The who, what, why, where, and when of Clinical Practice Guidelines (CPGs)

Figure 1  Number of guidelines in PubMed.
Appraisal of Guidelines for Research and Evaluation (AGREE) II

DOMAIN 1. SCOPE AND PURPOSE
DOMAIN 2. STAKEHOLDER INVOLVEMENT
DOMAIN 3. RIGOUR OF DEVELOPMENT
DOMAIN 4. CLARITY OF PRESENTATION
DOMAIN 5. APPLICABILITY
DOMAIN 6. EDITORIAL INDEPENDENCE
OVERALL GUIDELINE ASSESSMENT
Grading of Recommendations Assessment, Development and Evaluation (GRADE) - Quality

High Quality (Level A): Further research is very unlikely to change our confidence in the estimate of effect.

Moderate Quality (Level B): Further research is likely to have an important impact on our confidence in the estimate of effect, and may change the estimate.

Low Quality (Level C): Further research is very likely to have an important impact on our confidence in the estimate of effect, and is likely to change the estimate.

Very Low Quality (Level D): Any estimate of effect is very uncertain.
Grading of Recommendations Assessment, Development and Evaluation (GRADE) - Recommendation

Strong recommendation: Most informed patients would choose the option recommended, and clinicians can structure their interactions with patients accordingly.

Weak recommendation: Patient choices will vary based upon their values and preferences, and clinicians must help to ensure that patient care stays true to these values and preferences.

GRADE also allows for “good practice points.” These are recommendations that can be made when it is deemed they will be helpful to the clinician, such as recommendations for shared decision making, but there is no direct evidence to support the recommendation.
Good practice statements typically represent situations in which a large body of indirect evidence, made up of linked evidence including several indirect comparisons, strongly supports the net benefit of the recommended action.

Box 1 Examples of good practice statement previously mistakenly presented as GRADEd recommendations

For patients with congenital adrenal hyperplasia, we recommend monitoring patients for signs of glucocorticoid excess [5].

Triage (ie, take different courses of action for low vs. higher pretest probability) people with tuberculosis symptoms [6].

Health services should be made available, accessible, and acceptable to sex workers based on the principles of avoidance of stigma, nondiscrimination, and the right to health [7].

In patients presenting with heart failure, initial assessment should be made of the patient’s ability to perform routine/desired activities of daily living [8].

Table 1. Questions guideline panels considering good practice statement should ask themselves

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>Is the statement clear and actionable?</td>
</tr>
<tr>
<td>ii</td>
<td>Is the message really necessary?</td>
</tr>
<tr>
<td>iii</td>
<td>Is the net benefit large and unequivocal?</td>
</tr>
<tr>
<td>iv</td>
<td>Is the evidence difficult to collect and summarize?</td>
</tr>
<tr>
<td>v</td>
<td>If a public health guideline, are there specific issues that should be considered (eg, equity)</td>
</tr>
<tr>
<td>vi</td>
<td>Have you made the rationale explicit?</td>
</tr>
<tr>
<td>vii</td>
<td>Is this better to be formally GRADEd?</td>
</tr>
</tbody>
</table>

A word of caution is required: good practice statements may be subject to abuse.
Levels of Evidence:

IA - Evidence from meta-analysis of randomized controlled trials

IB - Evidence from at least one randomized controlled trial

IIA - Evidence from at least one controlled study without randomization

IIB - Evidence from at least one other type of quasi-experimental study

III - Evidence from non-experimental descriptive studies, such as comparative studies, correlation studies, and case-control studies

IV - Evidence from expert committee reports or opinions or clinical experience of respected authorities, or both
Grades of Recommendations:

A - Directly based on Level I evidence

B - Directly based on Level II evidence or extrapolated recommendations from Level I evidence

C - Directly based on Level III evidence or extrapolated recommendations from Level I or II evidence

D - Directly based on Level IV evidence or extrapolated recommendations from Level I, II, or III evidence
Typically “evidence-based” guideline recommendations are not based on “solid” evidence.

<table>
<thead>
<tr>
<th>EVIDENCE LEVEL</th>
<th>Cardiology</th>
<th>Infectious disease</th>
<th>Endocrinology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 or A based on RCTs</td>
<td>11%</td>
<td>14%</td>
<td>6%</td>
</tr>
<tr>
<td>3 or C based on opinion</td>
<td>48%</td>
<td>55%</td>
<td>35%</td>
</tr>
</tbody>
</table>
The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies

Table 2  Appraisal of Guidelines, Research and Evaluation domain scores of guidelines over time (total sample=608)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope and purpose</td>
<td>44</td>
<td>61</td>
<td>60</td>
<td>71</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Stakeholder involvement</td>
<td>18</td>
<td>38</td>
<td>33</td>
<td>37</td>
<td>0.01</td>
</tr>
<tr>
<td>Rigour of development</td>
<td>14</td>
<td>41</td>
<td>43</td>
<td>44</td>
<td>0.003</td>
</tr>
<tr>
<td>Clarity and presentation</td>
<td>32</td>
<td>56</td>
<td>55</td>
<td>68</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Applicability</td>
<td>10</td>
<td>30</td>
<td>18</td>
<td>23</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Editorial independence</td>
<td>17</td>
<td>30</td>
<td>28</td>
<td>33</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Top Score = 100%

Engaging the right people, quality of evidence appraisal, providing useful tools, and competing interests have not improved in 14 years (1993-2007)

Recent examples of Guideline Quality/Rigour
AGREE II (Appraisal of Guidelines for Research and Evaluation)
is the instrument typically used - 207 guidelines

avg 55% - neuropathic pain - 16 CPGs - range 27%-88% - BMC Anesthesiology 2016;16:12
avg 30% - hypertension - 11 CPGS - range 8%-86% - PLoS ONE 2013 8(1): e53744
avg 32% - asthma - 18 CPGs - range 8%-64% - Chest 2013 144: 390-7
avg 48% - diabetes - 24 CPGs - range 0%-81% - PLoS ONE 2013 8(4): e58625
avg 20% - vancomycin - 12 CPGs - range 4%-73% - PLoS ONE 2013 9(6): e99044
avg 18% - hypertension (China) - 17 CPGs - range 1-36% - BMJ Open 2015;5:e008099
avg 8% - respiratory (China) - 109 CPGs - range 0%-27% - Chest 2015;148:759-766
Glycemic Control for Patients With Type 2 Diabetes Mellitus
Our Evolving Faith in the Face of Evidence

René Rodríguez-Gutiérrez, MD, MSc; Victor M. Montori, MD, MSc
Evidence since 1998 for
Tight glycemic control (A1c 6.5%-7%) vs less tight (A1c 7%-8.5%)

Endpoints - End Stage Renal Disease/dialysis, renal death, blindness or clinical neuropathy
5 large trials, 8 meta-analyses, 2 follow-up trials
31 estimates of outcomes
   2 (6%) suggested benefit
   29 (94%) suggested NO benefit

Endpoints - all-cause mortality, CV mortality, non-fatal MIs, stroke, amputations/PVD
5 large trials, 10 meta-analyses, 5 follow-up trials
78 estimates of outcomes
   10 (13%) suggested benefit
   64 (82%) suggested NO benefit
   4 (5%) suggested harm

Circ Cardiovasc Qual Outcomes. 2016;9:00-00. DOI: 10.1161
despite this, over the last 10 years -
“practice guidelines and published statements offer a consistent and confident consensus, with 100% of the guidelines and 77% to 100% of the statements in favor of tight glycemic control to prevent microvascular complications”

Overall estimates of benefits and harms (micro and macro)

11% of estimates = a benefit
4% of estimates = harm
85% of estimates = no benefit
CONTRIBUTORS

54% non–family physician specialists
17% family physicians - 8% if industry sponsored
11% other clinicians
8% non-clinician scientists
6% nurses
3% pharmacists

69% of guidelines didn’t report conflicts of interest
Guideline sponsorship

2009 - 2,300 guidelines in the National Guideline Clearinghouse

Guideline development

41% - medical speciality societies
22% - government agencies/nonprofit
17% - professional associations
9% - disease specific societies
4% - independent expert panels

at least 2/3 are being developed by groups with a clear potential for important biases

http://www.ncbi.nlm.nih.gov/books/NBK22928/
Prevalence of financial conflicts of interest among panel members producing clinical practice guidelines in Canada and United States: cross sectional study

~50-80% of panel members on guidelines have financial COIs

BMJ 2011;343:d5621 doi: 10.1136/bmj.d5621

EVIDENCE BASED MEDICINE

Why we can’t trust clinical guidelines

Despite repeated calls to prohibit or limit conflicts of interests among authors and sponsors of clinical guidelines, the problem persists. Jeanne Lenzer investigates
Adding “value” to clinical practice guidelines

James P. McCormack PharmD  Peter Loewen PharmD

5 Canadian Guidelines for blood pressure, cholesterol, glucose, and bone density

197 PAGES - 90,000 WORDS

99 (0.1%) words - relevant to patients’ values and preferences

Can Fam Physician 2007;53:1326-27
Looked for info on

Risk estimation (magnitude)
Impact of treatment on risk
Potential harms (magnitude)

“The information presented in these documents is glucose-centric and not organized or presented in a way that could be construed as supporting shared decision making”
Their response

“would like to thank McCormack et al for their thoughtful letter regarding the American Diabetes Association’s Standards of Medical Care in Diabetes”

“agrees that shared decision making is a valuable aspect of diabetes care … that process would be incredibly labor intensive and would make the Standards long and unwieldy”

“Clinical guidelines are the foundation for evidence-based medicine”
Guidelines

Hypertension Canada’s 2016 Canadian Hypertension Education Program Guidelines for Blood Pressure Measurement, Diagnosis, Assessment of Risk, Prevention, and Treatment of Hypertension

~11,800 words - 20 pages

Total mention of values and preferences - 0.19% of the words

“Practitioners are advised to consider patient preferences, values, and clinical factors when determining how to best apply these recommendations at the bedside”

“In the absence of Canadian data to determine the accuracy of risk calculations, avoid using absolute levels of risk to support treatment decisions”
Patients’ Expectations of the Benefits and Harms of Treatments, Screening, and Tests
A Systematic Review

Tammy C. Hoffmann, PhD; Chris Del Mar, MD, FRACGP

BENEFIT - 88% of study authors concluded that participants overestimated benefits

HARM - 67% underestimated harm
Evaluating physician understanding of harms and benefits of common tests and therapies

Paper survey to residents and attending internal medicine physicians – 18 questions – 117 people responded

79% overestimated benefit
66% overestimated harm
67% were unconfident

<table>
<thead>
<tr>
<th>Estimate of benefit in absolute terms</th>
<th>&lt;1%</th>
<th>1 to 5%</th>
<th>5 to 10%</th>
<th>10 to 20%</th>
<th>20 to 45%</th>
<th>45 to 70%</th>
<th>70 to 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green cells are the correct answer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild HTN 5 years</td>
<td>11</td>
<td>25</td>
<td>22</td>
<td>19</td>
<td>7</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Aspirin with risk factors 5 years</td>
<td>8</td>
<td>32</td>
<td>29</td>
<td>17</td>
<td>8</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Aspirin in CVD 5 years</td>
<td>0</td>
<td>16</td>
<td>29</td>
<td>30</td>
<td>16</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Warfarin in Afib 1 year</td>
<td>3</td>
<td>31</td>
<td>29</td>
<td>17</td>
<td>12</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Hip fracture osteoporosis 5 years</td>
<td>3</td>
<td>24</td>
<td>30</td>
<td>24</td>
<td>13</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Death from bleed with PPI 5 years</td>
<td>21</td>
<td>22</td>
<td>20</td>
<td>19</td>
<td>9</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Cancer diagnosis among + screening</td>
<td>4</td>
<td>14</td>
<td>23</td>
<td>35</td>
<td>18</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Major bleeding with ASA 5 years</td>
<td>21</td>
<td>46</td>
<td>21</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Major bleeding with warfarin 1 year</td>
<td>14</td>
<td>42</td>
<td>30</td>
<td>11</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Unnecessary biopsy with screening 10 years</td>
<td>1</td>
<td>9</td>
<td>15</td>
<td>33</td>
<td>26</td>
<td>15</td>
<td>0</td>
</tr>
</tbody>
</table>

JAMA Aug 29 2016
307 subjects using a written questionnaire and interview

## Results

<table>
<thead>
<tr>
<th>Patients</th>
<th>Median acceptable absolute % benefit threshold</th>
<th>% that would take a “safe” drug for 5 years</th>
<th>Absolute % benefit they felt they were getting from their drug</th>
<th>% who wanted to be told percent chance of benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post MI patients</td>
<td>20</td>
<td>32</td>
<td>69</td>
<td>70</td>
</tr>
<tr>
<td>On drugs</td>
<td>20</td>
<td>29</td>
<td>74</td>
<td>68</td>
</tr>
<tr>
<td>No drugs</td>
<td>30</td>
<td>21</td>
<td>56</td>
<td>-</td>
</tr>
</tbody>
</table>

Clin Med 2002;2:527-33
Ability of clinicians to make an estimate of CHD risk

53 residents, 8 fellows, 18 attending physicians

The mean degree of over-estimation compared to the Framingham estimate:

- low-risk scenarios - 7.8 times
- medium-risk scenarios - 2.8 times
- high-risk scenarios - 1.5 times
Factors involved in deciding to start preventive treatment: qualitative study of clinicians’ and lay people’s attitudes

David K Lewis, Jude Robinson, Ewan Wilkinson

Qualitative study using semi-structured interviews

“Many of the preferences expressed by the clinicians and lay people in this study are at odds with recommendations in guidelines”
Differing perceptions of intervention thresholds for fracture risk: a survey of patients and doctors

Did NOT ask patients to consider side effects or drug cost, just the dosing regimen, in the decision

“A typical patient in our study required a 50% absolute fracture risk and 50% relative risk reduction (giving an absolute risk reduction of 25%) before considering long-term drug therapy”

A prominent current guideline … recommends pharmacologic intervention at thresholds of 10-year risk of 20% for major osteoporotic fracture or 3% for hip fracture

125 (77%) of doctors would recommend treatment
24 (21%) of our patient cohort would consider treatment justified

Osteoporos Int 2012;23:2135–40
“In three quarters of the cancer studies ... the majority of patients preferred shared or autonomous decision making. In contrast, this was true for only about half of the studies with non-disease specific study populations”

“the number of patients who prefer participation has increased over the past three decades so that the majority of patients prefer to participate in decisions”
Guidelines and the Law

“As per the Canadian Medical Association Handbook on Clinical Practice Guidelines, guidelines should NOT be used as a legal resource in malpractice cases as “their more general nature renders them insensitive to the particular circumstances of the individual cases.”
Even an authoritative CPG may NOT be found to be determinative of a standard of care.

It is prudent for physicians to be aware of authoritative clinical practice guidelines relevant to their practices. If a clinical decision may be perceived as being contrary to a recognized and accepted CPG, a physician, where appropriate, may consider the following steps: consult with a colleague or relevant specialist, discuss reasonable treatment options with the patient, and document the patient's consent for the chosen treatment.

If deviating from an established CPG, physicians should consider documenting the rationale for doing so, as well as any discussions with the patient about such variance.
Many courts (UK, US, CA)

“The reasonable-patient standard … requires physicians and other health care practitioners to disclose all relevant information about the risks, benefits, and alternatives of a proposed treatment that an OBJECTIVE PATIENT would find material in making an intelligent decision as to whether to agree to the proposed procedure”

JAMA 2016;315:2063-4
Guidelines should provide ballpark estimates of what happens if you DON’T treat/test/screen and if you DO treat/test/screen
An Example of a Guideline that Promotes Discussion Rather than Treatment

Simplified lipid guidelines

Prevention and management of cardiovascular disease in primary care

G. Michael Allan MD CCFP  Adrienne J. Lindblad ACPR PharmD  Ann Comeau MN NP CCN(C)  John Coppola MD CCFP
Brianne Hudson MD CCFP  Marco Mannarino MD CCFP  Cindy McMinis  Raj Padwal MD MSc
Christine Schelstraete  Kelly Zarnke MD MSc FRCP  Scott Garrison MD PhD CCFP  Candra Cotton
Christina Korownyk MD CCFP  James McCormack PharmD  Sharon Nickel  Michael R. Kolber MD CCFP MSc

Can Fam Physician 2015;61:857-67