

VIEWPOINT

Con: Can biomarkers be gold standards in Alzheimer's disease?

Kenneth Rockwood^{1,2*}

Abstract

As Alzheimer's disease remains a clinical diagnosis, and as clinical diagnosis can be difficult, it makes sense to look for so-called biomarkers. A biomarker predicts who is likely to have the illness and who is not. Some biomarkers might even correlate with a clinically meaningful response to treatment. Developing biomarkers is often characterized as searching for a diagnostic gold standard that can seem appealing in its promise of certainty. Even so, considering both the economic history of the gold standard and the results of neuropathological studies, framing the search for measurable, biological correlates of dementia syndromes in this way is likely to be self-defeating. Instead of considering biomarkers as providing certainty through referent criterion validation, currently it makes more sense to test their construct validity and their predictive ability. This means that while biomarkers should inform, they will not dictate clinical meaningfulness. For the foreseeable future, even were they to inform diagnosis, biomarkers cannot substitute for understanding whether patients and caregivers find a given dementia treatment effective. Instead, clinicians should recognize their own determining role, both in dementia diagnosis and in the evaluation of treatment. These roles will best be executed by hearing what patients and caregivers tell us about dementia, and its response to treatment.

The syndrome of dementia is greatly feared by many people, robbing them of their dignity, independence and ability to lead a meaningful life – at least on the terms set by their prodementia selves. The diagnosis of dementia can be difficult, especially for nonexperts. First, the

syndrome of dementia needs to be distinguished from somewhat less onerous diagnoses, such as deafness, depression and delirium. Next, when the dementia is established, its cause needs to be sought out.

In practice, the syndromic diagnosis is more complicated still – even for experts – because it is common to identify people who have verifiable memory symptoms, and sometimes even mild decline in one of two other cognitive domains. What can be especially tricky is that many such people appear not (yet) to have reached a threshold of functional impairment that would allow a dementia diagnosis to be made with confidence. This state of cognitive impairment without important cognitive decline is best known as mild cognitive impairment. Of some interest is the fact that this state places the person so affected at much higher (but not certain) risk for developing unequivocal dementia. For these reasons – diagnostic uncertainty, the need to distinguish, early on, those at greatest risk, and (as discussed below) the need to monitor treatment effects – much effort is now being devoted to developing and validating so-called biomarkers. But how reasonable is it that biomarkers will meet these important challenges?

Biomarker is the term given to 'measurable biological characteristics that can either serve as indicators of normal or pathogenic processes in the body, or as tools to track pharmacological responses to therapeutic drugs' [1]. Their validation in dementia is inspiring an enthusiasm that seems imprudent to many people, because it glosses an essential aspect of how to test tests.

Validity can be understood in many ways, but a trinitarian approach of content, construct and criterion validity is well accepted [2]. *Content* validity refers to whether a test stands to reason; its assessment is largely qualitative. *Construct* validity assesses whether a candidate measure correlates with like measures, and not with unlike measures. *Criterion* validity has two components – the ability to predict outcomes (tested here in forecasting disease progression, or the response to treatment) and validation against a referent, which in medicine is often referred to as validation against a diagnostic gold standard.

The problem is not just that we have no gold standard in dementia; the problem is that even were we able to

*Correspondence: kenneth.rockwood@dal.ca

²Divisions of Geriatric Medicine & Neurology, Department of Medicine, Dalhousie University, Halifax, Nova Scotia, Canada B3H 2E1

Full list of author information is available at the end of the article

measure a uniquely dementia-associated pathological process for diagnostic purposes, it is unlikely to meet the level of near-magical thinking needed now required to accept a test as a gold standard. Most important, we can never have a high level of certainty about how biomarkers relate to clinical meaningfulness without involving patients and caregivers, the very factor that biomarkers are aiming to supplant.

The experience with what were long considered gold standard dementia biomarkers is salutary. Given their long iconic role, many of us believed – if there was some way to measure them prior to autopsy – the plaques and tangles that are demonstrable by histopathology would have allowed what the National Institute of Neurological Disorders and Stroke/Alzheimer's Disease and Related Disorders Association diagnostic criteria for Alzheimer's disease (AD) called a 'definitive' diagnosis [3]. But those criteria date back more than a generation – to 1984 – and are beginning to show their age.

The idea of a definitive diagnosis resting on plaque and tangle counts has been under erosion since at least the 1997 Nuns Study report of the mediating role played by even single lacunes in disease expression [4]. Amongst many reports from the new millennium, the Cognitive Function and Ageing Study investigators – who examined autopsy cases in a prospective cohort study of older, community-dwelling people – found that more people aged 85+ had pathological evidence of dementia than had clinical evidence of dementia. What is more, no cut-off point of degenerative or ischemic lesions optimized the dementia diagnostic accuracy [5]. So if not even neuropathology can offer a gold standard diagnosis, perhaps the fault is not with the plaques and tangles but with the idea of a criterion diagnosis, especially in dynamic biological systems.

What, then, is (or was) the gold standard? In economic history, the term refers to the practice of fixing a country's currency against a specified amount of gold. But history has not been kind to that proposition; notwithstanding the current economic upheavals, not even the gold standard is a gold standard, an idea in which few economists see merit [6]. Indeed, not since President Richard Nixon ended the Bretton Woods agreement in 1971 has a major world currency used the gold standard.

So why does the dementia community persist with using a metaphor for an idea long discredited by the people who understand it professionally? Beyond inertia, and the typical economic naiveté of people in this line of work ('a physician and his money are soon parted', as the financial planners say) is the potential advantage offered to the pharmaceutical industry. At present, pharma must contend with comparatively messy clinical interactions to know whether their drugs work. Their desire to move to

what is typically described as the cleaner objective of need chiefly to change a number, preferably one that can be obtained with an easy-to-order blood or imaging test, is understandable. As a pharma attendee at a recent dementia biomarkers meeting put it – 'we need our own cholesterol test'.

Such a test would offer many advantages. It would have the patina of objectivity, which is often equated to non-arbitrariness. This test could dramatically reduce the dimensionality of dementia, which otherwise requires evaluation of cognition, function, behavior, quality of life, caregiver burden and costs. But regulators have been careful to distinguish between those biomarkers that aid in the diagnosis of an illness, and surrogate markers that can substitute for a clinically meaningful endpoint, such as a measure of how a patient feels, functions or survives [7]. To be validated as a surrogate measure, any candidate measure for AD must show a unique link to the disease pathophysiology, must be specific and sensitive in relation to clinically diagnosed dementia, and must be able to distinguish degrees of risk in outcomes. Many considerations mitigate against any current measure meeting these criteria anytime soon. Not least of these are the heterogeneity of disease progression in cognitive impairment and the tendency for the most common dementia phenotypes – late-onset AD in older patients with multiple vascular risk factors – to have more than one operational cause of their dementia.

The opportunity cost in terms of a better appreciation of AD and its treatment is another reason to be skeptical about viewing biomarkers as the gold standard. In the early days of AD drug trials, evaluation of executive function was all but unknown. Caregiver reports that 'my dad is more like himself' were dismissed as anecdote. Twenty years later, while frontal lobe testing had become more evident, even replicable patient accounts ('I feel like the fog has lifted/the smoke has cleared/a window has been opened/I was standing off stage and now I am front and centre') receive short shrift. We are embarking on a new era of AD treatment, past the cholinesterase inhibitors. The lesson from that early experience is not just that we should use the clock drawing test, but that we should listen to what patients and caregivers have to say, with the idea that they will have more to teach us about how our new treatments work – and with that with how AD, and the brain, operates.

The alternative to a gold standard is to accept and better prosecute a strategy of combing construct validation with criterion validity that focuses on predicting important outcomes. This means employing more than one class of measure – not just a biomarker, but judgment-free tests (such as neuropsychological ones) and judgment-based tests (such as clinical interviews, and patient and caregiver questionnaires). It

means improving clinical measures and planning to retain them beyond some preliminary phase of validation for biomarkers. Importantly, it means that the emphasis should be on predicting outcomes (especially the outcome of predicting whose cognitive disorder is most likely to progress, or who is most likely to respond to treatment, and how they will respond) as the highest standard for all measures. These are goals worth aiming for and, especially in framing the question of what constitutes successful treatment and whether that goal is being attained, we will find it easier to get there if we see a central role for the experience of patients (and their caregivers) beyond how their biomarkers change.

Abbreviations

AD, Alzheimer's disease.

Competing interests

The author declares that he has no competing interests.

Acknowledgements

This paper endorses the systematic assessment of patient and caregiver preferences in dementia, which is facilitated by software commercialized by DementiaGuide Inc., of which the author is president, chief scientific officer and majority shareholder. KR receives career support from the Dalhousie Medical Research Foundation as the Kathryn Allen Weldon Professor of Alzheimer Research. Support for the research summarized in the present article came from the Canadian Institutes for Health Research, Janssen Ortho Canada, the Alzheimer Society of Canada, the Mathematics of Information Technology and Computer Science program of the Canadian Research Network, and the Fountain Innovation Fund of the Queen Elizabeth II Health Sciences Foundation. DementiaGuide Inc. has received funding from the Atlantic Canada Opportunities Agency, and the National Research Council of Canada, through the Industrial Research Assistance Program.

Author details

¹Centre for Health Care of the Elderly, Capital District Health Authority, 1421-5955 Veterans' Memorial Lane, Halifax, Nova Scotia, Canada, B3H 2E1. ²Divisions of Geriatric Medicine & Neurology, Department of Medicine, Dalhousie University, Halifax, Nova Scotia, Canada B3H 2E1.

Published: 25 June 2010

References

1. Biomarkers Definitions Working Group: **Biomarkers and surrogate endpoints: preferred definitions and conceptual framework.** *Clin Pharm Ther* 2001, **69**:89-95.
2. Streiner DL, Norman GR: **'Precision' and 'accuracy': two terms that are neither.** *J Clin Epidemiol* 2006, **59**:327-330.
3. McKhann GM: **Metachromatic leukodystrophy: clinical and enzymatic parameters.** *Neuropediatrics* 1984, **15**(Suppl):4-10.
4. Snowdon DA, Greiner LH, Mortimer JA, Riley KP, Greiner PA, Markesbery WR: **Brain infraction and the clinical expression of Alzheimer disease. The Nun Study.** *JAMA* 1997, **277**:813-817.
5. Neuropathology Group of the Medical Research Council Cognitive Function and Aging Study: **Pathological correlates of late-onset dementia in a multicentre, community-based population in England and Wales.** *Lancet* 2001, **357**:169-175.
6. Frieden JA: **The dynamics of international monetary systems: international and domestic factors in the rise, reign, and demise of the classical gold standard.** In *The Gold Standard in Theory and History*. 2nd edition. Edited by Eichengreen B, Flandreau M. New York: Routledge; 1997:207-229.
7. Katz R: **Biomarkers and surrogate markers: an FDA perspective.** *NeuroRx* 2004, **1**:189-195.

doi:10.1186/alzrt40

Cite this article as: Rockwood K: **Can biomarkers be gold standards in Alzheimer's disease?** *Alzheimer's Research & Therapy* 2010, **2**:16.