Why fair tests are needed: a brief history*

Fair tests of treatments are those that take steps to obtain reliable information about treatment effects by reducing the misleading influences of biases and the play of chance. Why do we need such tests in health care? Have not doctors, for centuries, "done their best" for their patients? Sadly, health professionals in most if not all spheres of health care have harmed their patients inadvertently, sometimes on a very wide scale, because treatment decisions were not informed by what we now consider to be reliable evidence about the effects of treatments. Indeed, patients themselves sometimes harm other patients when, on the basis of untested theories and limited personal experiences, they encourage the use of treatments that turn out to be harmful.

The question is not whom we can blame but whether we can reduce the harmful effects of inadequately tested treatments. And it seems we can—to a great extent—first, by avoiding applying untested theories about the effects of treatment in practice, and second, by wider use of fair tests of treatments.

WHY THEORIES ABOUT TREATMENTS MUST BE TESTED IN PRACTICE

People have often been harmed because treatments have been based on theories about how disease should be treated without being tested in practice. For example, for centuries we believed the theory that illnesses were caused by "humoral imbalances." So patients were bled and purged, made to vomit and take snuff, in the belief that this would end the supposed imbalances. As long ago as the 17th century, Jan Baptista van Helmont (Figure 1) was impertinent enough to challenge the medical authorities of the time to assess the validity of their theories in a fair test of treatment (1).

Figure 1. Jan Baptista van Helmont.

BLOODLETTING

By the beginning of the 19th century, British military surgeons had begun to show the harmful effects of bloodletting (2, 3), and a few decades later, the practice was again challenged by the Parisian physician Pierre Louis (4). Yet at the beginning of the 20th century, orthodox practitioners in Boston who were not using bloodletting to treat pneumonia were still being judged negligent (5). Indeed, without citing supporting evidence, Sir William Osler, one of the most influential physicians of the past century, advised his readers that "during the last decades we have certainly bled too little. Pneumonia is one of the diseases in which a timely venesection may save life. To be of service it should be done early. In a full-blooded, healthy man with a high fever and bounding pulse the abstraction of from twenty to thirty ounces of blood is in every way beneficial" (6).

SIDS AND SLEEPING POSITION

Although the need to test theories in practice has been recognized for hundreds of years, this important principle is still too often ignored. For instance, based on an untested theory, Benjamin Spock, the influential American child health expert, informed the readers of his best-selling book, *Baby and Child Care*, that a disadvantage of babies sleeping on their backs was that if they vomited, they would be more likely to choke. Dr Spock therefore advised his millions of readers to encourage babies to sleep on their tummies (7). We now know that this advice, apparently rational in theory, led to crib death in tens of thousands of infants (8).

CLASS 1 ANTIARRHYTHMICS

The use of drugs to prevent heart rhythm abnormalities in people who have had myocardial infarctions provides another example of the dangers of applying untested theory in practice. Because heart rhythm abnormalities are associated with increased risk for sudden death, the theory suggested that class 1 antiarrhythmics should reduce these early deaths (Figure 2). However, years after the drugs had been licensed and adopted in practice, 2 systematic reviews of randomized trials showed that they actually increase the risk for sudden death after heart attack. Indeed, it has been estimated that, at the peak of their use in the late 1980s, these drugs may have killed as many as 70 000 people every year in the United States alone (9)—many more than the total number of Americans who died in the Vietnam War.

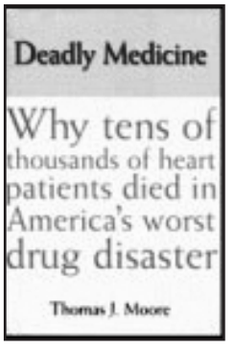

Figure 2. *Deadly Medicine* by Thomas J. Moore.

Misplaced confidence in the validity of a theory as a guide to practice has also resulted in some treatments being rejected inappropriately because researchers did not believe that they could work. Theories based on the results of animal research, for example, sometimes correctly predict the results of treatment tests in humans—but not always (10). Based on the results of experiments in rats, some researchers became convinced that there was no point in giving thrombolytic drugs to patients who had had heart attacks more than 6 hours previously. If such patients had not participated in some of the fair tests of these drugs, we would not know that thrombolytic drugs can be beneficial (11).

**Adapted from the James Lind Library, a resource for the public, illustrating the evolution of fair tests of treatments (www.jameslindlibrary.org)*

Observations in clinical practice or in laboratory and animal research may suggest that particular treatments will or will not benefit patients; but as these and many other examples make clear, it is essential to use fair tests to find out whether, in practice, these treatments do more good than harm or vice versa.

WHY TESTS OF MEDICAL TREATMENTS MUST BE FAIR TESTS

Failure to test theories about treatments in practice is not the only preventable cause of treatment tragedies. These have also occurred because the tests used to assess the effects of treatments have been unreliable and misleading. Fair tests entail taking steps to reduce the likelihood that we will be misled either by the effects of biases or by the play of chance.

For example, theory suggested that giving the synthetic sex hormone, diethylstilboestrol (DES), to pregnant women who had previously had miscarriages and stillbirths would increase the likelihood of a successful outcome of later pregnancies. Some of the tests done had not adequately controlled for biases and suggested that the theory was correct: The drug reduced miscarriages and stillbirths. Although other "fair" tests had suggested that DES was useless, the unreliable evidence, together with aggressive marketing, led to DES being prescribed to millions of pregnant women over the next few decades. The consequences were disastrous: Some of the daughters of women who had been prescribed DES developed cancer of the vagina, and some developed such health problems as malformations of the reproductive organs and infertility (12).

Problems resulting from inadequate tests of treatments continue to occur. Again, as a result of unreliable evidence and aggressive marketing, millions of women were persuaded to use hormone replacement therapy (HRT), not only because it could reduce unpleasant menopausal symptoms, but also because it was claimed that it would reduce their chances of having heart attacks and strokes (Figure 3). When these claims were assessed in fair tests, the results showed that far from reducing risks, HRT actually *increased* risk for heart attacks and strokes in addition to causing other undesirable effects (13, 14).

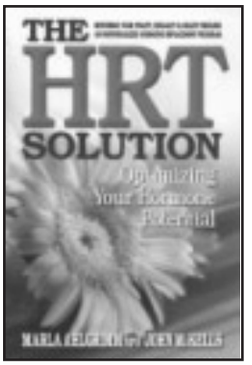

Figure 3. *The HRT Solution: Optimizing Your Hormone Potential* by Marla Ahlgrmm and John M. Kells.

These examples of the need for fair tests of treatments are a few of the many hundreds that illustrate how treatments can do more harm than good. Improved general knowledge about fair tests of treatments is needed so that—along with a healthy dose of skepticism—we can more critically assess claims about the effects of treatments. That way, we will all become more able to judge which treatments are likely to do more good than harm.

Iain Chalmers
Editor, James Lind Library
Oxford, England, UK

References

1. **van Helmont JB.** Oriatrike, or Physick Refined: the Common Errors Therein Refuted and the Whole are Reformed and Rectified [translated by J Chandler]. London: Lodowick-Loyd;1662:526.
2. **Robertson R.** Observations on Diseases Incident to Seamen, Whether Employed on, or Retired from Actual Service, for Accidents, Infirmities or Old Age. London: Cadell and Davies;1807.
3. **Hamilton AL.** Dissertatio Medica Inauguralis de Synocho Castrensi (Inaugural medical dissertation on camp fever). Edinburgh: J Ballantyne;1816.
4. **Louis PC.** Recherches sur les Effets de la Saignée dans Quelques Maladies Inflammatoires et sur l'action de l'émétique et des Vésicatoires dans la Pneumonie (Researches on the Effects of Bloodletting in Some Inflammatory Diseases, and on the Influence of Tartarized Antimony and Vesication in Pneumonitis). Paris: Baillière;1835.
5. **Silverman W.** In: Chalmers I, McIlwaine G, eds. Perinatal Audit and Surveillance. London: Royal College of Obstetricians and Gynaecologists; 1980:110.
6. **Osler W.** Principles and Practice of Medicine. London: Appleton;1892:530.
7. **Spock B.** Baby and Child Care. New York: Pocket Books; 1966:163-4.
8. **Gilbert R, Salanti G, Harden M, See S.** Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. Int J Epidemiol. 2005;34:874-87.
9. **Moore TJ.** Deadly Medicine. New York: Simon and Schuster;1995
10. **Pound P, Ebrahim S, Sandercock P, Bracken MB, Roberts I.** Where is the evidence that animal research benefits humans? BMJ. 2004;328:514-7.
11. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomised trials of more than 1000 patients. Fibrinolytic Therapy Trialists' (FTT) Collaborative Group. Lancet. 1994;343:311-22.
12. **Apfel RJ, Fisher SM.** To Do No Harm: DES and the Dilemmas of Modern Medicine. New Haven: Yale University Press;1984.
13. **McPherson K.** Where are we now with hormone replacement therapy? BMJ. 2004;328:357-8.
14. **Rossouw JE, Anderson GL, Prentice RL, et al.** Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA. 2002; 288:321-33.