EBM 32 years later: Achieved goals and unresolved issues

Gordon Guyatt
Clarity Research Group
Department of Clinical Epidemiology
McMaster University
Conflicts

- Co-Chair GRADE working group
- CSO MAGIC Evidence Ecosystem Foundation
- UpToDate consultant
Plan for talk

- Achieved goals
  - Fundamental Guide to Clinical Practice
  - GRADE and systematic reviews, guidelines

- Achieved goals/Unresolved Issues
  - Hierarchy - ? Has it penetrated
  - COVID – ? Can it continue?
  - Electronic layered format, SOFs - ? Dissemination

- Unresolved issues
  - Can we educate optimally and successfully
  - Can we get right balance rigor and simplicity
Evidence-Based Medicine

An internist sees a 70-year-old man whose main problem is fatigue. The initial investigation reveals a hemoglobin of 90 g/L. The internist suspects iron deficiency anemia. How might she proceed?

The way of the past
When faced with this situation during her training just a few years earlier, the internist was told by the attending physician that one ordered serum ferritin and transferrin saturation and proceeded according to the results. She now follows this path. If both results come back below the laboratory's cut-off, she will not order a serum ferritin level, but not transferrin saturation, which is less powerful and adds no useful information. She also finds that her laboratory's normal range for the test is misleading. The internist estimates the pretest likelihood of iron deficiency and orders the test. When the result is available, she uses data from the article to determine the sensitivity and specificity associated with the ferritin value obtained, and then decides on further management.

Discussion
The way of the future described above depicts an important advance in the inclusion of new evidence into clinical practice. Clinicians were formerly taught to look to authority (whether a textbook, an expert lecturer, or a local senior physician) to resolve issues of patient management. Evidence-based medicine uses additional strategies, including quickly tracking down publications of studies that are directly relevant to the clinical problem, critically appraising these studies, and applying the results of the best studies to the clinical problem at hand. It may also involve applying the scientific method in determining the optimal management of the individual patient (3).

For the clinician, evidence-based medicine requires skills of literature retrieval, critical appraisal, and information synthesis. It also requires judgment of the applicability of evidence to the patient at hand and systematic approaches to make decisions when direct evidence is not available. The primary purpose of ACP Journal Club is to help make evidence-based medicine more feasible for internists by extracting new, sound clinical evidence from the morass of the biomedical literature so that practitioners can get at it.

Gordon H. Guyatt, MD, MSc

References

interested in acquiring or enhancing these skills? Attend the ACP Annual Meeting, 11-13 April 1991, for workshops on searching the literature on MEDLINE and using the Clinical Literature to Solve Clinical Problems – The Editor

ACP Journal Club March/April 1991
Evidence-Based Medicine

A New Approach to Teaching the Practice of Medicine

Evidence-Based Medicine Working Group

EBM represents a new paradigm for medical practice
Less emphasis intuition, clinical experience, pathophysiologic rational
Instead evidence from clinical research

Users' Guides to the Medical Literature

I. How to Get Started

Andrew D. Oxman, MD, MSc; David L. Sackett, MD, MSc; Gordon H. Guyatt, MD, MSc;
for the Evidence-Based Medicine Working Group

25 part series in JAMA to 2000
Wide acceptance of basic principles
- Hierarchy of evidence, systematic reviews, V and P (SDM)
- Every undergraduate medical/health program
  - Curricula based on Users’ Guides
- 2007 BMJ EBM one of top 10 advance since 1840

Figure 1: Effect of steroids on mortality in community acquired pneumonia

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Corticosteroids</th>
<th>Control</th>
<th>Risk Ratio M-H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
</tr>
<tr>
<td>Blum 2015</td>
<td>16</td>
<td>392</td>
<td>13</td>
</tr>
<tr>
<td>Confalonieri 2005</td>
<td>0</td>
<td>23</td>
<td>8</td>
</tr>
<tr>
<td>El-Ghamrawy 2006</td>
<td>3</td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td>Fernández-Serrano 2011</td>
<td>1</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>Mark 1993</td>
<td>1</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>McCarthy 1972</td>
<td>3</td>
<td>40</td>
<td>9</td>
</tr>
<tr>
<td>Meijvis 2011</td>
<td>9</td>
<td>151</td>
<td>11</td>
</tr>
<tr>
<td>Naïve 2013</td>
<td>4</td>
<td>60</td>
<td>6</td>
</tr>
<tr>
<td>Sabry 2011</td>
<td>2</td>
<td>40</td>
<td>6</td>
</tr>
<tr>
<td>Snijders 2010</td>
<td>6</td>
<td>104</td>
<td>6</td>
</tr>
<tr>
<td>Torres 2015</td>
<td>6</td>
<td>61</td>
<td>9</td>
</tr>
<tr>
<td>Wagner 1956</td>
<td>1</td>
<td>52</td>
<td>1</td>
</tr>
</tbody>
</table>

Total (95% CI) 977 997 100.0% 0.67 [0.47, 0.97]

Total events: 52 79
Heterogeneity: Tau² = 0.03; Chi² = 11.84, df = 11 (P = 0.38); I² = 7%
Test for overall effect: Z = 2.12 (P = 0.03)
Grading quality of evidence and strength of recommendations

GRADE Working Group

Clinical guidelines are only as good as the evidence and judgments they are based on. The GRADE approach aims to make it easier for users to assess the judgments behind recommendations.

Summary

Users of clinical practice guidelines and other recommendations need to know how much confidence they can place in the recommendations. Systematic and explicit methods of making judgments can reduce errors and improve communication. We have developed a system for grading the quality of evidence and the strength of recommendations that can be applied across a wide range of interventions and contexts. In this article we present a summary of our approach from the perspective of a guideline user. Judgments about the strength of a recommendation require consideration of the balance between benefits and harms, the quality of the evidence, translation of the evidence into specific circumstances, and the certainty of the baseline risk. It is also important to consider costs (resource utilisation) before...
>110 organizations have adopted GRADE
<table>
<thead>
<tr>
<th>Study Design</th>
<th>Confidence in estimates</th>
<th>Lower if</th>
<th>Higher if</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trials</td>
<td>High</td>
<td>Risk of bias</td>
<td>Large Effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Serious</td>
<td>+ 1 Large</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very serious</td>
<td>+ 1 Very large</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>Inconsistency</td>
<td>Dose response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Serious</td>
<td>+1 Evidence of a gradient</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very serious</td>
<td></td>
</tr>
<tr>
<td>Observational studies</td>
<td>Low</td>
<td>Indirectness</td>
<td>All plausible confounding</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Serious</td>
<td>+1 Would reduce a demonstrated effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very serious</td>
<td>or</td>
</tr>
<tr>
<td></td>
<td>Very Low</td>
<td>Imprecision</td>
<td>+1 would suggest a spurious effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Serious</td>
<td>when results show no effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very serious</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Publication bias</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Likely</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very likely</td>
<td></td>
</tr>
</tbody>
</table>
Strength of Recommendation

- Strong recommendation
  - Benefits clearly outweigh risks/hassle/cost
  - Risk/hassle/cost clearly outweighs benefit

- Rating down
  - Close balance evidence and harms/burdens
  - Low certainty of evidence
  - Uncertainty, variability values and preference
  - Issues with feasibility, acceptability, equity
Variability in patient preference
- Strong, almost all same choice (> 90%)
- Weak, choice varies appreciably

Interaction with patient
- Strong, just inform patient
- Weak, ensure choice reflects values

Use of decision aid
- Strong, don’t bother; weak, use the aid
Plan for talk

- Achieved goals
  - Fundamental Guide to Clinical Practice
  - GRADE and systematic reviews, guidelines

- Achieved goals/Unresolved Issues
  - Hierarchy - ? Has it penetrated
  - COVID – ? Can it continue/extend?
  - Electronic layered format, SOFs - ? Dissemination

- Unresolved issues
  - Can we educate optimally and successfully
  - Can we get right balance rigor and simplicity
Case-Control

Observational Studies

Retrospective

Prospective

RCT

Systematic Review / Meta-analysis

Case Series

Nonsystematic Clinical Experience
From Evidence to Evidence-Based Resources

Hierarchy of Evidence

for Primary Studies

Different hierarchy of designs for each type of question:

- Therapy and harm
  1. Randomized trial
  2. Observational study
  3. Unsystematic observational study

- Prognosis
- Diagnosis
- Differential diagnosis

Level of Processing

- Guidelines
- Decision analyses
- Systematic reviews
- Primary studies

EBM Resources to Search for Answers

- Summaries and Guidelines
- Preappraised Research
  - Synopses and Systematic Reviews
- Nonpreappraised Research
  - and Clinical Queries
EBM and COVID-19

- Rapid conduct of platform clinical trials
- Vaccines, 3 drugs for non-severe, 4 for severe
- Rapid production of trustworthy guidelines
RAPID RECOMMENDATIONS

A living WHO guideline on drugs for covid-19

the *bmj* | *BMJ* 2020;370:m3379 | doi: 10.1136/bmj.m3379

---

**Recommendation 1**

Patients with severe and critical covid-19

We recommend corticosteroids

---

**Evidence profile**

<table>
<thead>
<tr>
<th>Within 28 days</th>
<th>Favours usual supportive care</th>
<th>No important difference</th>
<th>Favours corticosteroids</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality with critical illness</td>
<td>415</td>
<td>328</td>
<td>87 fewer</td>
</tr>
<tr>
<td>Mortality with severe illness</td>
<td>334</td>
<td>267</td>
<td>67 fewer</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Serious adverse events</th>
<th>Events per 1000 people</th>
<th>Evidence quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastrointestinal bleeding</td>
<td>48</td>
<td>No important difference</td>
</tr>
<tr>
<td>Super-infections</td>
<td>186</td>
<td>No important difference</td>
</tr>
<tr>
<td>Hyperglycaemia</td>
<td>286</td>
<td>46 fewer</td>
</tr>
<tr>
<td>Neuromuscular weakness</td>
<td>69</td>
<td>No important difference</td>
</tr>
<tr>
<td>Neuropsychiatric effects</td>
<td>35</td>
<td>No important difference</td>
</tr>
</tbody>
</table>

---

https://www.who.int/publications/i/item/WHO-2019-nCoV-therapeutics-2021.1
Plan for talk

- **Achieved goals**
  - Fundamental Guide to Clinical Practice
  - GRADE and systematic reviews, guidelines

- **Achieved goals/Unresolved Issues**
  - Hierarchy - ? Has it penetrated
  - COVID – ? Can it continue?
  - Electronic layered format, SOFs - ? Dissemination

- **Unresolved issues**
  - Can we educate optimally and successfully
  - Can we get right balance rigor and simplicity
Practitioners of evidence based care

Not all clinicians need to appraise evidence from scratch but all need some skills

BMJ VOLUME 320  8 APRIL 2000  bmj.com

10.1136/bmjebm-2020-111542

Understanding of research results, evidence summaries and their applicability—not critical appraisal—are core skills of medical curriculum

Kari A O Tikkinen 1,2, Gordon H Guyatt 3
Cross-sectional, paper-based survey
Academic centers in 8 countries,
Internal and family medicine, 531/610 (87%)

Figure 3: Understanding of the presentation approaches, \( n = 531 \)

<table>
<thead>
<tr>
<th>Understanding, % correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Difference</td>
</tr>
<tr>
<td>Relative Risk</td>
</tr>
<tr>
<td>Ratio of Means</td>
</tr>
<tr>
<td>SMD</td>
</tr>
<tr>
<td>MID units</td>
</tr>
<tr>
<td>Natural units</td>
</tr>
</tbody>
</table>

Problem: Clinicians don’t understand results
So how can they do shared decision-making?
Success of EBM, Users Guides and GRADE
- Make complex simple – some sacrifice
- Lost balance: Cochrane RoB 2
- Alternative instrument
  - 12 experts: methods paper 1 1\textsuperscript{st}/last, 2 others
  - Only one said no
- ROBINS-I right off the rails
  - Smart PhD at defence: “Torture”
- GRADE risks same problem
Conclusion

Achieved goals

- EBM has changed practice of medicine worldwide
- Three principles widely acknowledged (V and P, SDM less)
- GRADE now core in systematic reviews, guidelines
  - Guidelines becoming trustworthy, UpToDate, Dynamed

Partially and unfulfilled

- COVID: rapid trials, living reviews/guidelines, optimal format
- Education clinicians understand magnitude of effects, SDM
- Return optimal balance methodological rigor versus simplicity