

The Minimum Quality Criteria Set (MQCS) for Critical Appraisal: Advancing the Science of Quality Improvement

Presenting author: Lisa Rubenstein, MD, VA, RAND, UCLA

Susanne Hempel, PhD, RAND

Jodi Liu, MS, RAND

Margie Danz, MD, VA, RAND

Robbie Foy, MD, University of Leeds

Yee-Wei Lim, MD, National University of Singapore

Aneesa Motala, BA, RAND

Paul Shekelle, MD, PhD, VA, RAND



MQCS Is A Tested Critical Appraisal Tool That

- Is used to assess the *quality of literature* on the success, effectiveness or impacts of *quality improvement interventions* (QIIs)
 - Is accompanied by a manual

What Is QII Critical Appraisal?

- Retrospective quality review of QII articles
- Reporting guidelines such as SQUIRE (Standards for Quality Improvement Reporting Excellence <http://www.squire-statement.org>) = critical appraisal tools. Guidelines
 - Can be comprehensive
 - Need not be interpreted with high reliability
 - Can be aspirational, because prospective
 - Directed toward improving future work
 - Focused on improving article content

MQCS Tool



Minimum Quality Criteria Set (MQCS) – Version 1.0

ID: _____ Author, year: _____ Reviewer: _____

Intervention: _____ Outcome: _____

Domain	Minimum standard	Score
<p><u>1. Organizational Motivation: Organizational problem, reason, or motivation for the intervention</u></p> <ul style="list-style-type: none"> ◆ Consider quality of care problems; organizational problems; regulations, legal constraints, and external financial incentives at the target organization; or organizational motivation. 	<p>Names or describes at least one motivation for the organization’s participation in the intervention</p>	<p>Not met Met</p>
<p><u>2. Intervention Rationale: Rationale linking the intervention to its expected effects</u></p> <ul style="list-style-type: none"> ◆ Consider citations of theories, logic models, or existing empirical evidence that links the intervention to its expected effects. 	<p>Names or describes a rationale linking at least one central intervention component to intended effects</p>	<p>Not met Met</p>
<p><u>3. Intervention Description: Change in organizational or provider behavior</u></p> <ul style="list-style-type: none"> ◆ Consider the presented details that describe the change in the delivery of care, provider behavior, or structure of the organization needed to replicate the evaluated intervention including the involved key personnel. 	<p>Describes at least one specific change in detail including the personnel executing the intervention</p>	<p>Not met Met</p>

This Talk

- Describes the *development* of the MQCS (the Minimum Quality Criteria Set) for quality improvement intervention (QII) publications
- Describes the *application* of the MQCS to the QII literature
 - Article identification and screening
 - MQCS psychometric properties
 - Strengths and weaknesses in current QII publications

QII Evidence Review Poses Challenges

- Identifying articles without bias
- Location of information in articles
- Widely varying language for describing the QII process
- Heterogeneity
 - Contexts
 - Article goals
 - Phases
 - Evaluation methods

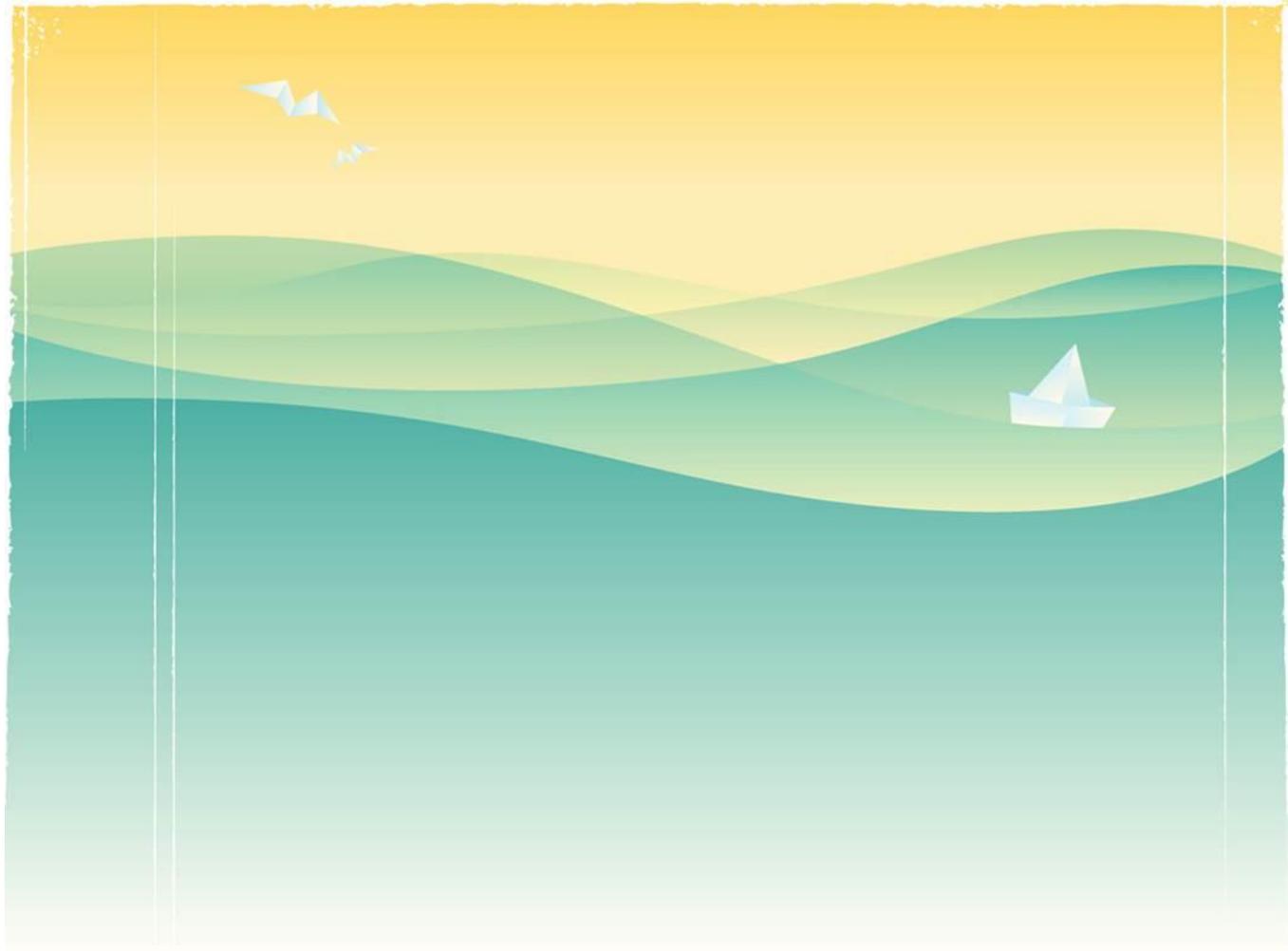
And Yet....

- Without reliable, unbiased evidence synthesis, we cannot learn across QII efforts
- Lack of reliable critical appraisal instruments for identifying high quality QII articles is an important barrier to evidence synthesis

Poll Question:

Have You (click on all that apply)

- *Participated in* a quality improvement intervention?
- *Written up* a quality improvement intervention?
- *Reviewed* quality improvement intervention literature?



MQCS *Development* Objectives

- Develop unbiased, broad identification of QII articles
- Develop a critical appraisal tool that is
 - Reliable and valid
 - Identifies high quality articles for evidence review without excluding important work
 - Inclusive relative to topics, countries of origin, methodologies
 - Applicable to the broad field of already published studies
 - Feasible (short, focused on key domains)

Methods: MQCS Development

- Step 1: Define QIIs for article identification
- Step 2: Review existing criteria sets
- Step 3: Convene an expert panel
- Step 4: Iteratively develop & test items (criteria)

9 Month Panel Process

Study Team Process

Survey & Telephone Consensus Panel

Telephone Panel

Survey & Face to Face Consensus Panel

QII Definition

Electronic Search, Screener, Strategy & Preliminary Criteria

Quality Criteria

Existing Criteria Review

SQUIRE Domains Survey

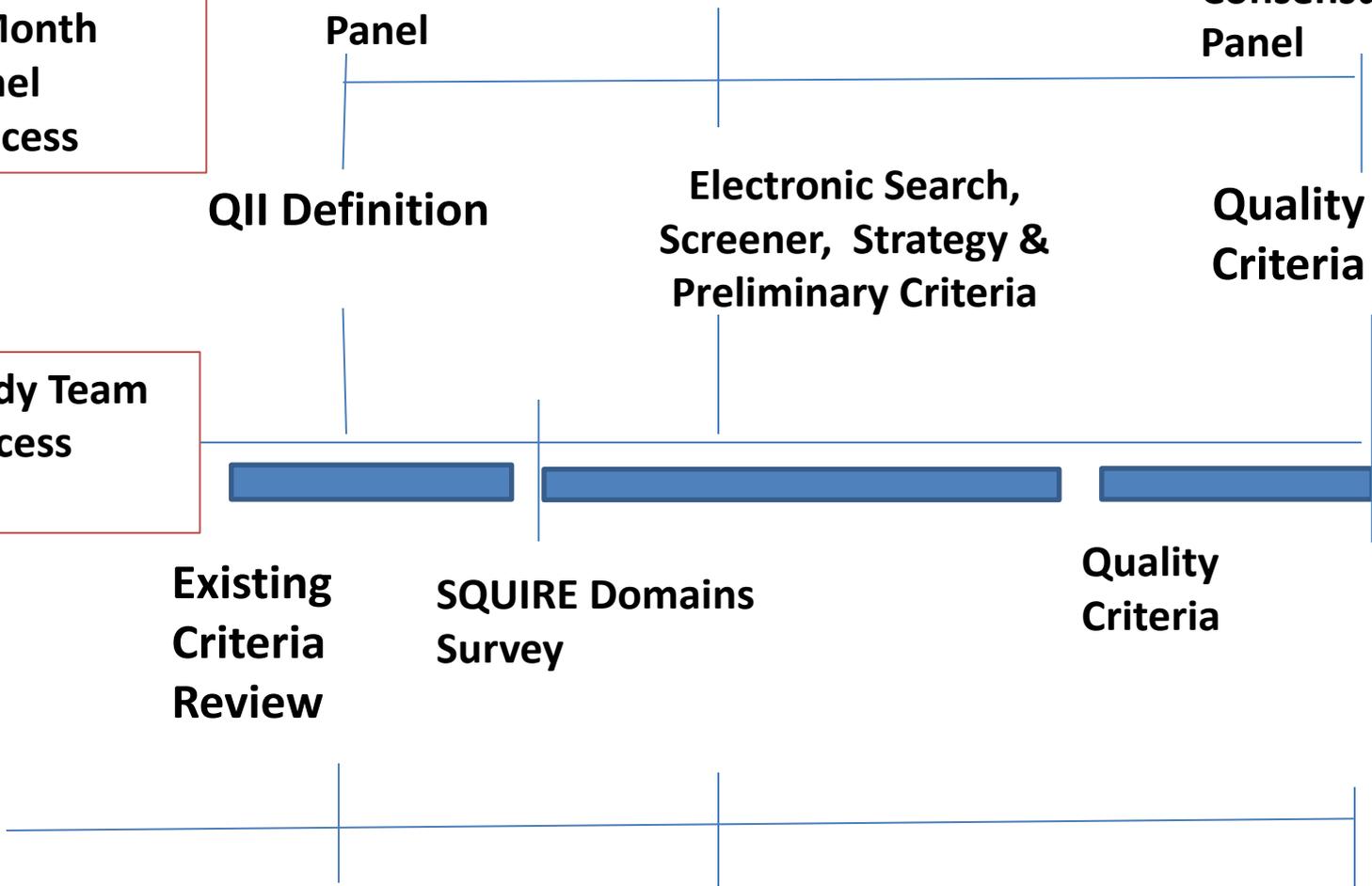
Quality Criteria

Iterative Development of MQCS Tool

Nov 2008

March 2009

July 2009



Step 1: Define QIs For Article Identification

- QI literature is a subset of 1) all articles important for QI itself and 2) all articles relevant to interventions intended to affect quality
 - Not discussed here, but referenced ([Rubenstein LV. Quality and Safety in Healthcare, 2008](#))
- Electronic search capability is imperative
 - Not discussed here, but referenced ([Hempel S. Implement Sci. 2011](#))

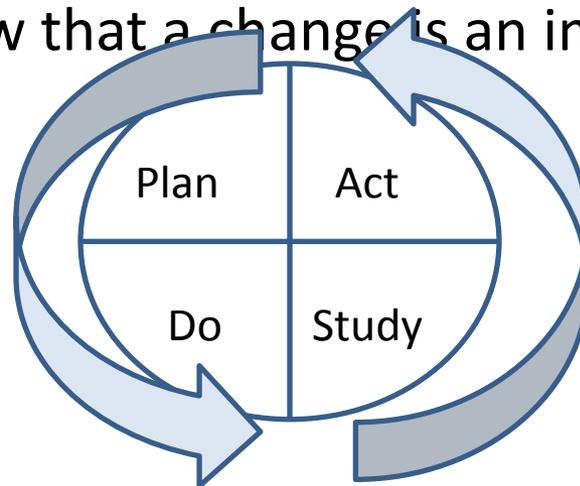
QII Definition for Article Identification

- First fundamental definitional difference from classical intervention research (e.g., trials) is that *QIIs involve and reflect organizations*
 - Are not about holding the hand of an individual provider/staff/patient, but about influencing a group's ongoing care and outcomes within an organization

QII Definition

- Second fundamental difference from classical intervention research is a non-linear model
 - E.g., Model for Improvement
 - What are we trying to accomplish?
 - What changes can we make that will result in an improvement?
 - How will we know that a change is an improvement?

IHI, Associates in
Process
Improvement



What is a QII? (Definition Suitable for Development of Article Title & Abstract Screening)

- “An effort to change/improve the clinical structure, process and/or outcomes of care by means of an organizational or structural change”

M S Danz. Qual Saf Health Care 2010

Even More Specifically...

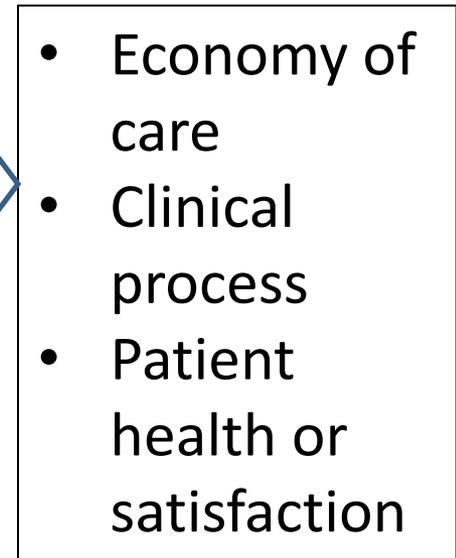
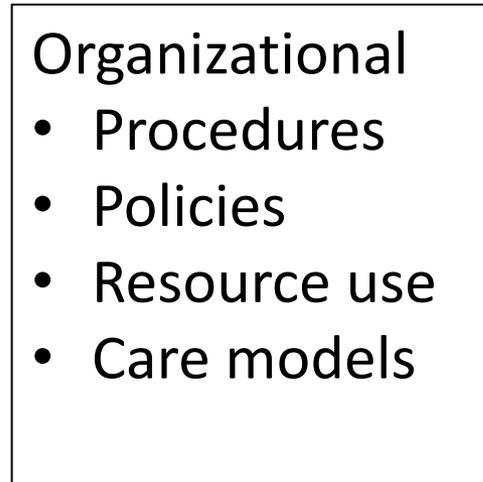
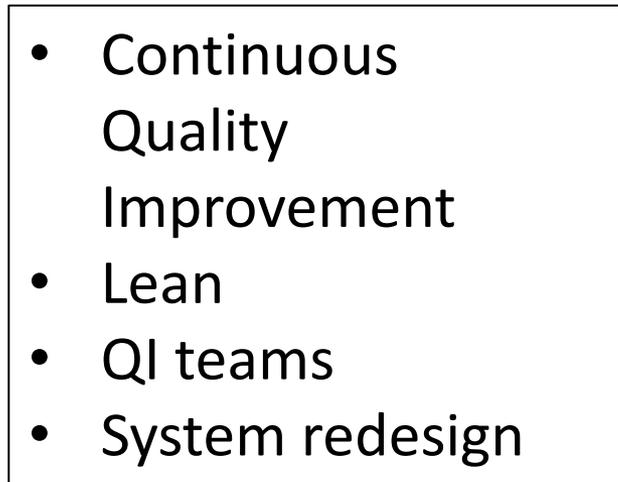
An improvement *effort*, or initiative
e.g.



Effort is designed
to make an
*organizational or
structural change*,
e.g.



To affect the
organization or
organizational
unit's *structure*,
process, or
outcomes of care,
e.g.



Step 2: Identify Existing Review Criteria and Guidelines (Study Team)

- Identify QII quality domains and items from key sources including
 - SQUIRE (Standards for Quality Improvement Reporting Excellence): “The SQUIRE Guidelines help authors write excellent, usable articles about quality improvement in healthcare” (<http://squire-statement.org/>)
 - MRC (Medical Research Council) guidance ([BMJ 2008;337:a1655](#))
- Identify evaluation design criteria (e.g., <http://www.equator-network.org.>)
- Reviewed > 20 additional sets of criteria

Step 3: Expert Panel

- ***Panel Members***: Frank **Davidoff**, Institute for Healthcare Improvement; Martin **Eccles**, Newcastle University Institute of Health and Society; Robert **Lloyd**, Institute for Healthcare Improvement; Vin **McLoughlin**, The Health Foundation; Brian **Mittman**, Department of Veterans Affairs; Shirley **Moore**, Case Western Reserve University; Greg **Ogrinc**, Dartmouth Institute; Drummond **Rennie**, University of California, San Francisco; Susanne **Salem-Schatz**, Independent Consultant; David P. **Stevens**, Dartmouth Institute; Edward H. **Wagner**, Group Health Center for Health Studies
- Funders and participating sponsors: RWJ (Lori **Melichar**); AHRQ (Denise **Dougherty**, Judith **Sangl**, Laurence **Kleinman**); Veterans Affairs (David **Atkins**)
- We benefited greatly from their input, but any errors are ours

Panel Determined the MQCS Scope

- Per panel, MQCS focuses on QII-specific criteria
 - Panel judged the heterogeneity of QII designs made evaluation design criteria inadvisable
 - Systematic reviewers can apply existing relevant evaluation design criteria in addition to MQCS if relevant
- All but two of the final MQCS criteria are based on Panel ratings
 - Two non-Panel-identified items focused on the quality of descriptions of evaluation methods—data sources for outcomes, and study design/comparators

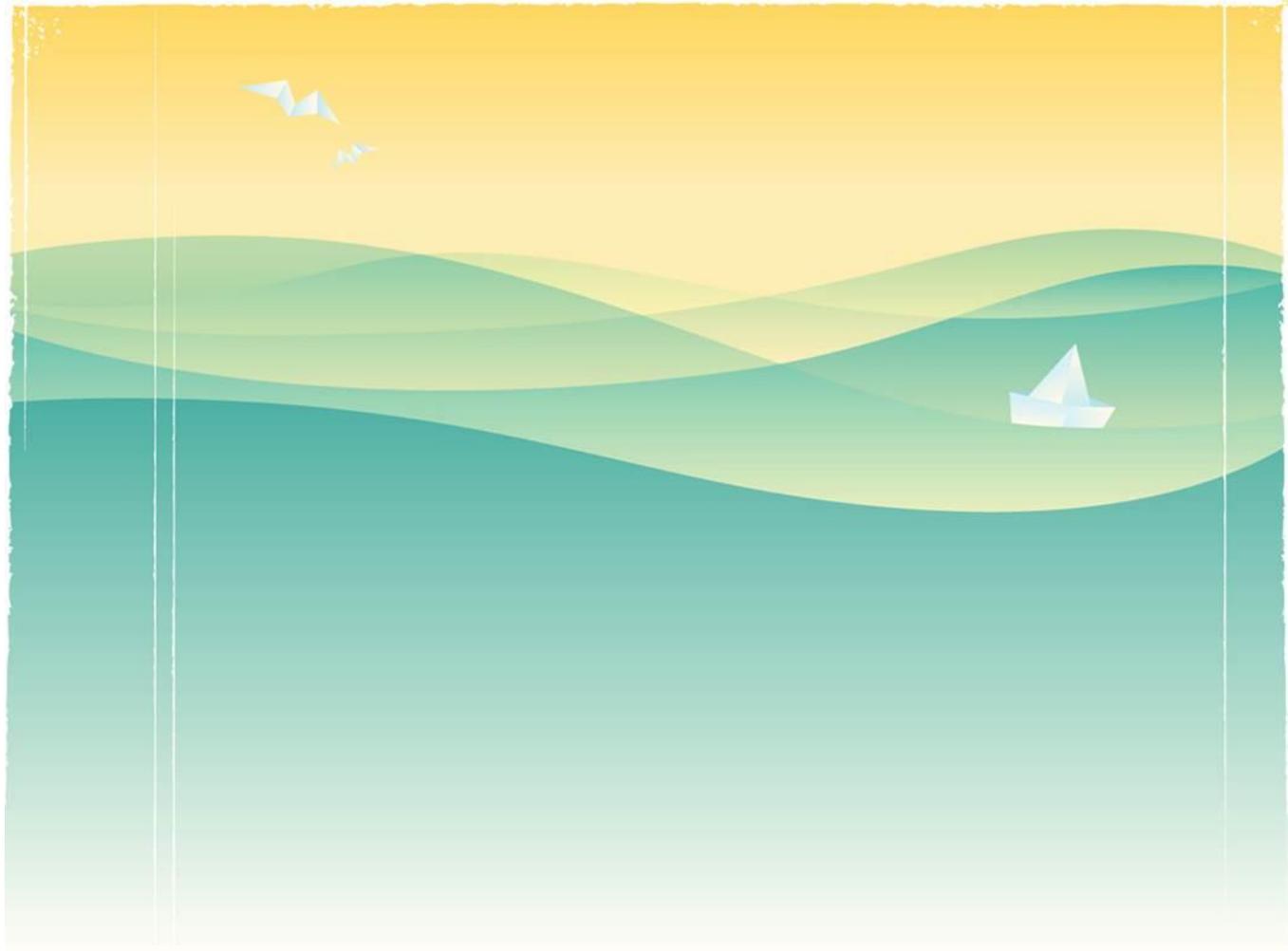
Step 4: Iterative MQCS Tool

Development

- Post-Panel abstraction form had 16 domains, one item per domain, yes/no answers, and scoring criteria
 - Two reviewers independently reviewed batches of nine articles and reconciled ratings
 - Adjustments made to items
- Kappas, % agreement, correlations calculated using final MQCS
- Developed accompanying manual (“*What to consider*”)

Example from the MQCS FORM

MQCS ITEM	Minimum standard	Score
<p>2. <u>Intervention Rationale</u>: Rationale linking the intervention to its expected effects</p> <p>Consider citations of theories, logic models, or existing empirical evidence that links the intervention to its expected effects.</p>	Names or describes a rationale linking at least one central intervention component to intended effects	Not met/ Met



MQCS *Application* Objectives

- Using a diverse set of electronically searched, screened, and reviewed QI articles
 - Assess MQCS psychometric properties
 - Assess areas of weakness in QI publications (adherence to each of 16 MQCS criteria)

Identification of QII Literature

- We applied MQCS to 54 articles resulting from electronic search & hand/machine screening for QII
 - Validated electronic search strategy yielded 9427 articles
 - 1600 randomly-selected were hand title/abstract reviewed; 7827 were machine-screened re: *did the article report empirical data on a QII (248 screened in)*
 - 24 QII exemplar articles from panelists were added
 - The resulting 272 title/abstract screened articles underwent full article screening by two reviewers

Resulting Articles

- Covered diverse topics: E.g., restructuring of teams, audit and feedback, falls, tuberculosis detection
- Included international articles from developed and developing countries

Article Yield

- Most articles in the QI field do not include empirical data on QII health-related outcomes
 - About 3% (272) of electronically searched QI articles had any empirical data on a QII
 - About 20% of the 272 articles hand screened in for qualitative or quantitative data focused on a health-related process or outcome
 - “Patient (or caregiver), provider behavior, or process of care health outcome”

Is the dearth of QII empirical evaluation publications primarily due to

Choose one:

- A scarcity of QI projects that gather empirical data on whether there was improvement
- Difficulty writing up and publishing empirical QII evaluations
- About equally due to both

Assess MQCS Tool Reliability

- Median inter-rater agreement across all items and articles was $\kappa = 0.57$; reviewer agreement = 83%
- Two low outliers
 - Spread ($\kappa = 0.13$; Agreement 67%)
 - Adherence/Fidelity ($\kappa = 0.9$; Agreement 56%)
- Inter-item correlations all below 0.60, mean 0.19 (indicating conceptual independence)

Percent of Articles Meeting Each MQCS Criterion

MQCS Item	% of Articles That “MET” (N = 54)
Organizational motivation for the QII	64%
Intervention rationale	67%
Intervention description	93%
Organizational characteristics	89%
Implementation activities	92%
Study /evaluation design	44%
Information about comparators	67%
Data sources for outcome	67%

Percent of Articles Meeting Each MQCS Criterion

MQCS Item	% of Articles That "MET" (N = 54)
Organizational motivation for the QII	64%
Intervention rationale	67%
Intervention description	93%
Organizational characteristics	89%
Implementation activities	92%
Study /evaluation design	44%
Information about comparators	67%
Data sources for outcome	67%

MQCS Item	% of Articles That “MET” (n = 54)
Timing of the intervention and evaluation	56%
Adherence/fidelity to the intervention	47%
Patient health-related outcomes	58%
Organizational readiness/barriers and facilitators	84%
Penetration/reach	85%
Sustainability of the intervention	83%
Ability to be spread or replicated	89%
Limitations description	64%

MQCS Item	% of Articles That “MET” (n = 54)
Timing of the intervention and evaluation	56%
Adherence/fidelity to the intervention	47%
Patient health-related outcomes	58%
Organizational readiness/barriers and facilitators	84%
Penetration/reach	85%
Sustainability of the intervention	83%
Ability to be spread or replicated	89%
Limitations description	64%

Key Areas for Improvement: Lowest Group (<60 % of articles “Met” the Criterion)

- Need systematic attention from the field
 - Study design description (44%)
 - Description of timing of the intervention and evaluation (56%)
 - Description of intervention adherence/fidelity (47%)
 - Description of patient/non-professional caregiver health outcomes (58%)

Would more clarity on definitions of the terms *adherence/fidelity* and/or of *spread* (the low outliers for reliability) be helpful to the QI field?

Choose one

- Yes or probably yes
- No or probably no

MQCS ITEM

Minimum Standard

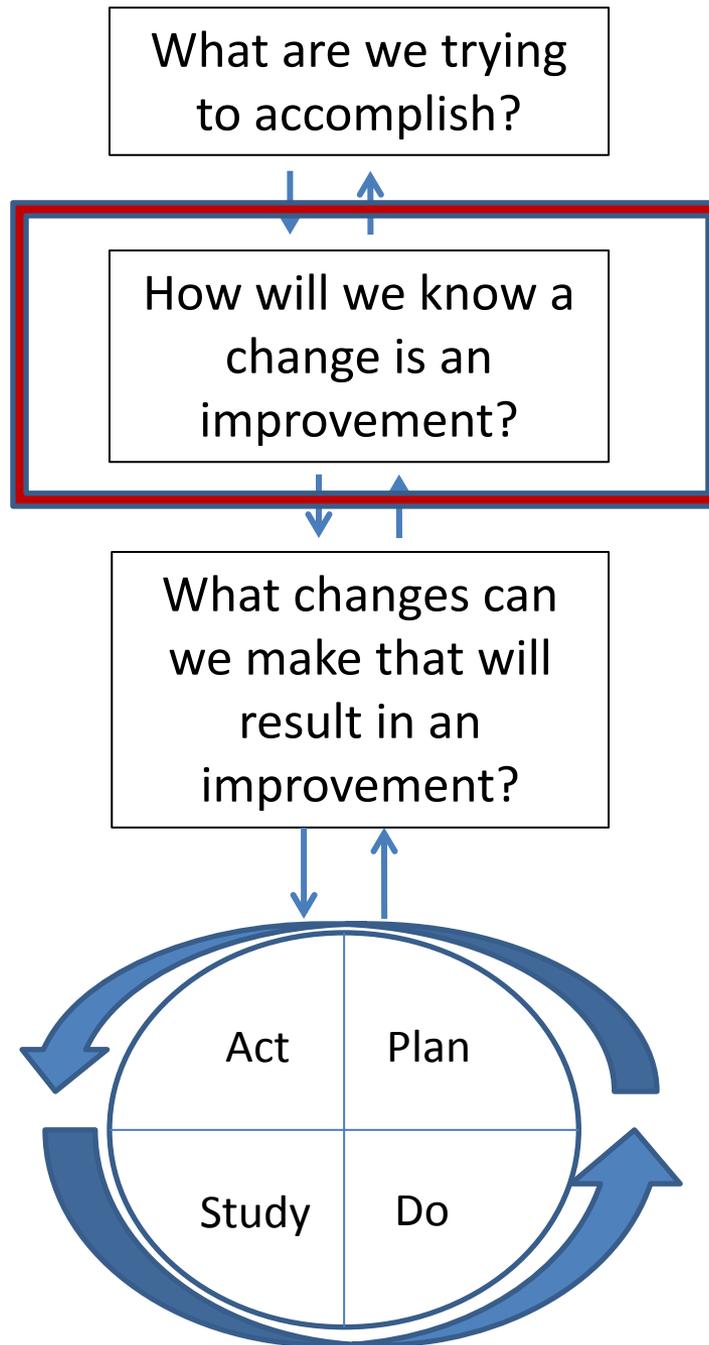
Study Design: Study design and comparator

Names the study design

Timing: Timing of intervention and evaluation

Describes the timing of the intervention and evaluation components [condensed from the original]

MQCS ITEM	Minimum standard
<u>Adherence / Fidelity</u>	Reports fidelity information for at least one intervention component, or describes evidence of adherence or a mechanism ensuring compliance to the intervention [condensed from the original]
<u>Health Outcomes:</u> Patient or non-professional caregiver health-related outcomes	Reports data on at least one health-related outcome



The Institute for
Healthcare
Improvement
(IHI)
(<http://www.ihio.org/resources/Pages/HowtoImprove/default.aspx>) and
Associates in
Process
Improvement
(<http://www.apiweb.org/>)

**Model for
Improvement**

Conclusions: Feasibility and Psychometrics

- It is feasible to reliably identify and review QII literature
 - The MQCS had acceptable psychometric properties for critical appraisal, and can support systematic review of diverse QII evaluations
 - Greater consensus/definition is needed for review of information addressing lowest reliability items
 - *QII adherence/fidelity*
 - *QII spread*

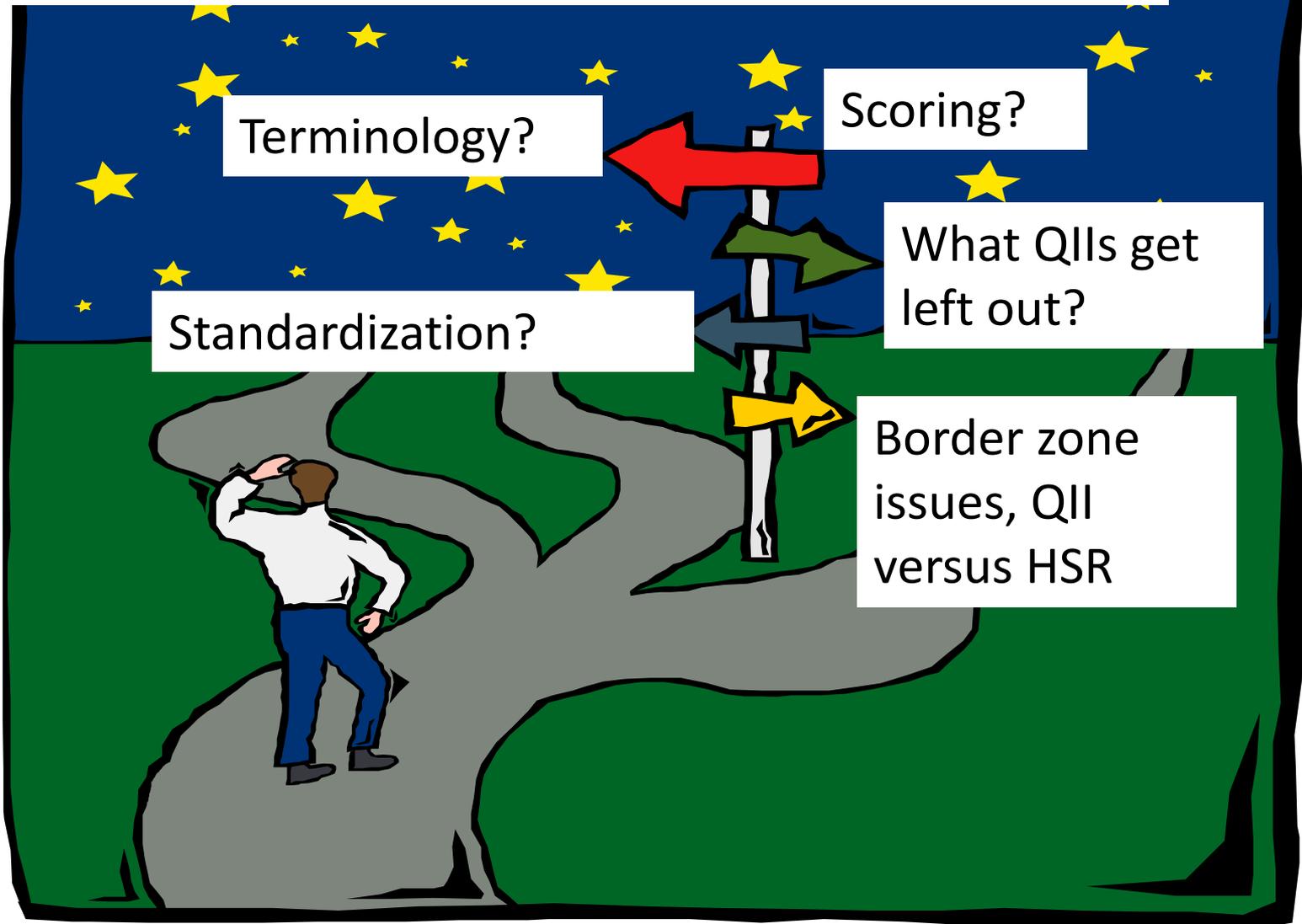
Conclusions: QII Articles

- QI practitioners, funders, and journals should focus on enabling QII publications that include any empirical evaluation data *and* a health-related outcome of some kind
- It would be helpful for evidence review if guidelines on placement of key information in QII articles, especially but not only the abstract, were refined and followed

Conclusions: QII article quality

- QII project designers and writers need to improve reporting on *study design* and the *timing of intervention and evaluation*
 - Quality was best for critical appraisal criteria focused on intervention description, and worst for criteria focused on evaluation

But...we're still on the path, not at the goal, for critical appraisal



Knowing is not
enough; we must
apply. Willing is not
enough; we must
do.

[Johann Wolfgang von Goethe](#)



Published Articles By Our Team on QI/Redesign Evidence Review Methods for QI/Patient Safety

- Rubenstein, L.V., Hempel, S., Farmer, M., Asch, S., Yano, E., Dougherty, D. & Shekelle, P. (2008). Finding order in heterogeneity: Types of quality improvement intervention publications. *Quality and Safety in Healthcare*,17(6), 403-8.
- Danz MS, Rubenstein LV, Hempel S, Foy R, Suttorp M, Farmer MM, Shekelle PG. Identifying quality improvement intervention evaluations: is consensus achievable? *Qual Saf Health Care*. Aug 2010;19(4):279-283.
- Foy R, Hempel S, Rubenstein L, Suttorp M, Seelig M, Shanman R, Shekelle P: Metaanalysis: Effect of Interactive Communication Between Collaborating Primary Care Physicians and Specialists. *Ann Intern Med* 2010, 152:247-258
- Hempel S, Rubenstein LV, Shanman RM, Foy R, Golder S, Danz M, Shekelle PG. Identifying quality improvement intervention publications - A comparison of electronic search strategies. *Implement Sci*. 2011;6:85.
- Soban LM, Hempel S, Munjas BA, Miles J, Rubenstein LV. Preventing pressure ulcers in hospitals: A systematic review of nurse-focused quality improvement interventions. *Jt Comm J Qual Patient Saf*. Jun 2011;37(6):245-252.
- Khodyakov D, Hempel S, Rubenstein L, Shekelle P, Foy R, Salem-Schatz S, O'Neill S, Danz M, Dalal S. Conducting Online Expert Panels: A Feasibility and Experimental Replicability Study. *BMC Med Res Methodol*. 2011 Dec 23;11(1):174.
- O'Neill SM, Hempel S, Lim YW, Danz MS, Foy R, Suttorp MJ, Shekelle PG, Rubenstein LV. Identifying continuous quality improvement publications: what makes an improvement intervention 'CQI'? *BMJ Qual Saf*. Dec 2011;20(12):1011-1019.
- Foy R, Ovretveit J, Shekelle PG, Pronovost PJ, Taylor SL, Dy S, Hempel S, McDonald KM, Rubenstein LV, Wachter RM. The role of theory in research to develop and evaluate the implementation of patient safety practices. *BMJ Qual Saf*. 2011 May;20(5):453-9.

Articles published (continued)

- Shekelle PG, Pronovost PJ, Wachter RM, Taylor SL, Dy SM, Foy R, Hempel S, McDonald KM, Ovretveit J, Rubenstein LV, Adams AS, Angood PB, Bates DW, Bickman L, Carayon P, Donaldson L, Duan N, Farley DO, Greenhalgh T, Haughom J, Lake ET, Lilford R, Lohr KN, Meyer GS, Miller MR, Neuhauser DV, Ryan G, Saint S, Shojanian KG, Shortell SM, Stevens DP, Walshe K. Advancing the science of patient safety. *Ann Intern Med*. 2011 May 17;154(10):693-6.
- Ovretveit JC, Shekelle PG, Dy SM, McDonald KM, Hempel S, Pronovost P, Rubenstein L, Taylor SL, Foy R, Wachter RM. How does context affect interventions to improve patient safety? An assessment of evidence from studies of five patient safety practices and proposals for research. *BMJ Qual Saf*. 2011 Jul;20(7):604-10.
- Taylor SL, Dy S, Foy R, Hempel S, McDonald KM, Ovretveit J, Pronovost PJ, Rubenstein LV, Wachter RM, Shekelle PG. What context features might be important determinants of the effectiveness of patient safety practice interventions? *BMJ Qual Saf*. 2011 Jul;20(7):611-7.
- Dy SM, Taylor SL, Carr LH, Foy R, Pronovost PJ, Ovretveit J, Wachter RM, Rubenstein LV, Hempel S, McDonald KM, Shekelle PG. A framework for classifying patient safety practices: results from an expert consensus process. *BMJ Qual Saf*. 2011 Jul;20(7):618-24.
- Danz MS, Hempel S, Lim YW, Shanman R, Motala A, Stockdale S, Shekelle P, Rubenstein L. Incorporating evidence review into quality improvement: meeting the needs of innovators. *BMJ Qual Saf*. Nov 2013;22(11):931-939.
- Rubenstein L, Khodyakov D, Hempel S, Danz M, Salem-Schatz S, Foy R, O'Neill S, Dalal S, Shekelle P. How can we recognize continuous quality improvement? *Int J Qual Health Care*. Dec 4 2013.
- Ovretveit J, Hempel S, Magnabosco J, Mittman B, Rubenstein L & Ganz D. Guidance for Research-Practice Partnerships (R-PPs) and Collaborative Research. *Journal of Health Organization and Management*. 2014. 28(1).115-126.
- Submitted : Hempel¹, Shekelle^{1,2}, Liu¹, Danz^{1,2}, Foy³, Lim⁴, Motala¹, Rubenstein- Critical Appraisal of Quality Improvement Intervention Publications: The Minimum Quality Criteria Set (MQCS)

Search Strategy for QI Publications

- No strategy was perfect
 - Tested against 3 expert article sets
- Best: ‘quality’ AND ‘improv*’ AND ‘intervention*’ yielded 13,572 articles
 - 62% (Cochrane, 26 articles); 24% (SQUIRE, 29); 44% AHRQ, 25) (tested on 2005-2007 studies)
- Yielded 13,572 articles

Types of QI Articles that Are Important (AHRQ expert-suggested article set)

- (1) Empirical articles on development and testing of QIIs
 - (1a) development of QIIs
 - 1(b) history, documentation, or description of QIIs
 - (1c) success, effectiveness or impact of QIIs
- (2) QI stories, theories, and frameworks
- (3) QI literature syntheses or meta-analyses
- (4) development and testing of QI tools

(Rubenstein, L.V. et al. (2008). *Quality and Safety in Healthcare*)

Research Team Developed and Panel Reviewed Screening Process

Title and Abstract Review

- Is the article potentially relevant to quality improvement?
- Is the article about a QII?
 - Two independent reviewers; included all articles endorsed by either one

Screeners: Whole Article

- *Setting*: Did the article report on **an intervention that was implemented in or by a healthcare delivery organization or organizational unit**?
- *Design*: Were qualitative or quantitative data on the **effectiveness, impacts, or success** of the intervention reported in the article?
- *Outcome*: Did the article report on **patient (or care-giver), provider behavior, or process of care health outcomes**?
- *Intervention*: Did the article suggest that the intervention **aimed to change how delivery of care is routinely structured** within a specific organization or organizational unit?

Key Areas for Article Improvement

- Middle Group (<85% but >60% “Met”):
These need work
 - Organizational motivation (64%)
 - Intervention rationale (67%)
 - Comparator description (67%)
 - Data source for outcomes (67%)
 - Limitations (64%)

MQCS ITEM

Minimum Standard

6. Study Design: Study design and comparator

- ◆ Consider the type of evaluation (e.g., post-only, pre-post, time series, parallel control group, randomized groups; same participants assessed multiple times or different samples) / how the authors evaluated whether the intervention worked

Names the study design

MQCS ITEM

9. Timing: Timing of intervention and evaluation

- ◆ Consider the clarity of the timeline of the intervention, e.g., when introduced, when fully implemented, when evaluated relative to the intervention implementation status, and a clear indication of whether baseline data (defined as before the intervention was introduced) was present.

Minimum Standard

Describes the timing of the intervention and evaluation to determine the presence of baseline data and the follow-up period after all intervention components were fully implemented

MQCS ITEM

Minimum standard

10. Adherence / Fidelity:

Consider reporting of compliance with the intervention for the duration of the study, fidelity data on intervention use, or described mechanisms that ensures compliance (e.g., provider reminder integrated in electronic health record that cannot be skipped).

Reports fidelity information for at least one intervention component, or describes evidence of adherence or a mechanism ensuring compliance to the intervention

MQCS ITEM

Minimum standard

11. Health Outcomes: Patient health-related outcomes

- ◆ Consider patient and non-professional care-giver health-related outcomes (including e.g., quality of life), but exclude satisfaction, provider-behavior (e.g., number of diagnostic tests ordered, knowledge) and process improvements.

Reports data on at least one health-related outcome