

What's Wrong With Current Guidelines?



In this session...

- What makes a guideline a Public Health Guideline?
- IOM-Identified Deficiencies in Guidelines
 - Articulation
 - Evidence Quality and Recommendation Strength
- Instruments for Appraising Guideline Quality

What is a public health guideline?

- Public health guidance makes recommendations for populations and individuals on activities, policies and strategies that can help prevent disease or improve health.
- The guidance may focus on a particular topic (such as smoking), a particular population (such as schoolchildren) or a particular setting (such as the workplace).
- Even clinical practice guidelines are population-based

Guidelines Define Recommendations

(Key Action Statements)

- Recommendations differentiate guidelines from reviews
- Recommendations tell health professionals **WHAT TO DO** and **WHEN**
- Recommendations can be stated as **IF-THEN** rules
- Recommendation is the unit of implementation (**NOT** the guideline)

WHO: Guidelines for the programmatic management of drug-resistant tuberculosis: 2011 update

- Patients with MDR-TB should be treated using mainly ambulatory care rather than models of care based principally on hospitalization

WHO: Guidelines for the programmatic management of drug-resistant tuberculosis: 2011 update

- Patients with MDR-TB should be treated using mainly ambulatory care rather than models of care based principally on hospitalization

IF patient with MDR-TB

THEN treat using mainly ambulatory care (rather than using models of care based on hospitalization)

Update to CDC's *Sexually Transmitted Diseases Treatment Guidelines, 2010*

- For all patients with gonorrhea, every effort should be made to ensure that the patients' sex partners from the preceding 60 days are evaluated and treated for *N. gonorrhoeae* with a recommended regimen

Update to CDC's *Sexually Transmitted Diseases Treatment Guidelines, 2010 (Slide 2)*

- For all patients with gonorrhea, every effort should be made to ensure that the patients' sex partners from the preceding 60 days are evaluated and treated for *N. gonorrhoeae* with a recommended regimen

IF (diagnosing) patient with gonorrhea

THEN make every effort to ensure that the patient's sex partners from the preceding 60 days are evaluated and treated

Continued Use of CDC Growth Charts for Children Aged 24–59 Months

- Use of the CDC growth charts for children aged 24–59 months is recommended.

Continued Use of CDC Growth Charts for Children Aged 24–59 Months (Slide 2)

- Use of the CDC growth charts for children aged 24–59 months is recommended.

IF child aged 24-59 months
THEN use CDC growth charts

Screening for Cervical Cancer: U.S. Preventive Services Task Force

- The USPSTF recommends screening for cervical cancer in women age 21 to 65 years with cytology (Pap smear) every 3 years or, for women age 30 to 65 years who want to lengthen the screening interval, screening with a combination of cytology and human papillomavirus (HPV) testing every 5 years.

Screening for Cervical Cancer: U.S. Preventive Services Task Force (Slide 2)

- The USPSTF recommends screening for cervical cancer in women age 21 to 65 years with cytology (Pap smear) every 3 years or, for women age 30 to 65 years who want to lengthen the screening interval, screening with a combination of cytology and human papillomavirus (HPV) testing every 5 years.

IF Woman AND Age is 21 to 65

THEN screen for cervical cancer with cytology every 3 yrs

IF Woman AND Age is 30-65 AND want to lengthen screening interval

THEN screen with cytology + HPV testing every 5 yrs

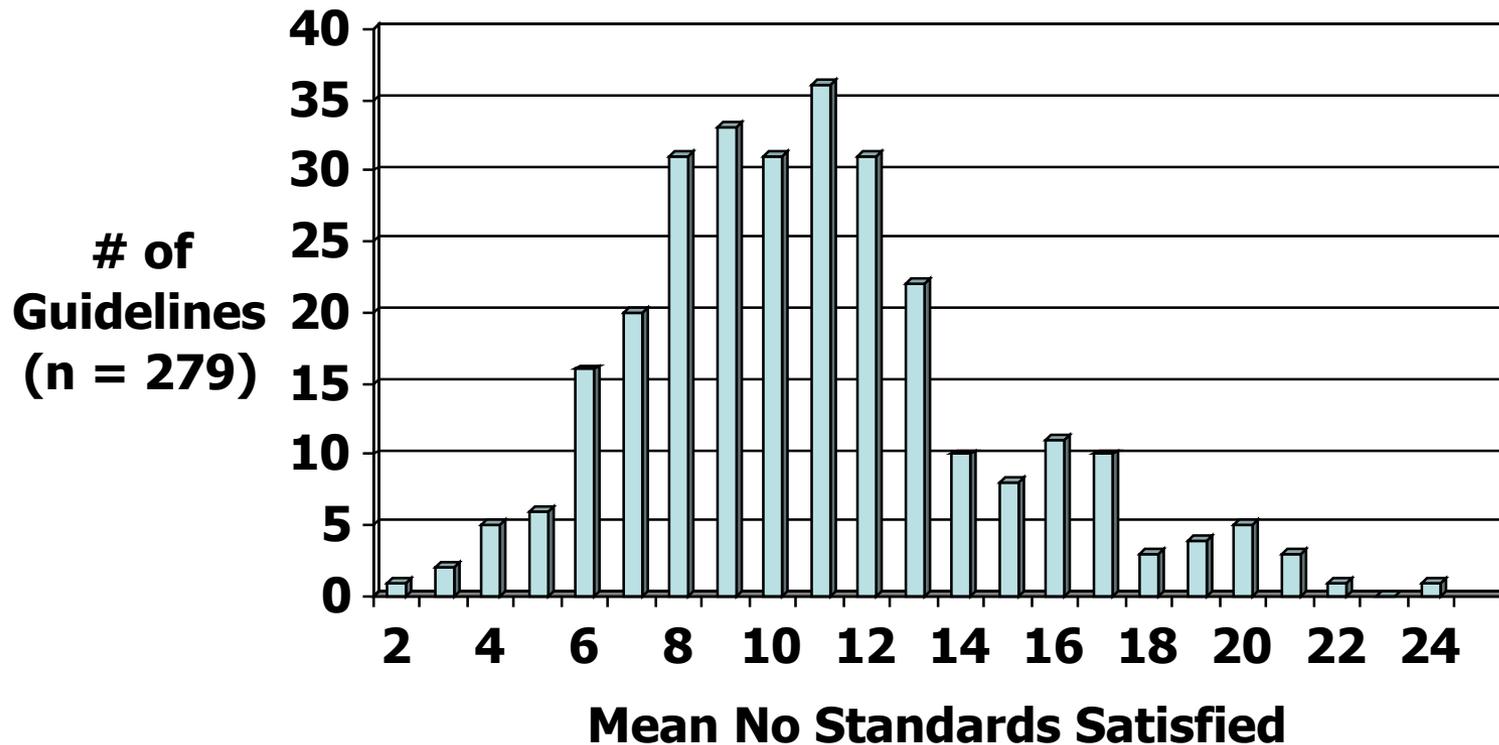
Screening for Cervical Cancer: U.S. Preventive Services Task Force (Slide 3)

- *AND woman who has a cervix*
- *AND NOT woman with a diagnosis of high-grade precancerous cervical lesion OR cervical cancer*
- *AND NOT woman with in utero exposure to diethylstilbestrol*
- *AND NOT woman who is immunocompromised*

Guidelines Have Problems...

- Cluzeau: majority of 60 UK guidelines failed quality criteria (Int J Qual Healthcare 1999)
- Grilli: 431 specialty society guidelines (Lancet 2000)
 - 82% did not apply explicit criteria to grade evidence
 - 87% did not report whether a literature search was performed
 - 67% did not describe type of professionals involved in development

Guideline Quality Evaluation



Shaneyfelt, JAMA 1999

Failure of CPGs to Meet IOM Standards: 2 More Decades of Little, If Any, Progress

Kung J. Arch Intern Med **2012**

- Randomly selected 114 guidelines from NGC
- 18 standards selected from IOM's 25 (7 are too vague!)
- Median number of standards satisfied 8 of 18 (IQR 7-10)
- Shaneyfelt wrote a commentary *"In Guidelines We Cannot Trust"*
 - *Variable and opaque development methods*
 - *Limited and conflicted panel composition*
 - *Lack of significant external review by stakeholders*

Clinical Practice Guidelines We Can Trust



**CLINICAL PRACTICE
GUIDELINES
WE CAN TRUST**

INSTITUTE OF MEDICINE
OF THE NATIONAL ACADEMIES

Guideline Problems

1. Development process is not transparent
2. Conflict of interest
3. Developer teams are insufficiently multidisciplinary and balanced
4. Existing knowledge is not thoroughly reviewed
5. **Articulation of recommendations is not clear**
6. Evidence foundations and strength of recommendations are not explicitly recorded
7. Reviewer input is not broad-based
8. New knowledge is not incorporated

British Airways Memorandum, quoted in *Pilot Magazine*, December 1996

The Landing Pilot is the Non-Handling Pilot until the decision altitude call, when the Handling Non-Landing Pilot hands the handling to the Non-Handling Landing Pilot, unless the latter "calls go around," in which case the Handling Non-Landing Pilot continues handling and the Non-Handling Landing Pilot continues non-handling until the next call of "land" or "go around" as appropriate.

In view of recent confusions over these rules, it was deemed necessary to restate them clearly.

Guidance on the use of glitazones for the treatment of type 2 diabetes

- For people with type 2 diabetes, the use of a glitazone as second-line therapy added to either metformin or a sulphonylurea--as an alternative to treatment with a combination of metformin and a sulphonylurea-- is not recommended except for those who are unable to take metformin and a sulphonylurea in combination because of intolerance or a contraindication to one of the drugs. In this instance, the glitazone should replace in the combination the drug that is poorly tolerated or contraindicated.

Guidance on the use of glitazones for the treatment of type 2 diabetes

- If a patient is unable to take the combination of metformin and sulfonylurea (because of intolerance or contraindication), then the clinician should prescribe a glitazone to replace the drug that is not tolerated.

Authors Should Be Explicit About

- **WHEN** {under what circumstances} ----- **Denominator**
- **WHO** {in the Intended Audience} ----- **Numerator**
- **Ought to** {with what level of obligation}
- **DO WHAT**
- **{To WHOM}** {which members of the target population} -----
- **HOW**
- **WHY**

Guidance on the use of glitazones for the treatment of type 2 diabetes

UNDER WHAT CIRCUMSTANCES?

- If a patient is unable to take the combination of metformin and sulfonylurea (because of intolerance or contraindication), the clinician *WHO?* should prescribe a glitazone to replace the drug that is not tolerated.

OUGHT?

To do WHAT?

AVUL

(Ambiguous, Vague, & Underspecified Language)

- **Ambiguous** statements are interpretable in more than one discrete way
 - Mixing ANDs and ORs
 - Abbreviations: CPZ: Compazine (antiemetic) or chlorpromazine (antipsychotic)
- **Vague**- lack a crisp threshold in a single dimension
 - “high fever,” “elevated LFTs”
- **Underspecified** - lack specificity in multiple dimensions
 - “sufficiently ill to warrant immediate antimicrobial therapy”

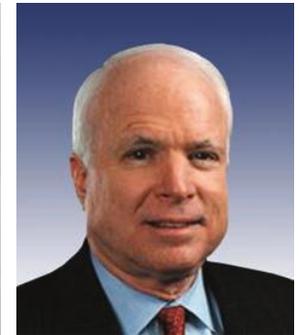
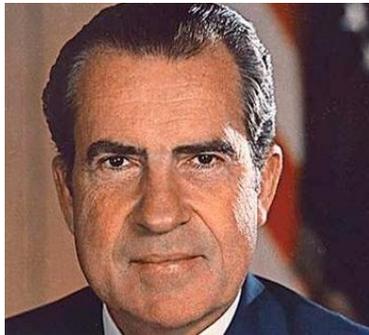
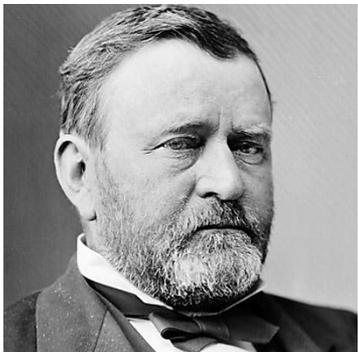
Deliberate vagueness and underspecification

- Insufficient evidence
- Inability to reach consensus
- Legal concerns (standard of care)
- Economic reasons

Authors should be transparent about
the reason for AVUL

VERBS

- Active voice
 - Passive masks the actor
 - “Mistakes were made...”
- Transitive: act upon an object
- Appropriate level of intended obligation is conveyed



The Perfect Non-apology Apology

Statement of fact is NOT a recommendation

- Adjuvant hormone therapy for locally advanced breast cancer results in improved survival in the long term.
- Clinicians should prescribe adjuvant hormone therapy for locally advanced breast cancer (when/unless?)...

Where's the Recommendation?

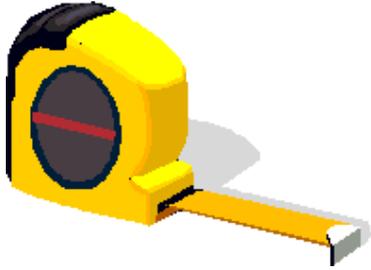
- Targeting high-risk patients at the ED visit for asthma education has been explored in two RCTs (Bolton et al. 1991; Cote et al. 2001) and in two observational studies (Kelso et al. 1995; Kelso et al. 1996). In one RCT, limited education in the ED in inhaler technique and use of a self-management action plan was compared to a comprehensive, structured educational program and usual care (Cote et al. 2001). ED revisits were not different among the groups in the first 6 months after the intervention, but declined significantly more in the structured education group by 12 months. However, reinforcement of self-management education was provided at the 6-month point only to the structured education group. In a second RCT, Bolton et al. (1991) provided three asthma education sessions to patients following a visit to the ED. Although there was significant attrition from attendance at sessions, followup was completed with 76 percent of the study sample, and adjusting for baseline differences, the intervention group had fewer ED visits than controls at 12-month followup ($p = .06$). In a race-specific reanalysis of the Bolton et al. (1991) study data, Ford et al. (1997) found that African-American and Caucasian patients learned equally well. It is the opinion of the Expert Panel that clinicians should offer brief and focused asthma education at the time of discharge from the ED (Evidence D). At the time of discharge from the ED all patients should receive a written asthma action plan with explanation to the patient and family of how to use it (Evidence B).

Hidden Recommendations

- Targeting high-risk patients at the ED visit for asthma education has been explored in two RCTs (Bolton et al. 1991; Cote et al. 2001) and in two observational studies (Kelso et al. 1995; Kelso et al. 1996). In one RCT, limited education in the ED in inhaler technique and use of a self-management action plan was compared to a comprehensive, structured educational program and usual care (Cote et al. 2001). ED revisits were not different among the groups in the first 6 months after the intervention, but declined significantly more in the structured education group by 12 months. However, reinforcement of self-management education was provided at the 6-month point only to the structured education group. In a second RCT, Bolton et al. (1991) provided three asthma education sessions to patients following a visit to the ED. Although there was significant attrition from attendance at sessions, followup was completed with 76 percent of the study sample, and adjusting for baseline differences, the intervention group had fewer ED visits than controls at 12-month followup ($p = .06$). In a race-specific reanalysis of the Bolton et al. (1991) study data, Ford et al. (1997) found that African-American and Caucasian patients learned equally well. **It is the opinion of the Expert Panel that clinicians should offer brief and focused asthma education at the time of discharge from the ED (Evidence D). At the time of discharge from the ED all patients should receive a written asthma action plan with explanation to the patient and family of how to use it (Evidence B).**

The Dreaded “Consider”

- When monitoring CD4+ counts frequently (e.g., every 1 to 3 months) is not possible, initiating chemoprophylaxis at a CD4+ count of >200, but <250 cells/ μ l, also **should be considered (BII)**.
- Twice-weekly continuation-phase therapy **may be considered** in patients with CD4+ counts ≥ 100 cells/ μ l (**CIII**).
- Routine laboratory monitoring during treatment, even when baseline laboratory abnormalities are not present, **could be considered**.



Measurement

- If you can't measure it, you can't manage it.

Peter Drucker

- If you don't measure it, you can't improve it.

Guideline Problems (Slide 2)

1. Development process is not transparent
2. Conflict of interest
3. Developer teams are insufficiently multidisciplinary and balanced
4. Existing knowledge is not thoroughly reviewed
5. Articulation of recommendations is not clear
6. Evidence foundations and strength of recommendations are not explicitly recorded
7. Reviewer input is not broad-based
8. New knowledge is not incorporated

Many guideline authors conflate two concepts that should remain distinct

- Quality of evidence
- Recommendation strength

Evidence Quality: Individual and Aggregate

- “Extent to which all aspects of a study’s design and conduct can be shown to protect against bias and inferential error”
- Study type: RCT > Observational study > Expert opinion/first principles
- Study quality and relevance
- Aggregate of studies
 - consistency of results
 - magnitude of effect

Evidence Quality

- An indication of the authors ‘ **confidence** in their appraisal of benefits and harms
- Based on an analysis of the validity, consistency, and applicability of the evidence supporting a recommendation

Evidence Quality Appraisal

- AHRQ found 82 different evidence quality rating schemes in 2002
 - Strong, moderate, no evidence
 - Grade A, Grade B, Grade C
 - IA, IB, IIA, IIB, III, IV
 - I, II, III-1, III-2, III-3, IV
 - 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B, 4, 5
 - 1++, 1+, 1-, 2++, 2+, 2-, 3, 4 ...

Guideline **authors** are committed to appraising the **quality** of scientific evidence

- As they should!
- But they regularly fall short in helping users understand how to use the information
 - “Users should almost always adhere to a recommendation with Grade A evidence, no?”
 - Possible to have Grade A evidence of effectiveness AND Grade A evidence of harm for the same intervention
- **Implementers** are rarely interested in evidence quality per se
- Implementers need to understand experts’ assessment of ***strength of recommendation***

Strength of Recommendation requires a **subjective judgment** about the balance between benefits and harms as well as an **objective appraisal** of evidence quality

When we are confident that
benefits exceed harms (and costs)...

Benefits



Harms,
Costs

We can make a **(strong) recommendation**
to...

When we are confident that anticipated harms (and costs) exceed benefits...

Benefits



Harms,
Costs

We can make a **(strong) recommendation not to...**

When anticipated benefits are **balanced** by harms and costs...

Benefits



Harms,
Costs

We can make a **weak recommendation** to...

When evidence quality is low...

Benefits



Harms,
Costs

We can make a **weak recommendation**
to...

Recommendations based on evidence quality

AGGREGATE EVIDENCE QUALITY	BENEFIT OR HARM PREDOMINATES	BENEFIT AND HARM BALANCED
<p>GRADE A Intervention: Well designed and conducted trials, meta-analyses on applicable populations Diagnosis: Independent gold standard studies of applicable populations</p>	<p>STRONG RECOMMENDATION</p>	<p>WEAK RECOMMENDATION BASED ON BALANCED BENEFITS AND HARMS</p>
<p>GRADE B Trials or diagnostic studies with minor limitations; consistent findings from multiple observational studies</p>	<p>RECOMMENDATION</p>	<p>WEAK RECOMMENDATION BASED ON BALANCED BENEFITS AND HARMS</p>
<p>GRADE C Single observational study or multiple studies with inconsistent findings or major limitations.</p>	<p>RECOMMENDATION</p>	<p>WEAK RECOMMENDATION BASED ON BALANCED BENEFITS AND HARMS</p>
<p>GRADE D Expert opinion, case reports, reasoning from first principles</p>	<p>WEAK RECOMMENDATION BASED ON LOW EVIDENCE QUALITY</p>	<p>NO RECOMMENDATION</p>
<p>GRADE X Exceptional situations where validating studies cannot be performed and there is a clear preponderance of benefit or harm</p>	<p>STRONG RECOMMENDATION</p> <p>RECOMMENDATION</p>	<p>NO RECOMMENDATION</p>

Guideline Users and Strong Recommendations

- Users should follow such guidance unless a clear and compelling rationale for acting in a contrary manner is present
- Optimal source for performance measures

Guideline Users and *Recommendations*

- Users generally should follow such guidance but also remain alert to new information and be sensitive to patient preferences.

Guideline Users and *Weak Recommendations*

- Patient preference should have a substantial role in influencing clinical decision making
- Hard to hold users accountable (performance measures)

COGS

- Conference on Guideline Standardization
 - Held at Yale 2003
 - Participants: developers, disseminators, and implementers
 - Delphi
 - Defined a checklist of guideline elements that should be described to assure validity and usability

Publication

| ACADEMIA AND CLINIC

Standardized Reporting of Clinical Practice Guidelines: A Proposal from the Conference on Guideline Standardization*

Richard N. Shiffman, MD, MCIS; Paul Shekelle, MD, PhD; J. Marc Overhage, MD, PhD; Jean Slutsky, PA, MSPH; Jeremy Grimshaw, MB, ChB, PhD; and Aniruddha M. Deshpande, MD

- Overview material
- Focus
- Goal
- Users / Setting
- Target population
- Developer
- Funding source / sponsor
- Evidence collection
- Recommendation grading criteria
- Method for synthesizing evidence
- Pre-release review
- Update plan
- Definitions
- Recommendations and rationale
- Potential benefits and harms
- Patient preferences
- Algorithm
- Implementation consideration

Guideline Quality Appraisal

Appraisal of **G**uidelines for **R**esearch & **E**valuation II



AGREE II

Instrument

[Agree Trust Link](#)

AGREE II Domains

- I. Scope and Purpose
- II. Stakeholder Involvement
- III. Rigour of Development
- IV. Clarity of Presentation
- V. Applicability
- VI. Editorial Independence

AGREE Instrument

STAKEHOLDER INVOLVEMENT

4. The guideline development group includes individuals from all the relevant professional groups.

Strongly Agree

4

3

2

1

Strongly Disagree

Comments

4.

This item refers to the professionals who were involved at some stage of the development process. This may include members of the steering group, the research team involved in selecting and reviewing/rating the evidence and individuals involved in formulating the final recommendations. This item excludes individuals who have externally reviewed the guideline (see Item 13). Information about the composition, discipline and relevant expertise of the guideline development group should be provided.



GuideLine Implementability Appraisal

*BMC Medical Informatics and Decision Making
2005*

Guideline Quality Appraisal Based on IOM Standards?

- National Guidelines Clearinghouse will tighten requirements
- Announcement June 2013

In this session...(Slide 2)

- What makes a guideline a Public Health Guideline?
- IOM-Identified Deficiencies in Guidelines
 - Articulation
 - Evidence Quality and Recommendation Strength
- Instruments for Appraising Guideline Quality