A metrologist’s perspective on psychometric methods

PMhealth – Psychometrics in the health sciences 2017

L R Pendrill
Research Institutes of Sweden, Metrology, Eklandagatan 86, 41261 Göteborg (SE), phone:+46 767 88 54 44, mailto:leslie.pendrill@ri.se
Variation in Primary Care Indicators

Potential causes of variation:
- Disease prevalence
- How doctors diagnose
- How data coders interpret diagnoses

Source: OECD analysis of data from Prince et al. (2013) and the United Nations.
“Strong commitment of countries to make better use of health data, to foster international cooperation in health research and ultimately to improve health system performance and outcomes for people”

“Health data necessary to improve quality, safety and patient-centeredness of health care services and to support scientific innovation, discovery and evaluation of new treatments and to redesign and evaluate new models of health service delivery”
SP – Swedish national metrology institute
Metrology – quality-assured measurement

**Traceability** => *measurements can be compared*
- Under both repeatability and reproducibility
- Even different measurement quantities

**Uncertainty** => *declared measurement quality*
- ’Fit-for-purpose’
- Quantified risks of decision errors

1,234 m ± 0.016 m

**This gives processes and products which have:**
- Interoperability
- Improved communication
- Can be traded, are safe and lie within specifications
Object: Mass
**Delivering calibration**

**Object:** Mass

---

**Delivery 1: Calibration:** how much does my weight really weigh?

\[ 1.2 \text{kg} \pm 0.2 \text{ kg} \]
Delivering calibration

Delivery 1: Calibration: how healthy/ill am I?

0,8 units ± 0,2 units

1,0 units =>

0,8 units
Delivering calibration

Object: Health

Mini Mental State Examination
Svensk Revidering (MMSE-SR)

Utarbetad av: S Palmqvist  B Terzis  C Strobel  A Wallin
I samarbete med Svensk Förening för Kognitiva sjukdomar (SFk), 2012

Delivery 1: Calibration: cognitive ability?

0,8 units ± 0,2 units
Virtual Pills: How Video Games Can Boost Senior Brain Fitness

Posted by Nan Bauroth on Sep 14, 2014 in Aging &amp; Health

http://www.lifelinesys.com/content/blog/aging-amp-health/virtual-pills-how-video-games-can-boost-senior-brain-fitness
Cognitive tests: Working memory test (n-back)
[go to stats]
Interoperability

Needs

Delivering calibration

Object: Health

Confidence level

My health?

Demands

Precise meaning of exchanged information which is preserved and understood by all parties

Planning of technical issues involved in linking computer systems and services

Cooperating partners with compatible visions, aligned priorities, and focused objectives

Aligned legislation so that exchanged data is accorded proper legal weight

Legal Interoperability

Legislative Alignment

Organisational Interoperability

Organisation and Process Alignment

Semantic Interoperability

Semantic Alignment

Technical Interoperability

Interaction & Transport

Political Context

Coordinated processes in which different organisations achieve a previously agreed and mutually beneficial goal

ISA
Interoperability Solutions for European Public Administrations

L R Pendrill
Reorienting health systems to become more people-centred

Invest in measures so that health systems deliver what matters most to people.

Too often:
• we only rely on measures of what health systems do, and how much they cost,
• rather than their effects on patients.
Historically, NHS focused on measuring ‘inputs’ such as attendances, hospital admissions, and waiting times.

- Easy to measure,
- but fail to capture whether patient’s care was good or bad, or even clinically effective.

- No substitute for measuring actual outcomes as well as costs involved over full cycle of care for patient’s problem.
- Recent efforts to capture patient “experience” are useful, but not same as outcomes

Key challenge:
- **defining outcomes** that matter for each condition,
- how to **measure** them.

- Thus far: efforts bottom up and different across organizations and geography.

Pressing need to develop **standardized sets of outcomes** by condition:
- to enable comparison and learning,
- put in place infrastructure and tools needed to collect and measure them across entire system
Patient-Centred Outcome Metrology for Healthcare Decision-Making

Stefan J Cano
Modus Outcomes, Spirella Building, Letchworth Garden City, SG6 4ET, UK, e-mail: stefan.cano@modusoutcomes.com

L R Pendrill
SP Metrology, Research Institutes in Sweden, Ekladas gatan 86, S-412 61 Göteborg, e-mail: leslie.pendrill@sp.se

Skye P Barbic
Faculty of Medicine, Department of Occupational Science and Occupational Therapy, the University of British Columbia, Vancouver, BC V6T 2A1, Canada, e-mail: skye.barbic@ubc.ca

William P Fisher, Jr.
BEAR Center, Graduate School of Education, University of California, Berkeley, CA 92720 USA, e-mail: wfisher@berkeley.edu
Joint Committee for Traceability in Laboratory Medicine (JCTLM)

JCTLM - Joint Committee for Traceability in Laboratory Medicine
JCTLM Executive Committee

JCTLM Working Groups:
- JCTLM Database WG: Reference Materials, Procedures and Measurement Laboratories
- JCTLM WG on Traceability: Education and Promotion
Recommendations of IFCC Task Force on Impact of Laboratory Medicine on Clinical Management and Outcomes (TF-ICO) include:

- “Effective collaboration with clinicians
- Determination to accept patient outcome and patient experience as primary measure of laboratory effectiveness.”

"Large number of tools available to measure person-centred care, but no agreement about which tools most worthwhile.

No ‘silver bullet’ or best measure covers all aspects of person-centred care.
Person-centred care (PCC)

ICHOM

About Standard Sets Measure TechHub Events & Media Support Us Login

All 21 Cardiovascular 3 Congenital anomalies 2 Digestive 1 Malignant neoplasms 5 Maternal and neonatal 1 Mental and behavioral disorders 1

Musculoskeletal 2 Neurological 2 Primary/preventative care 1 Sense organ 2 Urogenital 1

Pregnancy and Childbirth Maternal and neonatal

Inflammatory Bowel Disease Digestive

Overactive Bladder Urogenital

Colorectal Cancer Malignant Neoplasms
Assessment Center

In the past 7 days
I felt uneasy
- Never
- Rarely
- Sometimes
- Often
- Always

In the past 7 days
I found it hard to focus on anything other than my anxiety
- Never
- Rarely
- Sometimes
- Often
- Always

In the past 7 days
I felt like I needed help for my anxiety
- Never
- Rarely
- Sometimes
- Often
- Always

Physical Function
Satisfaction with Discretionary Social Activities
Satisfaction with Social Roles

Please enter your age

Computerized Adaptive Test (CAT) Report
Date: 05-May-15
Your age: 55
Your gender: Male
Computerized Adaptive Tests: Anxiety/Fear

Your score on the Anxiety/Fear CAT is 71. The average score is 50.
Your score indicates that your level of Anxiety/Fear is higher (worse) than:
- 98 percent of people in the general population
- 99 percent of people age 55-64
- 99 percent of males

http://www.nihpromis.org/software/demonstration
PMhealth June 2017
Person-centred care (PCC)

**OMERACT values**, include:

- Trust
- Respect
- Transparency
- Partnership
- Communication
- Diversity
- Confidentiality
- Co-learning with respect to patient involvement
1. To reach consensus on which measurement properties should be evaluated of Health-Related Patient-Reported Outcomes (HR-PROs) and how they should be defined

2. To develop standards for how these measurement properties should be evaluated in terms of study design and statistical analysis
Person-centred care (PCC)

Measurements in PCC

Compared with traditional measures (e.g. blood pressure or queueing times) in care:

• Patient assessment ≠ Professional assessment

Patient:
• More symptoms
• Greater impact on daily living

• More subjective
• ’Saturation’ effects: almost everyone’s happy!
• Large person-to-person scatter

Special tools needed!
Comparing instruments rating of physical disability

δ(Challenge)

W P Fisher Jr. 1999
Measuring Man:
- Status, function of person
- Test against specifications

Man as Measurement Instrument:
- Perception of product/service function, comfort etc
- Propose improvements in product
Five components of health:  
[International Classification of Functioning, Disability and Health (ICF)]
Ordinal data (e.g. ‘Satisfaction’) – incorrect use of statistics in traditional analyses

Item responses:
• only ordered structure
• not numerical value in mathematical sense

Statistical methods applicable to data from rating scales differ completely from traditional methods for quantitative variables

Calculations based on adding or subtracting ordinal data not appropriate
Measuring People

Rasch (1961)

\[ \delta(\text{Challenge}) \]

\[ \theta(\text{Ability}) \]

- Correct ordinal data treatment
- Better resolution
- Separation of person and item measures

\[ \log \left( \frac{P_{\text{success}}}{1 - P_{\text{success}}} \right) = \theta - \delta \]
13. Vilka 3 ord var det jag bad dig att lägga på minnet?

- HOTELL
- BANAN
- KANIN

Poäng: 3
Mini Mental State Examination
Svensk Revidering (MMSE-SR)

MMSE Items

More difficult

1. Orientation
2. Immediate recall
3. Orientation
4. Immediate recall
5. Orientation
6. Immediate recall
7. Orientation
8. Immediate recall
9. Orientation
10. Immediate recall
11. Immediate recall
12. Orientation & counting
13. Delayed recall
14 - 20.

Less difficult

Difficulty, \( \delta \)

-1.5

-1

-0.5

0

0.5

1

Average (CTT)

underestimated
Task difficulty, $\delta$

Metrological references

Difficulty

Mass
Balance as Measurement Instrument - Sensitivity ($C$)

$R = C \cdot S + "additional\ terms"$

**Stimulus ($S$):** Mass of weight

**Response ($R$):** Mass of weight $\times$ Balance sensitivity
Man as Measurement Instrument - Sensitivity (C)

\[ R = C \cdot S + \text{"additional terms"} \]

\[ R = \frac{\partial P_{\text{success}}}{\partial z} \cdot (z = \theta - \delta) \]

Response (R):
- Task challenge x 'Instrument' sensitivity

Stimulus (S): Task challenge

\[ P_{\text{success}} = \frac{e^{\theta - \delta}}{1 + e^{\theta - \delta}} \]

\[ \delta(\text{Challenge}) \]

\[ \theta(\text{Ability}) \]
Task difficulty, $\delta$

\[ \delta_j = \{\delta_j\} \cdot [\delta_j] \]

unit

'logistic measurement function'

\[
\log \left( \frac{P_{success,i,j}}{1 - P_{success,i,j}} \right) = \rho_s \cdot (\theta_i^* - \delta_j^*)
\]

\[ \delta_j^* = \frac{\delta_j}{\rho} \]

Set, $s$, of items rather than single item


\[
[\theta_i] = [\delta_j]
\]

'common unit' of measure
NeuroMet

EMPIR 15HLT04: Innovative measurements for improved diagnosis and management of neurodegenerative diseases
June 2016 – June 2019

Short Name | Organisation legal full name | Country
---|---|---
LGC | LGC Limited | United Kingdom
INRIM | Instituto Nazionale di Ricerca Metrologica | Italy
LNE | Laboratoire National d’Essais | France
PTB | Physikalisch-Technische Bundesanstalt | Germany
SP | SP Sveriges Tekniska Forskningsinstitut AB | Sweden
Charité | Charité - Christian-Albrechts-Universität zu Berlin | Germany
CHRU Mpt | CHRU Mpt | Switzerland
UCL | National Amyloid Centre and UCL Drug Discovery, University College London | United Kingdom
UEA | School of Nursing at the University of East Anglia | United Kingdom

WP No | Work Package Title | Active Partners (WP leader in bold)
---|---|---
WP1 | Patient cohort sampling and stratification | PTB, Charité, SP, INRIM, CHRU Mpt, LGC, UEA
WP2 | Minimally invasive methods for AD and PD diagnosis | LGC, LNE, CHRU Mpt, Charité, UCL, INRIM, UEA, SP
WP3 | Establishing traceability of AD and PD biomarker measurements | LNE, LGC, CHRU Mpt, Charité, UCL, SP
WP4 | Patient-centred outcome measures | LNE, INRIM,

WP6 | Management and coordination | CHRU Mpt, Charité,

Acknowledgments
The European Metrology Programme for Innovation & Research (EMPIR, Horizon2020, Art. 185) is jointly funded by the EMPIR participating countries within EURAMET (www.euramet.org) and the European Union in this EMPIR 15 HLT04 NeuroMet project (coordinator: LGC (UK))
WP1: Establishment of patient cohorts (MCI, AD) and matched control samples

Patient stratification by clinical assessment protocols and laboratory analysis

WP1: in-vivo MRI on AD patient cohort

WP2: Blood analysis of biomarker for early diagnosis on AD samples

WP3: Application of reference methods on AD patient samples

WP4: Patient-centred outcome measures:

- 4.1 novel clinical assessment protocols
- 4.2 correlated with biomarkers

Innovative measurements for improved diagnosis and management of neurodegenerative diseases
Mini Mental State Examination
Svensk Revidering (MMSE-SR)

Object: Health

Delivery 1: Calibration: cognitive ability?

0.8 units ± 0.2 units
MMSE Items expand

Difficult, $\delta$

Less difficult

More difficult

Mini Mental State Examination
Svensk Revidering (MMSE-SR)
“‘Earlier I told you the names of three things. Can you tell me what those were?’”

“Repeat the phrase: ‘No ifs, ands, or buts.’”

“Name all three unrelated objects”
Using a Rasch scale to characterize the clinical features of patients with a clinical diagnosis of uncertain, probable, or possible Alzheimer disease at intake

Larry F. Hughes, Kyle Perkins, Benjamin D. Wright and Heather Westrick
11. Immediate recall

13. Delayed recall

Less able

Optimum point

More able

Sample range

Room for improvement

Scale range

Less difficult

More difficult
Mini Mental State Examination

Anatomic Correlation of the Mini-Mental State Examination: A Voxel-Based Morphometric Study in Older Adults

Mikael Dionesio¹, Sébastien Celle², Guillaume T. Doreni², Frédéric Rozina², Samir Henni¹, Robert Bartha³, Olivier Beauchet⁴, Cedric Amouyel⁵,∗

¹ Université d’Angers, Laboratoire Angévin de Recherche en Ingénierie des Systèmes (LARIS)-EA715-F. 49000 Angers, France, 2 Service de Physiothérapie, Centre Hospitalier de Saint-Malo, 3 University of Angers, Département de Neurologie et de Réadaptation, CHU Angers, F-49000, France, 4 Service de Physiothérapie, Centre Hospitalier de Saint-Malo, 5 University of Angers, Département de Neurologie et de Réadaptation, CHU Angers, F-49000, France

r = 0.62, P<0.001
r = 0.23, P=0.026
r = 0.19, P=0.031
r = 0.35, P=0.011
• PMhealth – Swedish national workshop in psychometrics in health sciences 2017 (Kristianstad, June)
• EMPIR mini-symposium, Innovative measurements for improved diagnosis and management of neurodegenerative diseases, (RI:SE, Göteborg, July)
• ENBIS 2017 (Naples, Sept.)
• Metrologie Congres 2017 (Paris, Sept.)
Quality-assured measurement of perception

Measurement of innovation

Comfort Assessment for heavy incontinent

Measures of accessibility

EMRP NEW04 uncertainty

Konsten att mäta upplevelse och intryck av plastprodukter

BULLETIC LIMIT Data:
- Velocity of bullet that perforates vest with 50% probability

STANDARDS
- U.S. Dept. of Justice, National Institute of Justice Standard 0101.06 (2008)

Gösta Erman Laboratory for Sensory Research

Rise Institutes
Thank you for your attention

adj prof Leslie PENDRILL
RISE Metrology, Göteborg (Sweden)
leslie.pendrill@ri.se