical practice guidelines and will discuss issues surrounding its use in the context of health care decisions.

Methods/context/ project
The Appraisal of Guidelines Research and Evaluation (AGREE) Instrument was developed by an international group of researchers through a multi-staged process of item generation, selection and scaling, field testing and refinement procedures. A small working group generated an initial list of 82 items from validated appraisal instruments and relevant literature that addressed five theoretical quality domains. They also wrote an accompanying user guide. After refinement the reduced list and user guide were circulated to all the AGREE partners and to 15 international experts for their views on the clarity, comprehensiveness, relevance and ease of use. The field test was conducted in winter 1999-2000. 100 guidelines selected from 11 participating countries were evaluated independently by 194 appraisers with the instrument. Following improvement the instrument was further field tested on a random sample of 3 guidelines per country by a new set of 70 appraisers in Autumn 2000.

Results:
The final version of the instrument contains 23 items grouped into six quality domains with a 4-point Likert scale to score each item (scope and purpose, stakeholder involvement, rigour of development, clarity and presentation, applicability, editorial independence). 95% of appraisers found the instrument useful to assess guidelines. Reliability was acceptable for most domains (Cronbach's Alpha ranged from 0.64 to 0.88; ICC ranged from 0.57 to 0.91). Increasing the number of raters resulted in substantially higher ICCs. Guidelines produced as part of an established guideline programme had significantly higher scores on editorial Independence and after the publication of a national policy had significantly higher quality scores on rigour of Development ($p<.005$). Guidelines with technical documentation had higher scores on that domain ($p<.0001$). Kendall's Tau B rank correlation coefficients between the appraisers' domain scores and their overall assessments were all highly significant ($p<0.001$), providing some evidence of criterion validity.

Conclusions:
This is the first time an appraisal instrument for clinical guidelines has been developed and tested internationally. The instrument is sensitive to differences in important aspects of clinical guidelines, and can be used consistently and easily by a wide range of professionals from different backgrounds. Adoption of common standards should improve the consistency and quality of the reporting of guideline developments worldwide and provide a framework to encourage international comparison of guidelines.

Design of CPGs – how to (better) include patient's preferences for better outcomes?

Farquhar Cynthia, Associate Professor in Reproductive Medicine, Department of Obstetrics and Gynaecology, University of Auckland

Producing CPGs that will improve outcomes for consumers is the overarching concern of every guideline developer. It is significant that the AGREE instrument includes seeking the views and preferences of patients' as one of the criteria for assessing the quality of a clinical practice guideline. Taking account of consumer concerns is the challenge to all guideline developers. Clinical researchers are also increasingly being challenged to design projects with endpoints that are relevant to patients, and not just health care providers.

Research into barriers to successful guideline implementation frequently reflect that patients want treatments not recommended by guidelines. Therefore it is important that guideline developers seek patients' views both in order to include their views and perspectives so that the most appropriate studies and research is reviewed, and also to assist with the effective uptake of the guideline.

There are various ways of seeking patients' views; inclusion of consumer representatives on guideline development teams, formal surveys of patients and the use of focus groups to test consumer

resources, are all possible methods. More recently the use of decision analyses and qualitative research methods have been used to get to the crux of patient decision making, especially within different cultural groups. The New Zealand Guidelines Group recommends a minimum of two consumers for each guideline development group, and is currently developing training programmes designed for consumer participants on guideline teams.

Successful guideline development ultimately needs to influence practitioners, policy makers and the consumers and in particular to consider what information consumers need. The production of consumer resources based on the guidelines has been a useful approach to dissemination and overcoming barriers. In some cases, resources designed for particular cultural groups are also being developed. Stories about guidelines in the popular media have also aided dissemination and reduce the barriers by communicating directly with patients.

Efficient and effective implementation strategies

Grimshaw, Jeremy Professor; Thomas, Ruth; Maclennan G; Vale L; Fraser C; Ramsay Craig Dr; Eccles Martin Professor; Whitty Paula Dr; Matowe Lloyd Dr; Wensing Michel Dr; Dikstra Rob Dr; Cinical Epidemiology Programme, Health Research Institute, University of Ottawa, Canada

Objective:
To describe the current knowledge base about the effectiveness and efficiency of guideline dissemination and implementation strategies.

Methods / Context / Project:
Systematic review of rigorous evaluations of guideline dissemination and implementation strategies. Critical appraisal of related economic evaluations.

Results:
235 studies reporting 309 comparisons met the inclusion criteria. The overall quality of the studies was poor. Seventy three per cent of comparisons evaluated multi faceted interventions although the maximum number of replications of a specific multifaceted intervention was 11 comparisons. Overall the majority of comparisons reporting dichotomous process data (86.6%) observed improvements in care; however, there was considerable variation in the observed effects both within and across interventions. The majority of interventions observed modest to moderate improvements in care. For example the median absolute improvement in performance across interventions ranged from 13.1% in 13 cluster randomised comparisons of reminders, 8.1% in four cluster randomised comparisons of dissemination of educational materials, 7.0% in five cluster randomised comparisons of audit and feedback and 6.0% in thirteen cluster randomised comparisons of multi faceted interventions involving educational outreach. We found no relationship between the number of component interventions and the effects of multifaceted interventions. Only 29% comparisons reported any economic data. Overall the methods of the economic evaluations and cost analyses were poor. The viewpoint adopted in economic evaluations was only stated in ten studies. The methods to estimate costs were comprehensive in about half of the studies, and few studies reported details of resource use. Due to the poor quality of reporting of the economic evaluation, data on resource use and cost of guideline development, dissemination and implementation was not available for most of them.

Conclusions:
There is an imperfect evidence base to support decisions about which guideline dissemination and implementation strategies are likely to be efficient under different circumstances. Decision makers need to use considerable judgement about how best to use the limited resources they have for guideline dissemination and implementation to maximise population benefits based upon consideration of the potential clinical areas for clinical effectiveness activities, the likely benefits and costs required to introduce guidelines and the likely benefits and costs as a result of any changes in provider behaviour.

Cultural heterogenity in CPGs – Only a problem of local tailoring?

Lorenz, Wilfried; Koller, Michael; Kopp, Ina; Sitter, Helmut; Celik, Ilhan; Nies, Christoph, University of Marburg

Cultural heterogeneity in Clinical Practice Guidelines (CPGs) is often used as an argument of criticism against international guidelines of high quality – somehow like developing protective measures in international trading. We elucidate a totally different perspective applying social psychology as the basic discipline to analyse culture's impact on clinical decision making (Philipchalk, 1995).

As a first example, we compared clinical algorithms for the management of anastomotic leakage and sepsis following colorectal carcinoma surgery from 10 hospitals of 5 countries. We measured algorithm complexity by CASA scores (Pearson et al. 1992) and obtained a variation for solving the same problem by a factor of 20.

As a second example, we present the outcomes of a metaanalysis on minimal-invasive versus conventional surgery in symptomatic gallstone disease in various countries and the results from a qualitative analysis based on asking patients and doctors to rank individual outcomes according to their personal preferences. The outcome variable "restauration of full physical fitness" was ranked very important, but was not investigated in any of the trials included in the metaanalysis. Apparently, different outcomes than those analysed in EBM are important for patients in Germany.

As a third example, quality of life was assessed as an outcome in patients with solid tumours of various organs and was shown to depend totally on psychological baseline variables such as negative affect and perceived social stigma. Negative affect, however, depended on cultural background and national origin as shown in the study of Diener and Suh (1999).

Cultural heterogeneity in CPGs is more than a problem of local tailoring: it is a main criterion of good quality.

Measuring the effect of CPGs on patient outcome

Slutsky, Jean PA, MSPH, Director National Guideline Clearinghouse and National Quality Measures Clearinghouse, Agency for Healthcare Research and Quality, USA

Objective:
This presentation will discuss assessing the implementation of guidelines into practice through the use of well-designed quality measures.

Methods / Context:
The difficulty in transferring research findings into the hands of clinicians has been well documented. One United States researcher has postulated it can take as long as seventeen years for research findings to make their way into practice (Balas EA, Boren SA. Managing Clinical Knowledge for Health Care Improvement. Yearbook of Medical Informatics. Schattauer, 2000: 65-70). High quality clinical practice guidelines, when implemented, have been shown to result in better patient outcomes. Like original research findings, CPGs are also hard to implement into practice. Several studies and reviews have suggested coordinated multi-faceted implementation and measurement strategies are the most effective way to change practice. Quality measures, like CPGs, must be well-developed and rational in order to measure accurately and fairly. The large number and variable rigor of quality measures makes choosing them carefully especially important.

Results:
There have been several successful attempts to identify quality attributes of CPGs (AGREE, National Guideline Clearinghouse, German Guideline Clearinghouse). These activities have provided an impetus for the CPG developer community to produce their best CPGs. Attempts to similarly describe quality measures have been more diverse and less coordinated. Despite the heterogeneity of efforts, a broad common framework of attributes of quality measures has emerged. Within, this framework, many

challenges have been identified: how to identify longitudinal changes while updating measures to reflect the most current science; ensuring that measures are accurately measuring what they are intended to measure; and basing measures on the best available evidence, systematic review, or CPGs. Preliminary surveys of quality measures in the United States against a framework of attributes show that a good number of measures are poorly documented, many are proprietary and, therefore, the documentation is not freely available, and many do not measure accurately what they purport to measure.

Conclusions:
One approach to providing critical information about quality measures is to develop a taxonomy of measures. Evidence base for the measure, documentation of measurement data sources and validation attempts can be housed in a large relational database accessible through the Internet. Quality measures can be effective tools for measuring the effect of CPGs on patient outcomes. However, in order to be interpreted in a meaningful way, the measures must be well developed, rigorous, reliable, and valid. Ultimately, doing the right thing and the right time in patient care is what produces the best patient outcomes.

Updating of CPGs

Experiences from the U.K. (& USA)

Eccles, Martin, Prof., Centre for Health Services Research, University of Newcastle upon Tyne, Shekelle, Paul, Dr., Grimshaw, Jeremy, Prof., Woolf, Steven, Prof.

Objective:
To present a set of principles by which to consider when clinical practice guidelines (CPGs) should be updated and to describe the consequences of using them.

Methods / Context / Project:
Considerable resources are being expended internationally on the development of CPGs. Whilst there is increasing consensus about methods for developing evidence-based guidelines, less attention has been paid to the process for assessing when guidelines should be updated. The most common advice is for guidelines to include a scheduled review date. However this could result in wasted resources if a full update is undertaken prematurely within a slowly evolving field, or in guidelines in a rapidly evolving field becoming out-of-date before the scheduled review. Some guidelines state that they should be updated when new information becomes available, however it is unclear how this should be operationalised, and we are unaware of any systematic attempts to devise such methods. In this presentation, we describe a set of principles and a pragmatic model for assessing whether guidelines need to be updated.
The framework suggests that changes in any or all of the following six areas may necessitate updating of a guideline: the evidence on the existing benefits and harms of interventions; the outcomes considered important; the available interventions; the evidence that current practice is optimal; the values placed on outcomes; the resources available for health care. This framework was applied to a set of 17 CPGs using a mixture of expert judgement and limited literature review.

Results:
For 13 guidelines, new evidence and expert judgement indicated that an update was required, three guidelines were judged still valid, and for one guideline we could reach no conclusion. Survival analysis indicates that about half the guidelines were out of date in 5.8 years. The time point at which no greater than 90% of the guidelines were still valid was 3.6 years (95% CI 2.6-4.6).

Conclusions:
It was possible to operationalise the set of principles. More than three quarters of the guidelines needed updating. As a general rule, guidelines should be reassessed for validity every 3 years.