# EBM notebook

## Why fair tests are needed: a brief history

Why do we need fair tests of treatments in health care? Have not doctors, for centuries, "done their best" for their patients? Sadly, there are many examples of doctors and other health professionals harming their patients because treatment decisions were not informed by what we consider now to be reliable evidence about the effects of treatments. With hindsight, health professionals in most if not all spheres of health care have harmed their patients inadvertently, sometimes on a very wide scale. Indeed, patients themselves have sometimes harmed other patients when, on the basis of untested theories and limited personal experiences, they have encouraged the use of treatments that have turned out to be harmful.

The question is not whom we might blame, but whether the harmful effects of inadequately tested treatments can be reduced. They can, to a great extent, firstly, by avoiding applying untested theories about the effects of treatment in practice, and secondly, by wider use of fair tests of treatments. What are fair tests of treatments? They are tests that take steps to obtain reliable information about treatment effects by reducing the misleading influences of biases and the play of chance.

### WHY THEORIES ABOUT TREATMENTS MUST BE TESTED IN PRACTICE

#### Van Helmont 1662

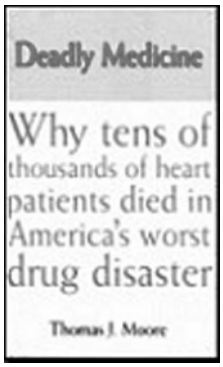
People have often been harmed because treatments have been based on our theories about how disease should be treated without testing the theories in practice. For example, for centuries we believed the theory that illnesses were caused by "humoral imbalances." So patients were bled and purged, made to vomit and take snuff, in the belief that this would end the supposed imbalances. As long ago as the 17th century, a Flemish doctor 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]

By the beginning of the 19th century, British military surgeons had begun to show the harmful effects of bloodletting[2,3] A few decades later, the practice was also challenged by the Parisian physician Pierre Louis (1835).[4] Yet at the beginning of the 20th century, orthodox practitioners in Boston, USA, 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 last 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 recognised 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 the cot deaths of tens of thousands of infants.[8]

#### Class I anti-arrhythmics




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 an increased risk of sudden death, the theory suggested that these drugs should reduce such early deaths. However, years after the drugs had been licensed and adopted in practice, 2 systematic reviews of randomised trials showed that they actually increase the risk of sudden death after heart attack. Indeed, it has been estimated that, at the peak of their use in the late 1980s, they may have been killing as many as 70 000 people every year in the US alone[9]—many more than the total number of Americans who died in the Vietnam War.

Misplaced confidence in the validity of 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 this is not always the case.[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 experienced heart attacks more than 6 hours previously. Had not such patients participated in some of the fair tests of these drugs we would not know that they can benefit from treatment.[11]

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: that 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 cancers of the vagina, and other children had other health problems, including malformations of their 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. When these claims were assessed in fair tests the results showed that, far from reducing the risks of heart attacks and strokes, HRT increases the risks of these life threatening conditions, as well as having other undesirable effects.[13 14]

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

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

IAIN CHALMERS
*Editor, James Lind Library, Oxford, UK*

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 the diseases incident to seamen*. 2nd edn.Vol.1. London: for the author, 1804.
3 Hamilton AL. *Dissertatio medica inauguralis de synocho castrensi* (Inaugural medical dissertation on camp fever). Edinburgh: J Ballantyne, 1816.
4 Louis PCA. *Recherches sur les effets de la saignée dans quelques maladies inflammatoires et sur l'action de l'émétique et des vésicatoires dans la pneumonie*. Paris: Librairie de l'Académie royale de medicine, 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*, 165th printing. New York: Pocket Books, 1966:163–4.
8 Gilbert R, Salanti G, Harden M, *et al*. 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, *et al*. Where is the evidence that animal research benefits humans? *BMJ* 2004;**328**:514–7.
11 Fibrinolytic Therapy Trialists' Collaborative Group. 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. *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 Women's Health Initiative Steering Committee. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: the Women's Health Initiative randomized controlled trial. *JAMA* 2004;**291**:1701–12.