

[Home](#)[Contents](#)[jameslindlibrary.org](http://jameslindlibrary.org)

## Why fair tests are needed

### Trying to do more good than harm

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 ([click here for examples](#)). 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 whether we must blame these people, but whether the harmful effects of inadequately tested treatments can be reduced. They can, to a great extent.

Acknowledging that treatments can sometimes do more harm than good is a prerequisite for reducing unintended harm (Gregory 1772; [Haygarth 1800](#); Fordyce 1802; [Behring 1893](#)). We then need to be more ready to admit uncertainties about treatment effects, and to promote tests of treatments to adequately reduce reduce uncertainties. Such tests are fair tests.

### Why theories about the effects of treatments must be tested in practice

People have often been harmed because treatments have been based only on theories about how disease should be treated, without testing how the theories played out in practice. For example, for centuries people believed the theory that illnesses were caused by 'humoral imbalances', and patients were bled and purged, made to vomit and take snuff, in the belief that this would end the supposed imbalances still, as long ago as the 17th century, a lone Flemish doctor was impertinent enough to challenge the medical authorities of the time to assess the validity of their theories by proposing a fair test of the results of their unpleasant treatments ([Van Helmont 1662](#)).



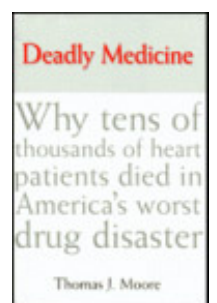
By the beginning of the 19th century, British military surgeons had begun to show the harmful effects of bloodletting for treating "fevers" (Robertson 1804; [Hamilton 1816](#)). A few decades later, the practice was also challenged by a Parisian physician ([Louis 1835](#)). Yet at the beginning of the 20th century orthodox practitioners in Boston, USA, who were not using bloodletting to treat pneumonia were still being judged negligent (Silverman 1980). Indeed, Sir William Osler, one of the most influential medical authorities in the world who was generally cautious about recommending unproven treatments, 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 [bleeding] 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" (Osler 1892).



Although the need to test the validity of theories in practice was been recognized by some people at least a millennium ago ([Ibn Hindu 10th-11th century](#)), this important principle is still too often ignored. For instance, based on 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 (Spock 1966). We now know that this advice, apparently rational in theory, led to the cot deaths of tens of thousands of infants (Gilbert et al. 2004).



The use of drugs to prevent heart rhythm abnormalities in people having heart attacks provides another example of the dangers of applying untested theory in practice. Because heart rhythm abnormalities are associated with an increased risk of early death after heart attack, the theory was that these drugs would reduce such early deaths. Just because a theory seems reasonable doesn't mean that it is necessarily right, however. Years after the drugs had been licensed and adopted in practice, it was discovered 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 United States alone (Moore 1995) – many more than the total number of Americans who died in the Vietnam War.



On the other hand, misplaced confidence in theoretical thinking 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. Based on the results of experiments in rats, some researchers became convinced that there was no point in giving clot-dissolving drugs to patients who had experienced heart attacks more than six 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 (Fibrinolytic Therapy Trialists' Collaborative Group 1994).

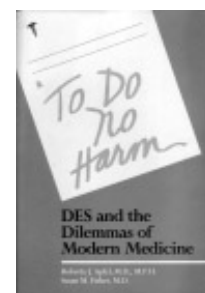


Observations in clinical practice or in laboratory and animal research may suggest that particular treatments will or will not benefit patients; but as these and many other examples make clear, it is essential to use fair tests to find out whether, in practice, these treatments do more good than harm,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

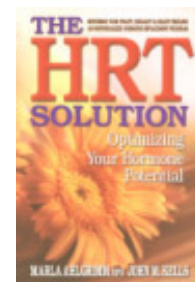
### Why tests of medical treatments must be fair tests

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 by the effects of [biases](#) and the [play of chance](#).

For example, in the 1950s, theory and poorly controlled tests yielding unreliable evidence suggested that giving a 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. Although fair tests had suggested that DES was useless theory and 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 (Apfel and Fisher 1984).



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 (McPherson 2004).



These examples of the need for fair tests of treatments are a few of many 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.

The principles of fair tests of treatments have been evolving for centuries - and they continue to evolve.

**Cite as:** Editorial commentary (2007). Why fair tests are needed. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Next essay:** [Why comparisons are essential](#)

**Select other essay:**

## References

Apfel RJ, Fisher SM (1984). To do no harm: DES and the dilemmas of modern medicine. New Haven, Ct: Yale University Press.

Behring, Boer, Kossel H (1893). Zur Behandlung diphtheriekranker Menschen mit Diphtherieheilserum. Deutsche Medicinische Wochenschrift 17: 389-393.

Fibrinolytic Therapy Trialists' Collaborative Group (1994). 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: 311-322.

Fordyce G (1802). A second dissertation on fever. London: J Johnson

Gilbert R, Salanti G, Harden M, See S (2005). Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. International Journal of Epidemiology 34: 874-87.

Gregory J (1772). Lectures on the duties and qualifications of a physician. London: Strahan and Cadell.

Hamilton AL (1816). Dissertatio Medica Inauguralis De Synocho Castrensi (Inaugural medical dissertation on camp fever). Edinburgh: J Ballantyne.

Haygarth J (1800). Of the imagination, as a cause and as a cure of disorders of the body: exemplified by fictitious tractors, and epidemical convulsions. Bath: R. Crutwell.

Ibn Hindu (10th-11th century CE; 4th-5th century AH). Miftah al-tibb wa-minhaj al-tullab [The key to the science of medicine and the students' guide].

Louis PCA (1835). 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 médecine.

McPherson K (2004). Where are we now with hormone replacement therapy? BMJ 328: 357-358.

Moore TJ (1995). Deadly Medicine. New York: Simon and Schuster.

Osler W (1892). Principles and Practice of Medicine. London: Appleton, p 530.

Robertson R (1804). Observations on the diseases incident to seamen, 2nd edn. Vol. 1, London: for the author.

Silverman W (1980). In: Chalmers I, McIlwaine G (eds). Perinatal Audit and Surveillance. London: Royal College of Obstetricians and Gynaecologists, 1980: 110.

Spock B (1966). Baby and Child Care. 165th printing. New York: Pocket Books, pp 163-164.

van Helmont JB (1662). Oriatrike, or physick refined: the common errors therein refuted and the whole are reformed and rectified [translated by J Chandler]. Lodowick-Loyd: London, p 526.

[Home](#)

[Contents](#)

[Comments welcome](#)