The political dimensions of Evidence-Based Medicine

Patrick Castel (SciencesPo, Center for the sociology of organizations)

SYMPOSIUM – 10 years KCE – ‘Bridging the Gap’
Introduction

- Evidence-Based Medicine is a self-styled paradigm
- The explicit goal of Evidence-based Medicine is to make decision-making more rational: assessment of a new drug or a new technology, diagnosis, treatment...
  - To de-emphasize the authority of individuals
  - A more effective use of available literature
  - A shift away from “intuition, clinical experience, and pathophysiological rationale” (EBM Working group, 1992)

From Evidence-Based Medicine to « Evidence-Based everything » (Fowler, 1997)? Policy, education, nursing...
Introduction

- Three major tools:
  - Randomized control trials;
  - Meta-analyses;
  - Clinical Practice Guidelines: “statements that include recommendations intended to optimize patient care that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options” (IOM 2011)

- Debates about guidelines largely overlap debates about EBM
Introduction

- EBM is not only a rational project but also a political one:
  - Quantification in medicine in general and EBM in particular have been debated and controversial for a long time;
  - EBM is at the core of power relationships between actors:
    - It is shaped by them;
    - It may be a source of destabilization and transformation of these relationships.
- We live in “a world of standards but not a standard world” (Timmermans & Epstein, 2010)
Part I. Quantification in medicine

A long-lived and hotly debated topic
1989: the term “EBM” (McMaster University)

But attempts to quantify and formalize medicine and related debates are far older:

- (Failed) attempts to quantify as soon as the 18th century.
- 19th century: Violent scientific controversy between Claude Bernard (biology and experimentation) vs. statisticians.
- Beginning of the 20th century: the emergence of the randomized clinical trial as “an utopian effort to establish a profession united by a faith in science” (Marks, 1997).
H. Marks (1997), *The Progress of Experiment*

- Early 20th century, RCT to protect the public from the excesses of capitalism – i.e. “to separate the pharmaceutical wheat from the chaff” (Marks).
- What is a well-controlled study? From reliance on the personal credibility of the investigator to a faith in the power of formal statistical methods.
M. Berg (1997), *Rationalizing Medical Work*

- Since the end of WWII, attempts to reduce the discrepancies between medical knowledge and actual medical practices.

- “Recent” conceptualizations (from 1970s): such discrepancies are due to:
  - physicians’ individual cognitive biases;
  - limited capacity to manipulate all relevant scientific data.

- The emergence of (evidence-based) guidelines
To sum up

- The emergence of clinical trials and guidelines:
  - As a (non-consensual and controversial) professional initiative,
  - But their diffusion is also due to the quality and cost-control imperatives in the Western health care systems → involvement of governments, national and international agencies, hospitals, insurance companies in the promotion and production of guidelines...
Part II. (Mainly theoretical, but still important) political debates about EBM
CPG as a tool to reduce medical authority?

- Some social scientists and some physicians believe/fear/complain that guidelines are designed:
  - to open the black-box of clinical judgement;
  - to make physicians more accountable;
  - to shift the power away from physicians towards health-services researchers and managers.

△ CPG = part of the ‘Audit Society’ (Power)?
'Resisting cookbook medicine'?

Although every patient is different, Emory Healthcare is finding that following standardized recipes saves lives.
EBM and health-care rationing

- EBM for cost-cutting?
- Ex: Strong controversies in the UK (NICE) → a limited access to unappraised treatments and a perpetuation of the geographical inequity problem? A ‘technocratic fix’? (Syrett, 2003)
- Ex: Daily Haemodialysis legislation in US (Medicare and Medicaid) → clinical trials to postpone the decision... and the medical expenses! (Gordon, 2006).
What about the patients’ experience and values?

- Does EBM subvert the integrity of doctor-patient communication and the singularity of doctor-patient relationship?
- EBM would strip patients of their stories and the meaning of their experience.
  - An exclusive focus on security and efficiency criteria;
  - A lack of assessment of important criteria for deaf children and their families: language learning, social integration, psychological effects...
Intermediate conclusion

- The above discussed critiques rarely rely on empirical research:

  - However, they represent important political concerns to take into account (they are not merely ‘irrational’)

- An important issue = ‘bringing the patient back in’:
  - “The next step” of guideline development?
  - ‘empowering’ patients is challenging (Callon & Rabeharisoa, 2004)
  - Clinical expertise and patients’ experience and values as second-class “citizens” in the realm of EBM?
Part III. EBM and power struggles

EBM as a strategic tool
Why Don't Physicians Follow Clinical Practice Guidelines?
A Framework for Improvement

Michael D. Cabana, MD, MPH
Cynthia S. Rand, PhD
Neil R. Powe, MD, MPH, MBA
Albert W. Wu, MD, MPH
Modena H. Wilson, MD, MPH
Paul-André C. Abboud, MD
Haya R. Rubin, MD, PhD

Clinical practice guidelines are "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances." Their successful implementation should improve quality of care by decreasing inappropriate variation and expediting the application of effective advances to everyday practice.

Despite wide promulgation, guidelines have had limited effect on changing physician behavior. Little is known about the process and factors involved in changing physician practices in response to guidelines.

Objective To review barriers to physician adherence to clinical practice guidelines.

Data Sources We searched the MEDLINE, Educational Resources Information Center (ERIC), and HealthSTAR databases (January 1966 to January 1998); bibliographies; textbooks on health behavior or public health; and references supplied by experts to find English-language article titles that describe barriers to guideline adherence.

Study Selection Of 5658 articles initially identified, we selected 76 published studies describing at least 1 barrier to adherence to clinical practice guidelines, practice parameters, clinical policies, or national consensus statements. One investigator screened titles to identify candidate articles, then 2 investigators independently reviewed the texts to exclude articles that did not match the criteria. Differences were resolved by consensus with a third investigator.

Data Extraction Two investigators organized barriers to adherence into a framework according to their effect on physician knowledge, attitudes, or behavior. This organization was validated by 3 additional investigators.

Data Synthesis The 76 articles included 120 different surveys investigating 293 potential barriers to physician guideline adherence, including awareness (n = 46), familiarity (n = 31), agreement (n = 33), self-efficacy (n = 19), outcome expectancy (n = 8), ability to overcome the inertia of previous practice (n = 14), and absence of external barriers to perform recommendations (n = 34). The majority of surveys (70 [58%] of 120) examined only 1 type of barrier.

Conclusions Studies on improving physician guideline adherence may not be generalizable, since barriers in one setting may not be present in another. Our review offers a differential diagnosis for why physicians do not follow practice guidelines, as well as a rational approach toward improving guideline adherence and a framework for future research.

JAMA. 1999;282:1458-1465

www.jama.com
Guideline as a professional tool?

- Professionals are the main producers of guidelines:
  - Of the 1223 CPG between 1999, 52% by professional organizations (Timmermans & Kolker, 2004).

- In some cases, guidelines have been developed by segments of medicine to restore their legitimacy:
  - Insurance physicians in the Netherlands (Berg et al., 2000)
  - Cancer center doctors in France (Castel & Friedberg, 2010; Castel, 2009)
  - NCCN vs. HMO in the US
EBM and intra-professional competition

The case of cancer care in France (Castel, 2009; Castel & Friedberg, 2010):

- For French CCC, a tool:
  - to improve their relations with community, local hospitals;
  - to reinstate themselves as national leaders in oncology.
- For individual physicians (esp. medical oncologists), a tool to improve their position in the management of cancer care.

Two lessons for guideline developers:

- Competition between medical groups may be a lever to involve physicians in the production and implementation;
- But there is also a risk to be turned into a tool by this competition.
Part IV. « A world of standards but not a standard world »
(Timmermans & Epstein, 2010)

Diversity of guidelines and of guideline organizations
Many guideline developers

- KCE (Belgium)
- NICE (UK)
- SIGN (Scotland)
- HAS Haute Autorité de Santé (France)
- ÄZQ Ärztliches Zentrum für Qualität in der Medizin (Germany)
- NCCN National Comprehensive Cancer Network (USA)
- And many, many more…
Levels of Evidence 1 (Sackett et al. 2000)

• 1A = Systematic Review of Randomized Controlled Trials (RCTs)
  – 1B = RCTs with Narrow Confidence Interval
  – 1C = All or None Case Series
• 2A = Systematic Review Cohort Studies
  – 2B = Cohort Study/Low Quality RCT
  – 2C = Outcomes Research
• 3A = Systematic Review of Case-Controlled Studies
  – 3B = Case-controlled Study
• 4 = Case Series, Poor Cohort Case Controlled
• 5 = Expert Opinion
Levels of evidence 2 (‘NICE guidelines manual’, 2007)

Level source of evidence

1++ High-quality meta-analyses, systematic reviews of randomised controlled trials (RCTs) or RCTs with a very low risk of bias
1+ Well-conducted meta-analyses, systematic reviews of RCTs or RCTs with a low risk of bias
1– Meta-analyses, systematic reviews of RCTs or RCTs with a high risk of bias
2++ High-quality systematic reviews of case–control or cohort studies; high-quality case–control or cohort studies with a very low risk of confounding, bias or chance and a high probability that the relationship is causal
2+ Well-conducted case–control or cohort studies with a low risk of confounding, bias or chance and a moderate probability that the relationship is causal
2– Case–control or cohort studies with a high risk of confounding, bias or chance and a significant risk that the relationship is not causal
3– Non-analytical studies (for example case reports, case series)
4– Expert opinion, formal consensus
Levels of evidence 3: NCCN Categories of Evidence and Consensus

**Category 1:** Based upon high-level evidence, there is uniform NCCN consensus that the intervention is appropriate.

**Category 2A:** Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate.

**Category 2B:** Based upon lower-level evidence, there is NCCN consensus that the intervention is appropriate.

**Category 3:** Based upon any level of evidence, there is major NCCN disagreement that the intervention is appropriate.

All recommendations are category 2A unless otherwise noted.
Levels of evidence 4 (GRADE)

Table 1 – Levels of evidence according to the GRADE system.

<table>
<thead>
<tr>
<th>Quality level</th>
<th>Definition</th>
<th>Methodological Quality of Supporting Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>We are very confident that the true effect lies close to that of the</td>
<td>RCTs without important limitations or overwhelming evidence from observational</td>
</tr>
<tr>
<td></td>
<td>estimate of the effect</td>
<td>studies</td>
</tr>
<tr>
<td>Moderate</td>
<td>We are moderately confident in the effect estimate: the true effect is</td>
<td>RCTs with important limitations (inconsistent results, methodological flaws,</td>
</tr>
<tr>
<td></td>
<td>likely to be close to the estimate of the effect, but there is a</td>
<td>indirect, or imprecise) or exceptionally strong evidence from observational</td>
</tr>
<tr>
<td></td>
<td>possibility that it is substantially different</td>
<td>studies</td>
</tr>
<tr>
<td>Low</td>
<td>Our confidence in the effect estimate is limited: the true effect may be</td>
<td>RCTs with very important limitations or observational studies or case series</td>
</tr>
<tr>
<td></td>
<td>substantially different from the estimate of the effect</td>
<td></td>
</tr>
<tr>
<td>Very low</td>
<td>We have very little confidence in the effect estimate: the true effect is</td>
<td></td>
</tr>
<tr>
<td></td>
<td>likely to be substantially different from the estimate of the effect</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 – Strength of recommendations according to the GRADE system.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>The desirable effects of an intervention clearly outweigh the undesirable</td>
</tr>
<tr>
<td></td>
<td>effects (the intervention is to be put into practice), or the undesirable</td>
</tr>
<tr>
<td></td>
<td>effects of an intervention clearly outweigh the desirable effects (the</td>
</tr>
<tr>
<td></td>
<td>intervention is not to be put into practice)</td>
</tr>
<tr>
<td>Weak</td>
<td>The desirable effects of an intervention probably outweigh the undesirable</td>
</tr>
<tr>
<td></td>
<td>effects (the intervention probably is to be put into practice), or the</td>
</tr>
<tr>
<td></td>
<td>undesirable effects of an intervention probably outweigh the desirable</td>
</tr>
<tr>
<td></td>
<td>effects (the intervention probably is not to be put into practice)</td>
</tr>
</tbody>
</table>
Assessing EBM’s evidence base

- Pragmatic and situated evaluations of what counts as evidence (and non-evidence)
  - ‘Evidence-searched guidelines’ (Knaapen, 2013)

- “Coming to terms with the external world” (Knaapen et al., 2010)

- Moreira (2005): five types of repertoires to justify a guideline’s content:
  - robustness (or science), acceptability (politics), usability (practice) and methodological adequacy (process).

- Guideline development as a textual activity (Knaapen et al., 2010)
Editorial

Clinical practice guidelines: towards better quality guidelines and increased international collaboration

R Grol\textsuperscript{\textregistered}1, FA Cluzeau\textsuperscript{2} and JS Burgers\textsuperscript{1}

1 University Medical Centre Nijmegen, Nijmegen, The Netherlands; 2 St George’s Hospital Medical School, London, UK


Keywords: practice guidelines; quality assessment; international network

An acknowledgement of the limitations of quantified evidence by many guideline developers

it must also be kept in mind that each country has its own norms and values that influence the content and presentation of guidelines. Therefore, the aim should not be to develop international guidelines, but to reach international agreement about the requirements for methodology and reporting of guidelines (De Maeseneer and Derese 1999). The AGREE instrument is an
# Meta-standardization of clinical guidelines

(source: Knaapen, 2012 « European regulation and harmonization of Clinical Practice Guidelines »)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Brouwers</td>
<td>Browman</td>
<td>Burgers</td>
<td>Burnand</td>
<td>Burgers</td>
<td>Burnand</td>
<td>Cluzeau</td>
</tr>
<tr>
<td>Eeder</td>
<td>Fervers</td>
<td>Grol</td>
<td>Cluzeau</td>
<td>Fervers</td>
<td>Fervers</td>
<td>Fervers</td>
</tr>
<tr>
<td>Grol</td>
<td>Mäkelä</td>
<td>Ollenschläger</td>
<td>Qureshi</td>
<td>Slutsky</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Littlejohns</td>
<td>Mäkelä</td>
<td>Ollenschläger</td>
<td>Qureshi</td>
<td>Slutsky</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Qureshi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IOM (2011): 8 standards for guidelines

Clinical Practice Guidelines We Can Trust

- Establishing transparency;
- Management of conflict of interest;
- Guideline development group composition;
- Clinical practice guideline–systematic review intersection;
- Establishing evidence foundations for and rating strength of recommendations;
- Articulation of recommendations;
- External review; and
- Updating.
But already criticized...

Rhode Island Medical Journal (2013 Jul 30)

Are the Institute of Medicine’s Trustworthiness Guidelines Trustworthy?

BENJAMIN K. YOUNG, MS; PAUL B. GREENBERG, MD

ing Evidence Foundations” and “Updating”). The IOM standards for the development of CPGs do not meet their own criteria of trustworthiness. Further study is needed to determine the best methodology to evaluate CPGs.

Acknowledgements
The authors thank Dr. Curtis Margo, MD, MPH, of the Department of Ophthalmology, Morsani College of Medicine, University of South Florida, for reviewing this manuscript. He received no financial compensation.

References
Networking and dissemination

About G-I-N
The Guidelines International Network, G-I-N, is a global network, founded in 2002, comprising about 90 organizations and 80 individuals from 43 countries and five continents. The network supports evidence-based health care and improved health outcomes by reducing inappropriate variation throughout the world.

G-I-N’s mission...
...to lead, strengthen and support collaboration and work within the guideline development, adaptation and implementation community.

G-I-N’s principal aims:
- Providing a network and partnerships for guideline organisations, implementers, end-users, researchers, students and other stakeholders
- Assisting members in reducing duplication of effort and improving the efficiency and effectiveness of evidence-based guideline development, adaptation, dissemination and implementation
- Promoting best practice through the development of opportunities for learning and building capacity
- Establishing high quality standards of guideline development, adaptation, dissemination and implementation.
Or the law of the market?

Guidelines for Guidelines: Measuring Trustworthiness

David F. Ransohoff, University of North Carolina at Chapel Hill, Chapel Hill, NC
Harold C. Sox, Dartmouth Institute for Health Policy and Clinical Practice, Lebanon, NH
See accompanying article doi: 10.1200/JCO.2012.46.8371

« We do not have an empirical basis for deciding which elements of CPG developments are essential to good outcomes. Metrics that distinguish the untrustworthy guideline (...) from the merely good and the excellent are necessary for creating a market for excellent trustworthy CPGs. When guideline developers compete on the basis of quality of the guideline, then quality will rise, recommendations may converge, the reasons for disagreement will be clearer, and the public will gain confidence in the medical profession’s efforts to do « what is best for patients ». »
Conclusion
Debates about EBM are typical for debates around standardization (not a specificity of the health sector)

Empirical studies:

- “advocate to stop treating standardization as inherently either good or bad” (Knaapen, 2013)
- conclude that one should not overestimate the universal or reductionist nature of EBM’s evidence base

A call for more studies (anthropology, ethnography, sociology, psychology...) which observe guideline development

A call to better identify and acknowledge the sociopolitical dimensions of guideline development and health technology assessment – to improve their acceptability
Selected references


