Methodology for Outcome Studies: How can the value of QI be demonstrated?

Ruth C. Carlos, MD MS
Professor and Assistant Chair of Clinical Research
University of Michigan

The Value Agenda

Measure outcomes that matter to patients

Porter and Lee, HBR 2013
QIBA
May 21, 2014
Defining and Measuring Value

Evidence base
• Limited for quantitative imaging
• Methodological challenges

Clinical outcomes
• Need to link use of QI to improved outcomes
• Outcomes dependent on quality and availability of therapy
• Attribution of outcomes

Care efficiency
• Time to initiation or change in therapy
• Treatment standardization
• Impact on throughput and resource use

The clinical scenario: diagnostic testing

- Patient visits clinician
- Clinician evaluates patient ➔ orders dx test(s)
- Radiologist supervises/interprets study ➔ issues report
- Clinician receives report, formulates management plan
- Clinician presents plan to patient
- Patient accepts or rejects plan

Patient health improves, worsens or remains unchanged
Challenges in evaluating QI

• QI test results are INTERMEDIATE outcomes
  – Health benefits and costs depend on disease severity, management plan, treatment effectiveness, patient compliance, etc...
  – Direct evidence linking results of QI to health benefits and costs can be difficult to obtain
Challenges in evaluating QI

• QI test results are INTERMEDIATE outcomes
  - Health benefits and costs depend on disease severity, management plan, treatment effectiveness, patient compliance, etc...
  - Direct evidence linking results of QI to health benefits and costs can be difficult to obtain
  - EX: Stage 4 GBM patient with extracranial metastases, QI demonstrates disease progression 6 weeks into a 12 week treatment course
    • Treatment continues in the absence of alternative therapies

“Moving target”

• Rapid technological advancement with continual development of new QI biomarkers
  - EX: DMIST ➔ TMIST
  - Longitudinal trials can be impractical or become obsolete quickly
Challenges in evaluating QI

• “Moving target”
  - Rapid technological advancement with continual development of new QI biomarkers
  - Changes in therapy or patient management affect outcome independent of “goodness” of QI marker

- Quantification of distribution and extent of emphysema predicts future cardiac disease
  - Smoking cessation can change outcome
  - Should outcome be smoking cessation or reduction of cardiac disease
Challenges to QI

- Multi-dimensionality of QI biomarkers
  - Presence/absence, location, length, intensity, type
  - Interaction / agreement with planar imaging characteristics
  - Measuring incremental value over qualitative imaging characteristics and/or other tests
  - Difficult to isolate the value of reporting quantitative markers over other imaging characteristics
Challenges to QI

• Clear cut-points rarely established or correlated to clinical outcome
  
  – Absolute value vs relative value (eg. hypointensity)
  – Unlikely vs highly likely (vs negative / positive)
  – Relate QI to probability of disease or predicted outcome
    • % fatty liver infiltration vs probability of NASH cirrhosis

• Disease is not binary
  
  – Tumor vs infection vs perfusion defect
  – Relate QI to probability of disease
Key considerations in technology assessment

• Is it true?
  – Technical capacity

• Is it meaningful?
  – Diagnostic accuracy

• Is it useful?
  – Clinical effectiveness
QI: Effect on Diagnostic Thinking

- May be reasonable proxy for effect on patient outcomes

- Sensitivity, specificity, NPV, PPV = standard measures of diagnostic accuracy

- Likelihood ratio = standard measure of potential effect on diagnostic thinking
  
  - \( \frac{\text{sensitivity}}{1 - \text{specificity}} \)

Effect on Diagnostic Thinking

- Likelihood ratio
  
  - How useful is a particular QI marker in a given clinical situation
  
  - \( LR > 10 \) or \(< 0.1 \) = large influence on diagnostic probability
  
  - \( LR \approx 1 \) = little / no diagnostic information
Effect on Diagnostic Thinking

- Likelihood ratio
  - How useful is a particular QI marker in a given clinical situation
  - LR >10 or <0.1 = large influence on diagnostic probability
  - LR ~1 = little / no diagnostic information

- LR can be used to determine post-test probability of disease
  - Requires knowledge of pretest probability of disease
  - Post-test odds of disease = pretest odds x LR

- Must know WHAT disease trying to detect (not always binary)
Effect on Diagnostic Thinking

- Aggregating clinical data, qualitative data and QI

- EX: Characterization of solitary pulmonary nodules
  - Comparing DCE-CT, DCE-MRI, FDG-PET, SPECT
  - Incorporating patient age, smoking history, history of prior malignancy, lesion size, other imaging characteristics, location

<table>
<thead>
<tr>
<th>Cigarettes</th>
<th>Other cancer</th>
<th>Spiculation</th>
<th>Upper Lobe</th>
<th>Diagnostic test</th>
<th>Positive LR</th>
<th>Post-test probability</th>
<th>25 years</th>
<th>55 years</th>
<th>75 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>No test</td>
<td>DCE CT</td>
<td>3.912</td>
<td>0.108</td>
<td>0.110</td>
<td>0.380</td>
<td>0.470</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>DCE MRI</td>
<td>4.565</td>
<td>0.124</td>
<td>0.361</td>
<td>0.662</td>
<td>0.256</td>
<td>0.548</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>FDG PET</td>
<td>5.440</td>
<td>0.144</td>
<td>0.402</td>
<td>0.700</td>
<td>0.291</td>
<td>0.591</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>SPECT</td>
<td>5.164</td>
<td>0.137</td>
<td>0.390</td>
<td>0.689</td>
<td>0.276</td>
<td>0.579</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>No test</td>
<td>DCE CT</td>
<td>3.912</td>
<td>0.227</td>
<td>0.510</td>
<td>0.389</td>
<td>0.697</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>DCE MRI</td>
<td>4.565</td>
<td>0.256</td>
<td>0.548</td>
<td>0.814</td>
<td>0.426</td>
<td>0.738</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>FDG PET</td>
<td>5.440</td>
<td>0.291</td>
<td>0.591</td>
<td>0.839</td>
<td>0.470</td>
<td>0.762</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>SPECT</td>
<td>5.164</td>
<td>0.276</td>
<td>0.579</td>
<td>0.832</td>
<td>0.446</td>
<td>0.752</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>No test</td>
<td>DCE CT</td>
<td>3.912</td>
<td>0.279</td>
<td>0.586</td>
<td>0.287</td>
<td>0.445</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>DCE MRI</td>
<td>4.565</td>
<td>0.313</td>
<td>0.603</td>
<td>0.848</td>
<td>0.483</td>
<td>0.775</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>FDG PET</td>
<td>5.440</td>
<td>0.380</td>
<td>0.645</td>
<td>0.869</td>
<td>0.527</td>
<td>0.804</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>SPECT</td>
<td>5.164</td>
<td>0.332</td>
<td>0.633</td>
<td>0.863</td>
<td>0.500</td>
<td>0.796</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>No test</td>
<td>DCE CT</td>
<td>3.912</td>
<td>0.170</td>
<td>0.430</td>
<td>0.230</td>
<td>0.430</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>DCE MRI</td>
<td>4.565</td>
<td>0.483</td>
<td>0.775</td>
<td>0.925</td>
<td>0.672</td>
<td>0.882</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>FDG PET</td>
<td>5.440</td>
<td>0.227</td>
<td>0.804</td>
<td>0.936</td>
<td>0.710</td>
<td>0.909</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>SPECT</td>
<td>5.164</td>
<td>0.500</td>
<td>0.796</td>
<td>0.933</td>
<td>0.674</td>
<td>0.904</td>
</tr>
</tbody>
</table>
QI: Effect on Therapeutic Planning

- Test should lead to change in patient management to demonstrate clinical effectiveness

- Challenging to compare management before and after testing

- Proxy measures for change in management
  - Change in intended treatment vs actual treatment
  - Clinician confidence in making correct therapeutic choice
  - Patient confidence in therapeutic choice
QI: Effect on Patient Outcomes

- Definition of patient outcomes
  - Life expectancy, event-free life expectancy, health status / functional status
  - Quality adjusted life years

- Generally requires RCT
  - Difficult to estimate effect of testing using observational design
  - Need to statistical methods to account for nonrandom events eg. propensity score adjustment for institutional / provider variables

Effect on Patient Outcomes

EX: Solitary pulmonary nodule

\[ P(\text{cancer})=0.10 \]
Effect on Patient Outcomes

EX: Solitary pulmonary nodule

P(cancer)=0.30

QI: Effect on Patient Outcomes

• Range of outcomes to be considered must be broader to adequately describe the patient experience
  
  – Patient-centered outcomes, patient-reported outcomes

QIBA
May 21, 2014
**Effect on Patient Outcomes**

EX: Solitary pulmonary nodule

![Diagram showing medical test result and patient outcomes](image)

\[ P(\text{cancer}) = 0.10 \]

**QI: Effect on Patient Outcomes**

- Effect of test result communication
  - Does the number make a difference?

- Alternative outcomes
  - Smoking cessation
  - Use of ancillary services (eg. among women with false positive mammography)
  - Adherence to screening (eg. among women with false positive mammography)
Effect on Patient Outcomes

EX: Breast cancer

P(Stage 1)=0.95

Effect on Patient Outcomes

- EX: Stage 1 breast cancer
  - Alternative patient-centered outcomes
    - Frequency of mastectomy instead of lumpectomy + radiation
    - Decision satisfaction with treatment choice
    - Cosmesis

QIBA
May 21, 2014
Potential clinical uses of QI and alternative RCT designs

- Test addition
- Triage
- Test replacement / substitution

Test addition: potential RCT designs

EX: Palliation vs resection of cholangioCA

Adapted from Bossuyt
**Test addition: potential RCT designs**

EX: Palliation vs resection of cholangioCA

![Diagram]

**Test as triage: potential RCT designs**

EX: 4D CT to determine need for arthroscopy

![Diagram]
Test addition: potential RCT designs

EX: 4D CT to determine need for arthroscopy

Test replacement: potential RCT designs

EX: 4D CT to determine need for arthroscopy
Test replacement: potential RCT designs

EX: Diffusion MRI to differentiate recurrent GBM vs radiation change

```
+  Concordant +  Treat
  Surgery

-  Discordant

-  Concerdant -  Observe
  Observe
```

Adapted from Bossuyt

Defining and Measuring Value

Evidence base
- Limited for quantitative imaging
- Methodological challenges

Clinical outcomes
- Need to link use of QI to improved outcomes
- Outcomes dependent on quality and availability of therapy
- Attribution of outcomes

Care efficiency
- Time to initiation or change in therapy
- Treatment standardization
- Impact on throughput and resource use
Additional measures of value

- Care efficiency
  - Time to initiation or change in therapy
  - Treatment standardization
  - Impact on throughput and resource utilization

- Use of registries eg. National Oncologic PET Registry
Paths to adoption of QI

- Value is the eye of the beholder
  - Demonstrate clinical value to clinicians, patients, payers

- Define value broadly
  - Impact on current care and care process; impact on future behavior

- Make it reportable
  - PQRS
  - NQF