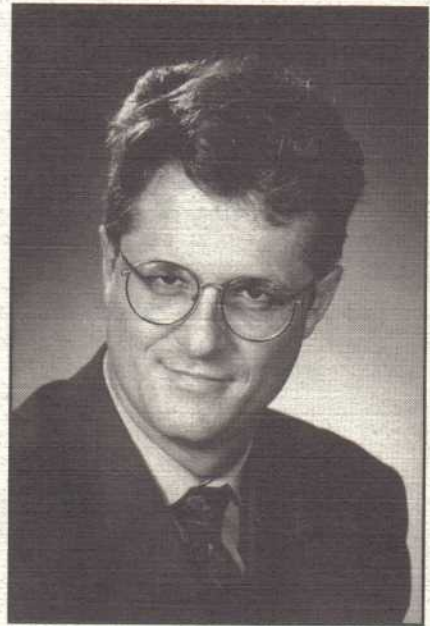


Health Care Outcome Measurement



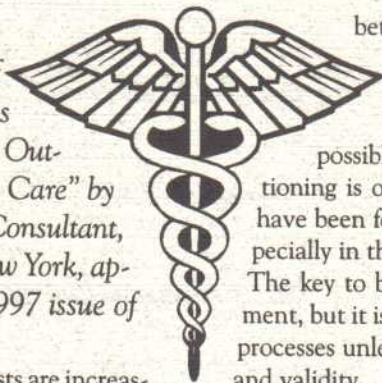
William P. Fisher, Jr. Ph.D.
LSU Medical Center, New Orleans

"The organizations that recognize the challenges, opportunities and rewards of measuring clinical outcomes will emerge as and remain market leaders." from "Clinical Outcomes: The New Driving Force in Health Care" by Raul A. Trillo, MD, Senior Health Care Consultant, Deloitte & Touche Consulting Group, New York, appearing on page 17 of the October 27, 1997 issue of American Medical News.

As everyone is well aware, health care costs are increasing at several times the general rate of inflation. Most health care consumers are also aware that health maintenance organizations (HMOs) are managing care in an effort to slow the spiraling costs, most usually by restricting access to care, as when referrals are required for specialist consultations, or when clinicians are required to follow procedural regimens in the care they provide.

What is less widely understood, however, is that HMOs and managed care produce, on average, only a one-time 7-9% reduction in costs, after which the increases continue unabated. Most approaches to cost reduction taken to date follow the model of quality control, in which the low-quality tail of a quality distribution is lopped off, with no overall change in the structure, process, or outcome of the care provided.

In contrast with the quality control approach is the quality assessment and improvement approach, in which the entire quality distribution is moved toward a higher standard. It is crucial at this point to recognize that costs and outcomes are opposite sides of the same coin. It is impossible to change anything that reduces costs without also affecting outcomes, and vice versa. The point is to be able to evaluate the relation



between cost and outcomes in ways that are sensitive to both the organization's mission to provide care and its bottom line.

Outcome measurement systems make it possible to show how much change in health or functioning is obtained per unit cost, and outcome measures have been focused on serving this accountability need, especially in the area of physical medicine and rehabilitation. The key to better outcomes per dollar is process improvement, but it is impossible to evaluate the effect of changes in processes unless outcomes are measured with high reliability and validity.

The vast majority of outcome measurement systems proposed to date mistakenly treat raw, ordinal summed scores as linear, interval measures. Accordingly, the various efforts underway ostensibly aimed at standardizing outcome measures in health care focus on the hopeless task of devising a single collection of items that will meet all users' needs. Though recognition of probabilistic measurement models in research publications is growing (see bibliography), there is not yet much widespread appreciation in health care for the strengths of models that 1) test data quality and the hypothesis that the variable is quantitative; 2) express each facet of the measurement design (item difficulties, person measures, rater harshness/leniency) in a common quality-assessed-and-improved metric; 3) accommodate missing data; 4) facilitate adaptive instrument administration, which adapts technology to the needs of people instead of vice versa; 5) remove from the measures rater and other identifiable and consistent bias factors that can be included in the model; and 6) provide a basis for standard metrics, i.e., universally-recognized, variable-specific quantities that can be read off any calibrated instrument shown to measure that variable.



R
E
H
A
B
I
L
I
T
Y
M
E
A
S
U
R
E
M
E
N
T

It is often instructive to observe where things have been if one desires a sense of where they are going. Outcome measurement research in health care employing Rasch's probabilistic models had its first applications in mental health and psychiatry, in the 1970s in Europe and North America (Hehl & Nussel, 1975, 1976; Kalinowski, 1985; Lewine, Fogg, & Meltzer, 1983; Maier & Philipp, 1986; Olsen & Savroe, 1984; Sørensen, Hansen, Andersen, et al., 1989). In the late 1970s or early 1980s, Ross Lambert, MD, an ophthalmologist at the Hines VA Hospital west of Chicago, and Benjamin D. Wright, PhD, became acquainted during early morning swims at a Hyde Park pool.

Lambert was involved in rehabilitating veterans suffering from low vision problems caused by accidents, diabetic retinopathy, or other problems. He needed an assessment tool that would enable therapists to document how well someone with severe visual impairments could perform travel activities, such as walking around at home, in the local neighborhood, in new places, as well as taking a bus or train, using an elevator, or shopping. University of Chicago graduate students, including Larry Ludlow, Matthew Schulz, Sheila Courington, David Zurakowski, Mark Wilson, Patrick Fisher, and this author worked as research assistants at Hines as a result of Lambert's interest in Rasch measurement.

In 1985, Lambert decided to become "double-boarded" and add a professional certification in physical medicine and rehabilitation to his ophthalmology certification. He became part of the first class of residents to rotate through Marianjoy Rehabilitation Hospital & Clinics, also in Chicago's western suburbs. At Marianjoy, Lambert learned that Medical Director, Richard Harvey, MD, had devised a rating-based functional assessment system, the Patient Evaluation Conference System, for monitoring the outcomes of care. Harvey took an immediate interest in testing data from the PECS system to see if they could meet the requirements for measurement specified in a Rasch model. He and Lambert used Wright's software to analyze the data. They presented the results to the Academy of Physical Medicine & Rehabilitation in 1987 (Harvey & Lambert, 1987; Lambert & Harvey, 1987; Lambert & Harvey, 1988; Lambert & Rao, 1989; Lambert & Wright, 1989; Lambert, Yokoo, Kilgore, et al., 1990).

Following the success of these initial analyses, Harvey brought in Burton Silverstein, PhD, in late 1987 to continue the work. Silverstein had just finished a post-doctoral fellowship at the University of Chicago. Harvey and Silverstein saw that the Rasch measurement research agenda held great potential for improving the PECS's capacity to support program evaluation and quality assessment applications, so in April,

1988, Karl Kilgore, PhD, was hired as Director of Research and Education at Marianjoy, and in August this author started as Research Associate. In 1989, Silverstein, Kilgore, and Fisher published a monograph on patient tracking and outcome assessment (Silverstein, Kilgore, & Fisher, 1989). Over the next several years, they together and separately published several articles on functional assessment in rehabilitation, and made many presentations on the topic.

With Harvey as editor and the submission of articles reporting advanced measurement research employing functional assessment instruments, the Archives of Physical Medicine and

Rehabilitation became the leader in rating scale measurement and practice among health care publications. A key moment arrived when the Archives published an article that criticized the use of ordinal rating scale data as though they were interval measures (Merbitz, et al., 1989) and concluded that rating scale data were incapable of providing a basis for the scientific measurement of outcomes. Several letters to the editor pointed out the possibilities for an enhanced scientific basis for rating scales that exist in Rasch's models, and the editors invited Wright and Linacre to write a special article expanding on this theme (Wright & Linacre, 1989).

After the 1989 Wright and Linacre article, research employing Rasch models began appearing as articles in the Archives and other journals (a sampling of the articles at hand includes: Cella, Lloyd, & Wright, 1996; Chang & Chan, 1995;

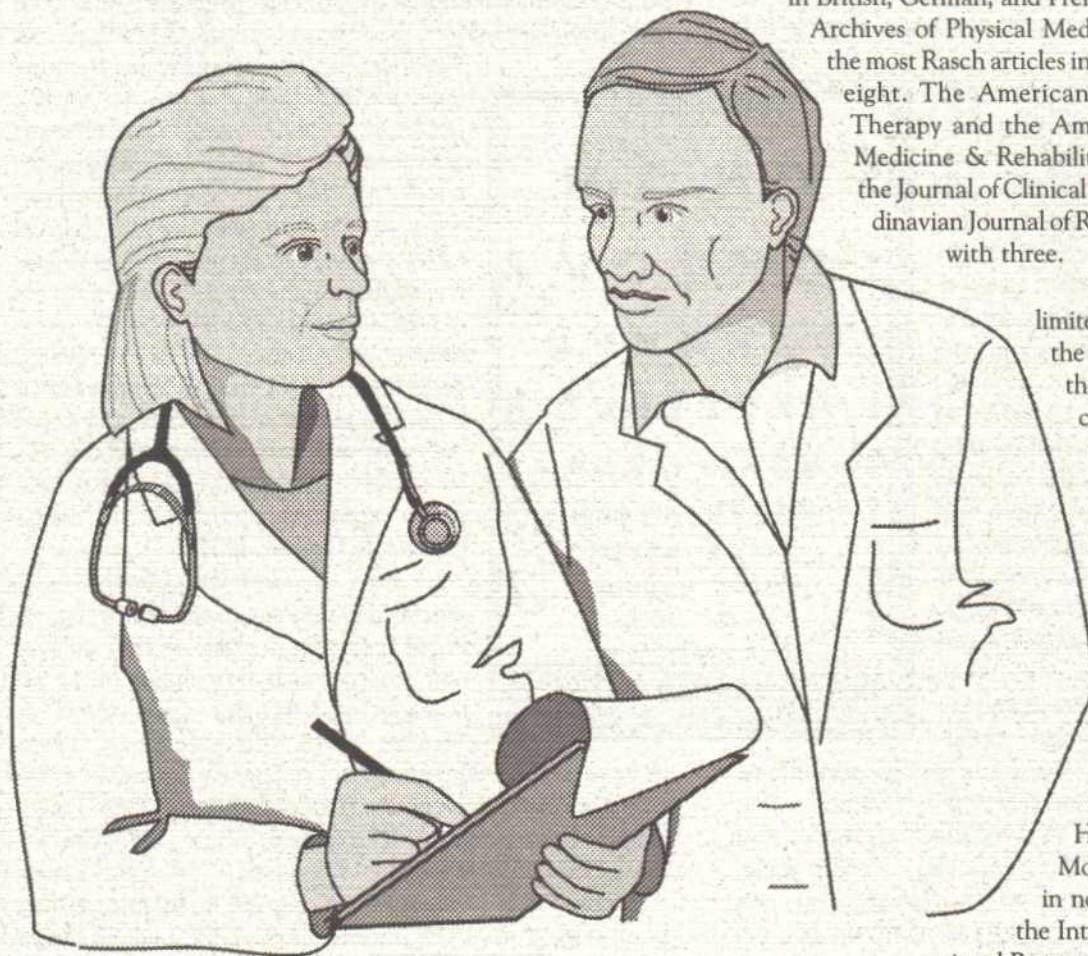
Daltroy, et al., 1992; Fisher, A., 1992, 1993; Fisher, W., 1993; Fisher & Fisher, 1993; Fisher, Harvey, & Kilgore, 1995; Fisher, Harvey, Taylor, et al., 1995; Granger & Wright, 1993; Grimby, et al., 1996; Haley & Ludlow, 1992a, 1992b; Haley, McHorney, & Ware, 1994; Heinemann, et al., 1994; Kilgore, Fisher, Silverstein, et al., 1993; Linacre, et al., 1994; Ludlow, Haley, & Gans, 1992; Lunz & Stahl, 1990, 1993; McArthur, Cohen, & Schandler, 1991; McHorney, Haley, & Ware, 1997; Pollack, Rheault, & Stoecker, 1996; Silverstein, Fisher, Kilgore, et al., 1992; Stucki, Daltroy, Katz, et al., 1996; Zhu & Cole, 1996), and not just as abstracts of annual meeting presentations. In 1991, a report on the Functional Independence Measure (FIM) employing Rasch models was made to the National Institute on Disability and Rehabilitation Research. The authors included Allen Heinemann, PhD, working at the Rehabilitation Institute of Chicago, and his colleagues Carl Granger, MD, and Byron Hamilton, PhD, of the Uniform Data System for Rehabilitation at the State University of New York in Buffalo, along with Wright and John Michael Linacre.

In 1993, the American Journal of Occupational Therapy published the proceedings of a 1991 conference sponsored by



the American Occupational Therapy Foundation and held at the University of Illinois-Chicago. Half of the papers elaborated on the scientific advantages of Rasch's models. Then in 1993, the journal *Physical Medicine and Rehabilitation Clinics of North America* published the proceedings of a 1992 conference hosted by Granger and Hamilton at SUNY-Buffalo; seven of the 13 articles were based on a Rasch analysis.

Since 1993, the research group at Marianjoy has moved to the Rehabilitation Foundation, Inc. (RFI), with Richard



Smith in charge of the measurement and evaluation work. Also in the last five years, the number and type of journals in health care publishing Rasch analyses has grown considerably. The *Journal of Clinical Epidemiology* has published three articles in the last several years, and a research report (Campbell, Kolobe, Osten, et al., 1995) employing a Rasch analysis in Physical Therapy was nominated as "the article of the year."

Researchers at Wayne State University, American University, and Indiana University have developed significant work in outcome measurement for physical and health education, especially as these concern persons with disabilities (Spray, 1987, 1990; Safrit, Cohen, Costa, 1989; Safrit, Zhu, Costa, et al., 1992; Zhu & Safrit, 1993; Cole, Wood, & Dunn, 1991; Zhu, 1996; Zhu & Cole, 1996; Zhu & Kurz, 1994). Although

this work situates itself within Item Response Theory, much of it, in fact, takes a strong measurement theory approach.

A MEDLINE search of the years 1993-1998 in the bibliographic database done in February, 1998, using the key word string, "Rasch analysis or Rasch measurement or Rasch model," produced 45 hits of articles appearing in 24 journals. Single articles have appeared in *Stroke*; *Aging*; *Pain*; *Neurology*; *Arthritis Care and Research*; *Biometrics*; and *Nutrition & Health*. Six articles appear in four Scandinavian journals, and one each in British, German, and French Canadian journals. The *Archives of Physical Medicine & Rehabilitation* has the most Rasch articles in the 1993-1998 period, with eight. The *American Journal of Occupational Therapy* and the *American Journal of Physical Medicine & Rehabilitation* both have five, with the *Journal of Clinical Epidemiology* and the *Scandinavian Journal of Rehabilitation Medicine* each with three.

The results of this search are limited to only what is included in the database. Not included was the 1997 special issue of *Physical Medicine & Rehabilitation: State of the Art Reviews*, edited by Richard Smith, which presents the proceedings of the First International Outcome Measurement Conference. Significant work in this area has also appeared in the Objective Measurement book series (Fisher, A., 1994; Ludlow & Haley, 1992; Ludlow & Haley, 1996; McArthur, Casey, Morrow, et al., 1992), as well as in non-medical journals, such as the *International Journal of Educational Research* (Fisher, A., et al., 1994).

To take advantage of Rasch's models for measurement we will need to establish the extent to which we can depend on these constructs as bases of comparison for the populations we serve. This calls for new ways of formulating research questions, reporting results, and collaborating, but most of all it requires a new awareness in the psychosocial sciences of the importance of metrology, the science of maintaining and improving the reference standard metrics through which we will most fully capitalize on scale-free measurement principles (Fisher, 1997a, 1997b, 1997c). For the latest on what's happening in the metrology movement among outcome measurement practitioners, be sure to attend the 2d International Outcome Measurement Conference at the University of Chicago, May 15-16.