

PAPER • OPEN ACCESS

## Patient-centred cognition metrology

To cite this article: S J Cano *et al* 2018 *J. Phys.: Conf. Ser.* **1065** 072033

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the [collection](#) - download the first chapter of every title for free.

## Patient-centred cognition metrology

**S J Cano<sup>1,5</sup>, J Melin<sup>2</sup>, WP Fisher<sup>3</sup>, AJ Stenner<sup>4</sup>, LR Pendrill<sup>2</sup> and the EMPIR NeuroMet 15HLT04 consortium<sup>5</sup>**

<sup>1</sup> Modus Outcomes, Spirella Building, Letchworth Garden City, SG6 4ET, UK, stefan.cano@modusoutcomes.com

<sup>2</sup> RI.SE, Eklandagatan 86, S-412 61, Göteborg, Sweden, jeanette.melin@ri.se, leslie.pendrill@ri.se

<sup>3</sup> BEAR Center, Graduate School of Education, University of California, Berkeley, CA 92720 USA, wfisher@berkeley.edu

<sup>4</sup> MetaMetrics, Inc., 1000 Park Forty Plaza Drive, Suite 120, Durham, NC 27713, USA & The University of North Carolina, Chapel Hill, North Carolina USA, [jstenner@Lexile.com](mailto:jstenner@Lexile.com).

<sup>5</sup> NeuroMet coordinator Milena Quaglia, LGC, Milena.Quaglia@lgcgroup.com

E-mail: stefan.cano@modusoutcomes.com

**Abstract.** We have yet to realise reconciliation between metrology in the physical sciences and measurement practices in the social sciences. Here, we build on our previous arguments by proposing a psychometric methodology to support metrology in the social sciences. This will be exemplified in the European EMPIR NeuroMet 15HLT04 project, part of which exploits causal Rasch models, novel construct specification equations, and subsequent improved patient centred cognition measurement in support of healthcare and policy decision-making in Alzheimer's disease. In addition, the programme's goal of forming key stakeholder networks, including clinicians, instrument manufactures, biopharma, national and international organisations, will ensure consistent uptake of the new measurement methods.

### 1. Introduction

This is the third in a series of IMEKO papers combining metrology, psychometrics, and philosophy to support the global comparability and equivalence of results in social measurement practices supporting high stakes decision making (e.g., education, healthcare). In our first paper, we concluded that measurement, whether physical or social, should have the same definition (i.e., the ratio of two magnitudes of the same *attribute*, where the denominator is the unit, providing for invariant comparisons) and the same broad goal (i.e., quantification of meaningful variables) [1]. By way of application, we turned the spotlight on patient (or person) centred outcome measures (PCOMs; e.g., questionnaires, biometric equipment, wearables) that quantify patients' health (e.g., pain, mood, and function) or experience of healthcare. To ensure decision makers within health systems have access to the necessary objective evidence, we proposed the requirement for the same kind of quality-assurance

<sup>5</sup> To whom any correspondence should be addressed.



of PCOMs as is established in physics and engineering [2]. i.e., ‘traceability’ and declared measurement uncertainty.

In our second paper, we emphasised the importance of international collaboration to realise a higher standard of social measurement practices, with five considerations: 1) key stakeholder involvement (including patients, clinicians, caregivers, policy makers); 2) a clear definition of the measurand in regards to the intended clinical use [3]; 3) a clear definition of the clinically allowable error of measurements; 4) international cooperation and consensus to navigate the complexities of the development of metrologically sound reference measurement systems; and 5) continued clinical validation of newly calibrated instruments. The ultimate benefits of this kind of international collaboration would include traceability, comparability, and (possible reduction) of measurement uncertainty.

In this paper, we bring together the two main threads of the above arguments, in the context of an ongoing European Union funded programme research into neurodegenerative diseases. First, we describe the central role of Rasch Measurement Theory [4], and especially causal Rasch models [5], to enable references for traceability and means of evaluating measurement uncertainty to be established. Second, we introduce the European EMPIR NeuroMet 15HLT04 project [6] which, as part of its remit, is producing improved PCOMs in support of healthcare clinical and policy decision-making in Alzheimer’s disease.

## **2. The central role of Rasch Measurement Theory in Patient Centred Outcome Metrology**

Georg Rasch, a Danish mathematician, argued that the core requirement of social measurement should be the same as that in physical measurement (i.e., ‘invariant comparison’). He developed the simple logistic model (now known as the Rasch model). In his landmark research in education and psychology, he demonstrated that his approach met the stringent criteria for measurement used in the physical sciences [7]. Rasch Measurement Theory models linearized ordinal (counted fraction) summed scores through paired comparisons of any two persons, any two items, or any one person and one item, defined by the logarithm of the relative probabilities [8]. The linearize measurements of persons and items are on the same scale with a common unit, freed up from the distributional properties of each other. Thus, the Rasch model realizes, mathematically, the requirements for scientific measurement of invariant comparisons of people, and items, on the same linear scale [9].

Rasch Measurement Theory provides the ideal infrastructure to support a metrological framework for the social sciences. This is due to the inherent properties of the Rasch model (i.e., parameter separability, statistical sufficiency, specific objectivity) embedded in Georg Rasch’s general philosophy of measurement [4], and subsequently formalized in the language of measurement traceability [10] and uncertainty [11]. However, straightforward fit to the Rasch model does not automatically ensure valid measurement. Psychometric statistics can be misleading when considered in isolation and cannot be expected to produce consistently meaningful results when considered apart from substantive evaluations of construct content [12, 13].

The evolution of a construct theory [14, 15] has been articulated through five levels. Level 1 measurement systems have no explicit construct theory and items are calibrated empirically from data analysis. At the highest level 5, measurement systems have a construct theory and specification equations (known as theory-referenced measurement); a conjoint model that measures reader (instrument) ability and text (object) complexity in a common unit (proposed by Stenner *et al.*), for instance, is an example of the most advanced form. Theory-referenced measurement affords three central benefits, as articulated through a comparison of measurement systems in the social sciences with those used in for example thermometry [5, 2]. Thus, for both types of measurement system there is: 1) an observational outcome, a response, such as a count; 2) a causal mechanism that transmits variation in the intended attribute (i.e., the stimulus such as reading text difficulty or temperature), via the instrument (e.g. the test person or thermometer) to the observed response; and 3) attribute

measures (i.e., the measurand) expressed in some unit (e.g., Lexiles or degrees Celsius). Once level 5 has been reached, we have successfully modelled causal relationships.

There are currently very few examples of measurement systems developed using theory-referenced measurement in the social sciences. However, the measurement of cognitive performance in clinical settings to assess the impact of dementia provides a promising medium [16-18].

### **3. Innovative measurements for improved diagnosis and management of neurodegenerative diseases**

Neurodegeneration is an increasingly urgent global public health issue. Alzheimer's disease is one of the most common neurodegenerative diseases. Currently, there are no minimally invasive diagnostic tools which allow for early diagnosis or monitoring of the progression of disease in patients. The European EMPIR NeuroMet 15HLT04 project (2016-2019) [6] combines expertise from National Metrology Institutes, clinical, and academic institutions. The goal of the project is to address the limitations currently constraining clinical innovation and uptake in neurodegenerative disease diagnosis and treatment, in measurement sensitivity and high measurement variability of recognised and novel biomarkers for AD (e.g. lumbar puncture, blood and saliva sampling), as well as in neuroimaging (non-invasive magnetic resonance imaging and spectroscopy).

Cognitive performance tests usually include tasks related to recall, language and praxis. They can be regarded as simple examples of person-centred measurements in contrast with the more 'technical' aspects of biomarkers, which are surrogates for the impact of the disease. As part of EMPIR NeuroMet 15HLT04 project, PCOMs of cognitive performance are examined, with the ultimate goal of improving the analysis of correlations with the various AD biomarkers in the manner of previous research modeling disease severity on the basis of data from the clinical laboratory [17-19]. Addressing specifically Mild Cognitive Impairment and Alzheimer's disease, work is now in progress deploying Rasch Measurement Theory to develop a 'NeuroMet Memory Score' based on results from administering an extensive battery of legacy instruments (e.g., MMSE, Corsi Block Test, Digital Span Test (DST)) on a NeuroMet cohort recruited by the Charité hospital in Berlin, as well as earlier test results. The resultant improved estimates across comprehensive spans of cognitive task difficulty and patient ability are in turn enabling formulation of novel construct specification equations for correlation using multivariate principal component regression with, respectively, task entropy (e.g. logarithm of number of digits in DST) and diverse biomarkers (e.g. for brain atrophy) [2].

Establishing new measurements, which can be ascribed a level-5 construct theory, should lead to more fit-for-purpose, better targeted and better administered cognitive measurement systems anchored in item banks of quality-assured metrological references for cognitive task difficulty and patient ability. By exploiting causal models, metrological references for cognitive task difficulty, akin to certified reference materials, promise to improve the quality of patient-centred care of Alzheimer's disease and other neurodegenerative diseases by enabling the traceable calibration of both additional cognitive tasks as well as the effects of intervention or disease progression on the cognitive ability of each individual patient.

An additional key aspect of the EMPIR NeuroMet 15HLT04 project, and one in line with our previous recommendations [3], is the formation of a stakeholder network consisting of consortium partners and relevant neurodegenerative disease stakeholders including clinicians, instrument manufacturers, biopharma, national and international organisations. By doing so, the programme is pre-empting and preparing for the widespread communication and uptake of improve measurement comparability of established biomarkers and PCOMs.

### **4. Conclusion**

Metrological quality assurance of person-centred outcome measures is essential if reliable decisions about care needed to diagnose treat and rehabilitate are to be made consistently throughout the

healthcare system. The interim findings of the EMPIR NeuroMet 15HLT04 project already show how causal models and novel construct specification equations can improve cognitive performance testing. This will lead to the provision of guidance and tools to clinicians for improved diagnosis, better prediction of future patient decline and more appropriate patient treatment. Overall, the NeuroMET project will combine the strengths of National Metrology Institutes and neurodegenerative disease clinicians to establish an infrastructure, providing guidance and quality-assured measurement tools to clinicians, academics and pharmaceutical companies for:

- Appropriate study design and definition of the uncertainty of NDD clinical assessment protocols to improve diagnosis and NDD progression monitoring;
- More accurate stratification of patient cohorts, with respect to NDD status, for enrolment into clinical trials and for more informed patient management;
- Improved measurement comparability for NDD biomarkers through optimised measurement procedures and development of SI traceable reference methods for key biomarkers.

## 5. Acknowledgments

This work has been performed as part of the 15HLT04 NeuroMet project which belongs to the European Metrology Programme for Innovation & Research (Horizon 2020), jointly funded by the EMRP participating countries within EURAMET ([www.euramet.org](http://www.euramet.org)) and the European Union.

## References

- [1] Cano SJ, Vosk T, Pendrill LR and Stenner AJ J 2016 *J. Phys. Conf. Series* **772** 012025
- [2] Pendrill LR 2018 *Meas. Sci. Technol.* **29** 034003
- [3] Cano SJ, Pendrill LR, Barbic SP and Fisher WP 2017 *Joint IMEKO TC1-TC7-TC13 Symp.: Measurement Science Challenges in Natural and Social Sciences* (Rio de Janeiro, Brazil)
- [4] Rasch, G 1977 *Danish Yearbook of Philosophy* **14** 58–94
- [5] Stenner AJ, Fisher W, Stone M, Burdick D 2017 *Frontiers in Psychology* **4** 1-14
- [6] EMPIR Project 15HLT04 NeuroMet 2016 *Innovative Measurements for Improved Diagnosis and Management Neurodegenerative Diseases 2016–2019* <https://www.euramet.org/research-innovation/research-empir/empir-calls-and-projects/empir-call-2015-health-si-broader-scope-normativeresearch-potential>; interim 18 month publishable summary, Feb 2018
- [7] Rasch G 1960 *Probabilistic models for some intelligence and attainment tests* (Copenhagen, Danish Institute for Education Research)
- [8] Andrich D 1988 *Rasch models for measurement* (Beverly Hills, Sage Publications)
- [9] Wright BD, Stone MH 1979 *Best test design: Rasch measurement*. (Chicago, MESA)
- [10] Fisher WP, Stenner AJ 2016 *Measurement: Interdisciplinary Research and Perspectives* **92** 489-496
- [11] Andrich D 2017 *Joint IMEKO TC1-TC7-TC13 Symp Measurement Science Challenges in Natural and Social Sciences* (Rio de Janeiro, Brazil)
- [12] Hobart JC, Cano S, Baron R, Thompson A, Schwid S, Zajicek JP and Andrich DA 2013 *Mult Scler* **19** 1773-1783.
- [13] Maul A 2017 *Measurement: Interdisciplinary Research and Perspectives* <https://doi.org/10.1080/15366367.2017.1348108>
- [14] Stenner AJ, Smith M 1982 *Perceptual and Motor Skills* **55** 415-426
- [15] Stenner AJ, Smith M, Burdick D 1983 *J Educ Measure* **20** 305-316
- [16] Cano SJ, Hobart JC, Posner H, Selnes O, Stern Y, Thomas R and Zajicek JP 2013 *J Alzheimers Dement* **9** S10-20
- [17] Hughes L, Perkins K, Wright BD, and Westrick H 2003 *J Alzheimer's Disease* **5** 367-373
- [18] Fisher WP, Burton E 2010 *J Applied Meas* **11** 271-287
- [19] Markward NJ, Fisher WP 2004 *J Applied Meas* **5** 129-141