

Caveat Emptor:
**Decision Support Tools for an
Emerging “Buyer’s Side” Health Care
Management**

10th Annual Disease Management Congress

Orlando, Florida

September 23, 2005

Thomas Wilson, PhD, Dr. PH

Founder & CEO

Population Health Impact Institute

www.PHIinstitute.org

513-677-8123

Outline

- Part I: Background: *Caveat Emptor*
- Part II: PHI Institute: Transparency & Credibility

Disclosure: IHPM has provided an unrestricted grant to PHII for two projects (focused interview and lit review). There is no other financial arrangement between the two organizations.

Background

Why Should Buyer's Beware? ...

Peer-Reviewed Literature: *Gold Standard or Crises of Confidence?*

JAMA publication (August 2005). Reviewed articles in major clinical, high impact journals, from 1990 to 2004.

16%: Original findings were later contradicted

16%: Subsequent findings were not as strong as first publication

Source: Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research JAMA. 2005 Jul 13;294(2):218-28.

Nature publication (June 2005) Responses to survey of 3,247 scientists.

6.0%: Yes to: “Failing to present data that contradicts one’s own previous research.”

12.5%: Yes to: “Overlooking others' use of flawed data or questionable interpretation of data”

15.5%: Yes to: “Changing the design, methodology or results of a study in response to pressure from a funding source.”

Source: Martinson BC, Anderson MS, de Vries R.. Scientists Behaving Badly. Nature 435, 737-738 (9 June 2005)

What about the Validity of non-Peer-Reviewed Studies?

A

V a c u u m

Peer-Reviewed Studies, Winston Churchill & Democracy

- In spite of their flaws, a peer-reviewed publication is better than all those other kinds of non-peer reviewed reports. We can read them, praise them, and critique it. Something we can't always do to many kinds of non-peer reviewed analyses.
- The flawed peer-review process is akin to Winston Churchill's description of democracy (1947): It is better than all those other forms of dissemination:

“Many forms of Government have been tried, and will be tried in this world of sin and woe. No one pretends that democracy is perfect or all-wise. Indeed, it has been said that democracy is the worst form of government except all those other forms that have been tried from time to time.”

The PHI Institute

Work-Site Lit Review: Methods & Interventions

Methods

Studies since 2000 in Medline: Key words: economic + occupational + health promotion.

Found 25 suitable studies; 15 were research studies.

Interventions

Dental, low back pain, HRA, integrated health management, executive physicals, vaccinations, depression / behavioral health, drug-free workplace.

The PHI Institute

Work-Site Lit Review: Findings

Study Designs

Prospective Cohort (pre-post), Cross-Sectional (pre-post), Participant-Non-Participant, Randomized Trial, Inferred Savings (QALYs), Trend Analysis

Findings

Most studies report positive economic value of the interventions; as do review articles (*i.e., “Nearly Everything Works”*).

However, we conclude it would be difficult / impossible for a decision maker to compare & contrast different interventions - *from a human capital investment perspective.*

The PHI Institute

Work-Site Lit Review: Implications

Goal:

Compare and Contrast

Defined Population Health Programs

Impact Methods and Results,

Openly & Credibly to

Empower Decision-Makers

The Population Health Impact Institute: A Non-Profit Organization

Mission: Promote and conduct independent & valid evaluations of defined population health programs through ...

Education

- Knowledge Transfer Workshops.

Research

- Papers and Articles based on Literature reviews, focus groups, surveys, health data.

Benchmarking

- VALIDITY: of methodology (assuming transparency of methods)
- VERIFICATION: of results (using multiple methods)
 - Methods-to-Method & Program-to-Program impact Crash TestsSM

Part II: The PHI Institute
Decision Support Tools & Framework

FOUNDATION:
TRANSPARENCY
of
Methods
&
Results
To
Permit Replication

*Nullius in verba**

**Slogan of the Royal Society of London (“Don’t take anyone’s word for it”)
from Sir Isaac Newton’s time to the present day*

PHI Institute's Five Evaluation Principles

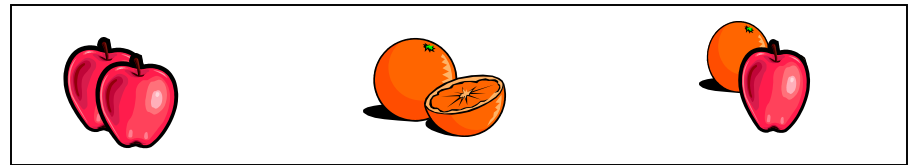
- 1) **Data Quality**
- 2) **Equivalence**
- 3) **Statistical Quality**
- 4) **Causality**
- 5) **Generalizability**

© 2005 Population Health Impact Institute. All Rights Reserved. PHI Institute encourages the private, non-commercial use of this framework, however, it reserves the exclusive right for benchmarking services.

Evaluation Principles Visualized

1. Data Quality

Measurement Error



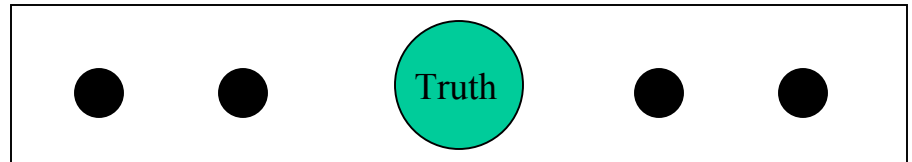
2. Equivalence

Differential Error



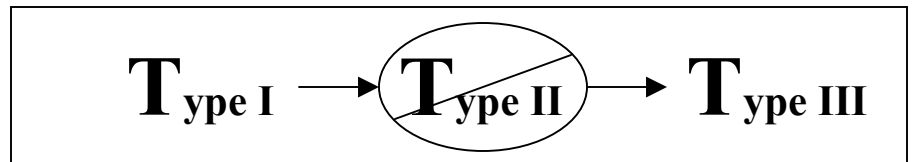
3. Statistical Quality

Non-Differential Error



4. Causality

Conceptual Error



5. Generalizability

Transmission Error



Illustrations of Principles

All articles cited here provoked responses in the form of letters-to-the-editor, most of these letters had responses from the authors (see reference slide at end)

Principle #1: Data Quality

Potential Measurement Error

Case: Definition of population in reference and intervention groups (in this case pre and post)

Potential problem: Different inclusion and/or exclusion criteria in intervention group than in reference group. *Measurement bias*

Original Source: Wilson TW, Gruen J, Thar W, Fetterolf D, Patel M, Popiel RG, Lewis A, Nash DB. Assessing ROI of defined-population disease management interventions. *Joint Commission Journal of Quality and Safety*. 2004; 30 (11); 614-21.

Principle 2: Equivalence

Potential Differential Error

Results: Positive ROI in CHF program

Case: Finding comparable reference group for op-in DM program, used non-participants and “propensity scoring” to select reference.

Potential problem: The number of people available for reference group (n=236) were considerably less than the number of participants (n=533). Many people were used more than once.

Original Source: Berg GD, Wadhwa S, Johnson AE. “A Matched-Cohort Study of Health Services Utilization and Financial Outcomes for a Heart Failure Disease-management Program in Elderly Patients.” *J Am Geriatr Soc* 52: 2004; 1–7.

Principle 3: Statistical Quality

Potential Non-Differential Error

Results: Positive ROI in Asthma program

Case: Statistical testing of reference group (based upon age) to show its similarity to intervention group.

Potential problem: The statistical test was stated as not-significant, when the difference was actually significant.

Original Source: Tinkelman D, Wilson S. “Asthma Disease Management: Regression to the Mean or Better?” *Am J Manag Care*. 2004; 10 (12): 948-54.

Principle 4: Causality

Potential “Conceptual Error”

Results: Positive ROI (8:1) in DM program

Case: ROI was based upon the difference in costs between those randomized into the intervention groups vs. the control group. A specific ROI was based upon changes among those with the actual intervention (n=961 with ROI of 68.0) compared to a like group

Potential problem: On recalculation, the ROI seems closer to 3.0 when calculating only on those who received the “care consideration” intervention, not all people randomized into the reference and control groups. Changes that occurred in the intervention population, but did not receive the intervention, were included in the 8:1 calculation
“False attribution” problem

Original Source: Javitt JC, Steinberg G, Locke T, Couch JB, Jacques J, Juster I, Reisman L. "Using a Claims Data-Based Sentinel System to Improve Compliance with Clinical Guidelines: Results of a Randomized Prospective Study." *Am J Manag Care*. 2005;11:93-102.

Principle 5: Generalizability

Potential “Transmission Error”

Results: Positive results on clinical measures in CHF population, no savings reported.

Case: Study with high “internal validity” (randomized control trial)

Potential problem: Potential low “external validity” CHF patients were already following evidence based guidelines. Had the study show an effect, it may have been expanded into other areas (see the authors letter to the editor)

Original Source: Galbreath AD, Krasuski RA, Smith B, et al. “Long-Term Healthcare and Cost Outcomes of Disease Management in a Large, Randomized, Community-Based Population With Heart Failure.” *Circulation*. 2004; 110(23):3518-26.

Observations

- Publication in a peer-reviewed journal is no guarantee of high validity, however, it does allow for further review, if warranted.
- All DM studies need to have “transparent” ROI and other impact methods so “buyers” can judge validity and make comparisons of results.

The Health Industry Forum of Brandeis University: Competitive Award to PHI Institute

Three Papers Planned

- 1) Peer-reviewed article on scoring system
 - Emphasis on need to improve of simple scoring based on study design (expert opinion, observational, quasi-experimental, and experimental)
 - Scoring should be based on principles, not study design.

- 2) PHI Institute's Delphi-panel made up of impartial evaluation experts to score 25 articles on disease management programs conducted in health plans-- based upon framework of Five Evaluation Principles.

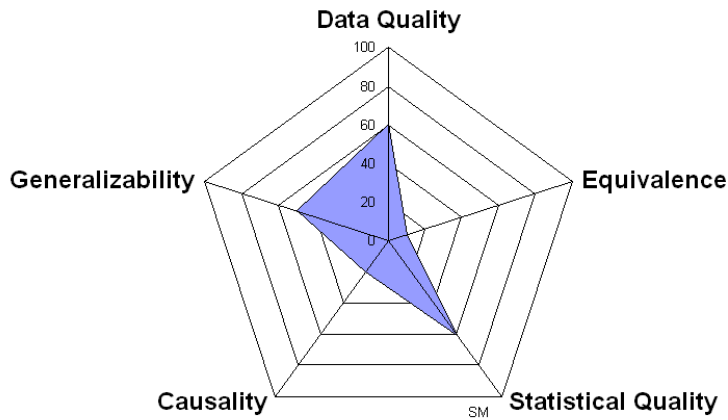
- 3) "Users Guide" to self-scoring of DM methods and results

Source: <http://sihp.brandeis.edu/forum/researchgrants.htm>

Example of Evaluation Scoring

Article 13

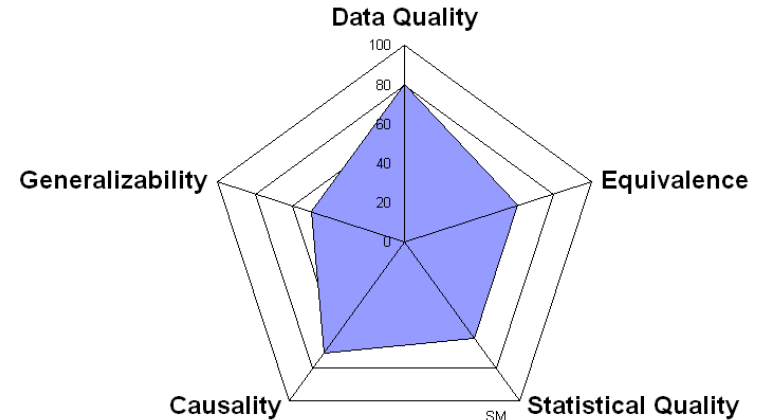
VALIDITY PENTAGONSM



© 2005 Population Health Impact Institute, Cincinnati, Ohio. For private use only. Commercial use prohibited. All other rights reserved.

Article 25

VALIDITY PENTAGONSM



© 2005 Population Health Impact Institute, Cincinnati, Ohio. For private use only. Commercial use prohibited. All other rights reserved.

© 2005 Population Health Impact Institute. All Rights Reserved. PHI Institute encourages the private, non-commercial use of this framework, however, it reserves the exclusive right for benchmarking services.

Future Directions of PHI Institute

- Validity scoring of work-site health promotion programs published in the peer-reviewed literature, using the same methods as the Brandeis project. (funded by IHPM with unrestricted educational grant)
- Develop “Transparency Standards” and other evaluation standards based on five evaluation principles with representation of three groups:
 - Evaluators
 - Buyers (employers, health plans, and government sector)
 - Sellers of defined population health management services

Conclusion:

Future of DM Measurement

- Compare and contrast, openly
 - *Depends on transparency*
- Compare and contrast, credibly.
 - *Depends on evaluation principles*

Bibliography on Principles

#1 DATA QUALITY: Wilson TW, Gruen J, Thar W, Fetterolf D, Patel M, Popiel RG, Lewis A, Nash DB. Assessing ROI of defined-population disease management interventions. *Joint Commission Journal of Quality and Safety*. 2004; 30 (11); 614-21.

Commentary: Duncan I. "Accuracy in the Assessment of Return on Investment of Defined Population Interventions" *Joint Commission Journal of Quality and Safety*. 2005; 31 (6); 357. Wilson T, Thar W, Gruen J. "Reply." *Joint Commission Journal of Quality and Safety*. 2005; 31 (6); 358.

#2 EQUIVALENCE: Berg GD, Wadhwa S, Johnson AE. "A Matched-Cohort Study of Health Services Utilization and Financial Outcomes for a Heart Failure Disease-management Program in Elderly Patients." *J Am Geriatr Soc* 52: 2004; 1-7.

Commentary: Wilson TW, Linden A. "Potential Bias in 'Controls' Used in a Heart Failure Disease Management Program." *J Am Geriatr Soc* 2005; 53(7): 1268-9.

#3 STATISTICAL QUALITY: Tinkelman D, Wilson S. "Asthma Disease Management: Regression to the Mean or Better?" *Am J Manag Care*. 2004; 10 (12): 948-54.

Commentary: Levin-Scherz, J. "To the Editor." *Am J Manag Care*. 2005; 11 (3): 134,136. Wilson T, Linden A. "To the Editor." *Am J Manag Care*. 2005; 11 (3): 136. Wilson S, Tinkelman D. "The Authors Respond." *Am J Manag Care*. 2005; 11 (3): 136-37.

#4 CAUSALITY: Javitt JC, Steinberg G, Locke T, Couch JB, Jacques J, Juster I, Reisman L. "Using a Claims Data-Based Sentinel System to Improve Compliance with Clinical Guidelines: Results of a Randomized Prospective Study." *Am J Manag Care*. 2005;11:93-102.

Commentary: Wilson TW, Linden A. "To the Editors." *Am J Manag Care*. 2005;11: 999. Javitt JC. "In Reply." *Am J Manag Care*. 2005;11: 999.

#5 GENERALIZABILITY: Galbreath AD, Krasuski RA, Smith B, et al. "Long-Term Healthcare and Cost Outcomes of Disease Management in a Large, Randomized, Community-Based Population With Heart Failure." *Circulation*. 2004; 110(23):3518-26.

Commentary: Linden A, Wilson T. "Letter." *Circulation*. 2005; 112: e11. Galbreath A, Freeman GL, Smith B, Krasuski RA, Stajduhar KC, Kwan MD, Ellis R. "Response." *Circulation*. 2005; 112: e11.

Contact Information

Thomas Wilson, PhD, Dr. PH

Founder & CEO

Population Health Impact Institute

10663 Loveland-Madeira Rd, #210B

Loveland, Ohio 45140

email: twilson@PHIinstitute.org

phone: 513-289-3743

web: www.PHIinstitute.org