

**CONFRONTING THE ILLUSION OF CERTAINTY:  
WHAT HAS CLINICAL EPIDEMIOLOGY CONTRIBUTED?**

---

**INAUGURAL ADDRESS: PROF JIMMY VOLMINK SEPT 2006**



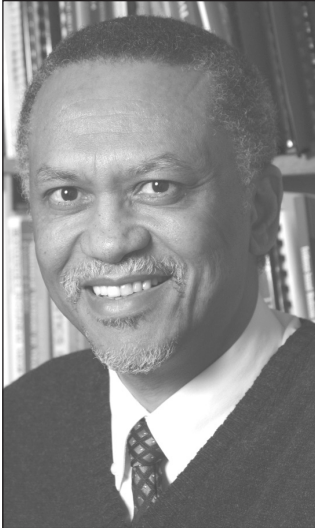
---

UNIVERSITEIT • STELLENBOSCH • UNIVERSITY

---



## ABOUT THE AUTHOR



Jimmy Volmink was born in Cape Town in 1957. He matriculated from Alexander Sinton High School in Athlone before obtaining his BSc (Biochemistry) and MBChB degrees from the University of Cape Town. He later obtained a DCH from the SA College of Medicine, a MPH from Harvard University and a DPhil from the University of Oxford.

He is Professor and Deputy Dean (Research) in the Faculty of Health Sciences, Stellenbosch University, and Director of the South African Cochrane Centre, Medical Research Council. Previously he has held appointments as Professor and Chair of Primary Health Care at the University of Cape Town and as Director of Research and Analysis at the Global Health Council, Washington DC.

Professor Volmink has a special interest in the evaluation of the effects of health care interventions. He has extensive experience in the application of randomised controlled trials, systematic reviews and meta-analysis in evaluating strategies and therapies for the control of tuberculosis, HIV/AIDS and cardiovascular disease. He also works with policy-makers and practitioners to promote the use of research evidence in decisions about health care. He has authored more than 100 peer reviewed journal articles, technical reports and book chapters.

He serves on the committees and advisory boards of a number of international organisations, including the World Health Organisation, Cochrane Collaboration, Wellcome Trust, Doris Duke Charitable Foundation and the International Clinical Epidemiology Network (INCLEN). He is also a member of the advisory boards of the on-line journals BioMed Central and PLoS Clinical Trials and of the evaluation board of the Faculty of 1 000 Medicine. In addition, he is a member of the Council of the Academy of Science of South Africa and chairs the Western Cape Clinical Guidelines Advisory Committee.

His hobbies include walking, reading and listening to jazz and classical music.

## ACKNOWLEDGEMENTS

This lecture is dedicated to many family members, friends and colleagues who have contributed to my personal and professional development in one way or another. This evening, I especially want to honour my late mom and dad, Johanna and James Volmink, who taught me the value of education and who provided the nurturing environment in which my potential could flourish; my wife, Blossom, without whose unfailing care and support few of my achievements would have been possible and my children, Lauren, Natalie and Michael, for being a constant source of pride and joy. I would also like to thank my mentors, Professors Wieland Gevers and Peter Folb – teachers at UCT during the dark days of apartheid, who always treated me with respect and fairness, and who continue to take a special interest in my career development. I owe a huge debt of gratitude to Sir Iain Chalmers, beloved friend and mentor, for profoundly influencing my thinking about health care and for introducing me to the global community of scholars in the field of clinical epidemiology. Finally, I wish to gratefully acknowledge my friends, Professors Bongani Mayosi and George Swingler and Dr Nandi Siegfried, whose passion for evidence-based practice is a continuous source of inspiration.

-----  
*Confronting the illusion of certainty:  
What has clinical epidemiology contributed?*

Inaugural lecture delivered on 19 September 2006

Prof J Volmink

Deputy Dean (Research)

Faculty of Health Sciences, Stellenbosch University

|                       |   |
|-----------------------|---|
| Editor:               | Mattie van der Merwe  |
| Language editing:     | Dr Edwin Hees   |
| Design:               | Heloïse Davis   |
| Photograph of author: | Jacques Botha, <a href="http://www.jbphoto.co.za">www.jbphoto.co.za</a> |
| Printing:             | Africa Digital Printing Services  |

ISBN: 0-7972-1132-2

# CONFRONTING THE ILLUSION OF CERTAINTY: WHAT HAS CLINICAL EPIDEMIOLOGY CONTRIBUTED?

## INTRODUCTION

Benjamin Franklin, famous polymath of the 18th century and one of the Founding Fathers of the United States, once said, “*In this world nothing is certain except death and taxes*”. Strangely enough, while we intuitively accept that there are few things in life one can be absolutely sure of, our minds nonetheless seem to have a strong inclination towards creating impressions of certainty (Gigerenzer 2002). Perhaps, because of our human need to make sense of life events, we prefer to see structure instead of chaos, cause and effect instead of chance, and certainty rather than uncertainty. As a result incomplete or ambiguous information detected by our senses is often ‘sold’ to our consciousness as a definite product (Gigerenzer 2002). This process is strongly influenced by our prior beliefs, which means that what we see is often determined by our mental model of what should be. In a phrase – “believing is seeing!” Have a look at this picture (Figure 1). Do you see Bill Clinton and Al Gore? Wrong! It’s Bill Clinton and Bill Clinton. Because you are so used



to seeing pictures of the former US President accompanied by his Vice-President, your mind in an attempt to make sense of the visual image ignored the fact that the faces in the picture are in fact identical. Things are not always what they seem to be. Sometimes certainty can be nothing but an illusion.

Given that we seem to be hard-wired for certainty, we might reflect for a moment on what value this may hold for us. Firstly, certainty can play a role in promoting our integration into a particular social group or culture. The desire for acceptance into a group may necessitate our adopting certain beliefs, values or practices without question. Certainty may also bring us hope or comfort in times of hardship or suffering. Think, for example, of the consolation in bereavement we derive from religious beliefs about the afterlife. A further benefit of certainty is its effect on our performance. The sportsman who convinces himself of his ability to win, finds that he is able to mobilise his inner resources and perform at a higher level than one who is gripped by self-doubt. You may be able to add to this list of benefits.

But certainty also has a dark side. Some of the most

dangerous people in history have been those who were so certain of their special status or beliefs that they were moved to marginalise, injure or eliminate people who differed from them. Hitler, for example, was absolutely certain that his actions were needed to rid the world of evil. George Bush felt sure that God told him to invade Iraq to find and destroy weapons of mass destruction. The rest is history.

## CERTAINTY AND UNCERTAINTY IN HEALTH CARE

What about certainty in health care? Let’s explore two examples from antiquity to set the scene. Colonic cleansing by means of enemas was used in Egypt as far back as 1 500 BC and the practice was known to be widespread in the ancient world, including China, India and Babylon. It was believed this treatment was necessary to rid the human body of toxins acquired from food. According to the Roman Naturalist, Pliny, the Egyptians got the idea from the ibis that was frequently seen using its long beak to wash the inside of its body with infusions of water from the River Nile (Editorial 1946).

In pre-revolutionary France enemas were used on a daily basis after dinner. They were thought to be good for skin complexion, maintenance of health and as a treatment for a variety of diseases. Louis XIV, for instance, apparently used enemas about 2 000 times during his lifetime and as a result stayed healthy for his entire life (or so it is believed). A large number of concoctions have been used in enemas over the years, such as salt water, bile, vinegar, wine, beer, oil, honey and even coffee!

By the beginning of the 20th century scientific evidence had accumulated that ‘autointoxication’ – the theory providing the rationale for colonic cleansing – was seriously flawed. Evidence was also available that the risks of the procedure outweighed any potential benefits. As a result the medical profession led by the American Medical Association strongly discouraged its use (Ernst 1997).

Yet today the practice remains common in many Western countries and, in fact, there has even been a resurgence of interest in this treatment in recent years. Alternative health practitioners in the US charge around \$100 for a colonic irrigation and many celebrities are queuing up to have the procedure. Modern-day indications for colon cleansing comprise a wide array of con-

ditions, such as alcoholism, allergies, arthritis, asthma, backache, bad breath, bloating, coated tongue, colitis, constipation, fatigue, gas, headache, hypercholesterolemia, hypertension and skin problems.

The strong belief in the value of colonic therapy is summed up in the report I found on one website:

*"The life of 109-year-old Norman W Walker, DSc, PhD, is testimony that colon cleansing is essential to your health. Dr Walker, ....stated emphatically, 'It is impossible when we eat two, three or more meals in a day not to have residue accumulating in the colon. Not to cleanse the colon is like having the entire garbage collecting staff in your city go on strike for days on end!'"*

(<http://swiftweb.com/ha/colonic.html>;  
accessed 26 August, 2006).

Before commenting on this example I would like to mention another – bloodletting (Trohler 2000). During the 19th century bloodletting was used as a cure for most infections. For instance, during the 1832 cholera epidemic in France, patients were starved and subjected to bloodletting by venesection and leeching. The practice of bloodletting was based on the theory that redness, heat and swelling seen in cases of inflammation were due to an excess of blood that led to a build-up of pressure. It seemed logical that the removal of blood would lessen the pressure and cure the inflammation. If the swelling was local, blood could be removed by leeches. In the case of general inflammation manifesting as a red face, general fever and strong pulse, venesection was needed to remove larger amounts of blood.

In 1835 a Parisian physician Pierre Louis published the results of his systematic observations of groups of patients and concluded (somewhat heretically) that bloodletting was not all it had been trumped up to be. Despite these findings the practice of bloodletting continued for a further 100 years! Even Sir William Osler, one of the most influential doctors in history, wrote in 1892:

*"... 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"*

(Osler 1892, cited in the James Lind Library  
<http://www.jameslindlibrary.org/>).

Today we know that this practice was responsible for killing hundreds of thousands of people!

What lessons may be learnt from these examples? Firstly, they illustrate the enormous power of personal testimony. People who subject themselves to colonic cleansing 'know' they are healthier and live longer, and experienced clinicians 'knew' that leeches were good for their patients – they had seen it with their own eyes! While it can be tempting to draw conclusions from personal experience or anecdotes, it must be kept in mind that anecdotal evidence suffers from serious limitations. Anecdotes can tell us that a certain outcome can occur but provide no information on how frequently that outcome occurs or whether an intervention can influence the frequency of the outcome. Secondly, these examples show how personal experience can collude with theory in leading us astray. In both situations an illusion of certainty was created by clinical observations being filtered through a distorted lens of incomplete or faulty understanding of the functions of the human body. Lastly, we learn something about the phenomenon of "positive bias." When a treatment is being evaluated there are at least three parties with an interest in the outcome – the patient, the clinician and those that profit commercially from the use of the treatment – and all three parties want to believe (though for different reasons) that the treatment works! In this situation evidence that a given treatment may be ineffective or harmful may be most unwelcome.

Given the inherent biases that bedevil the process of evaluation of care, it should be clear that specific steps are necessary to avoid the dangers of being misled. Historically, clinical epidemiology is the field most directly concerned with this challenge.

## CLINICAL EPIDEMIOLOGY

Most simply put, clinical epidemiology is the use of scientific methods in the evaluation of clinical practice. The field involves the study of diagnostic methods, disease prognosis, the benefits and risks of health care interventions, and factors that influence professional practice. Time, this evening, allows only for some limited remarks on treatment evaluation.

It should be declared at the start that clinical epidemiology is rooted in the scientific ideas of Francis Bacon, the 16th-century English philosopher who introduced the school of inductive reasoning (Vickers B, 2002). Bacon taught that truth should be sought through observation, experiment and testing hypotheses (empirical methods) in the real world and not on theory or "higher insight." Before Bacon deductive logic – i.e. reason based on first principles (rationalism) – was regarded as a sufficient basis for obtaining knowledge.

Importantly, Bacon pointed to the dangers of bias in

the pursuit of knowledge. He spoke of the need for investigators to free their minds of “idols” before starting induction and mentioned four kinds: 1) “Idols of the Tribe”, which were common to the human race; 2) “Idols of the Den”, which were peculiar to the individual; 3) “Idols of the Marketplace” from the misuse of language (which today might be referred to as “spin”; and 4) “Idols of the Theatre” from the abuse of authority. As we shall see later, these concepts are as relevant today as they were in the 1500s.

## THE EVOLUTION OF EMPIRICAL METHODS FOR EVALUATING HEALTH CARE

I will now briefly trace the evolution of empirical methods for evaluating treatments in health care. The James Lind Library which has brilliantly documented this development under the title of “fair tests of treatments in health care” will serve as the primary source for the material presented in this section (<http://www.jameslindlibrary.org/>).

### *Learning to make comparisons*

In medicine the first study involving a comparative (controlled) evaluation of treatment was conducted in 1747. At the time thousands of British sailors were dying at sea from scurvy, a disease that we know today is the result of a lack of vitamin C. Then, the cause of scurvy was not known and many treatments were being used to cure the disease. James Lind, a naval surgeon, decided to conduct a controlled trial to evaluate six of these therapies. He took 12 patients who were at a similar stage of scurvy, had a similar diet and were housed in similar parts of the ship on which he served and allocated them in pairs to the following treatments: a quart of cider a day; two spoonfuls of vinegar three times a day; 25 gutts of elixir vitriol three times a day; half a pint of sea-water daily; a concoction of nutmeg, mustard and garlic three times a day; and two oranges and a lemon every day. Lind’s trial clearly demonstrated what the best treatment was. He reported:

*“the consequence was, that the most sudden and visible good effects were perceived from the use of oranges and lemons; one of those who had taken them, being at the end of six days fit for duty”*

(Milne and Chalmers 2004).

Sadly, it took another 50 years for the Admiralty to sanction the provision of lemon juice in the navy. Within two years scurvy was virtually eliminated from the Royal Navy (Milne and Chalmers 2004).

Control groups are important for conducting reliable assessments of the effects of treatment, because they help us take account of two common situations: 1) people sometimes recover from illness without any treatment at all, and 2) natural fluctuations in the course of a disease may make it seem that a treatment is helping or hurting when in reality it is having no effect at all. Comparing the experience of those receiving a particular treatment with that of people not so treated therefore helps reduce the likelihood of falsely attributing a particular clinical outcome to the effect of treatment.

Claude Bernard, the famous French physiologist, underscored the importance of this principle in 1866:

*“Comparative experience is a prerequisite for experimental and scientific medicine, otherwise the physician may walk at random and become the sport of a thousand illusions”*

(cited in Trohler 2000).

### *Making unbiased comparisons*

While the use of controls was an important step forward in the evaluation of care, a further challenge needed to be met. A way needed to be found to ensure that people in comparison groups were similar enough at baseline so that any differences in clinical outcomes could be confidently attributed to treatment rather than to other factors, such as differences in illness severity between the groups.

Early methods involved allocating patients to experimental or control groups using clinical judgment or alternate allocation methods (using for example the day of the week), but none of these methods was good enough. The major advance finally came in the 1940s, when the medical statistician Sir Austin Bradford Hill succeeded in convincing the British MRC to use a truly random method for allocating patients with pulmonary tuberculosis to receive either streptomycin or placebo. Bradford Hill’s method was innovative for two reasons: 1) chance alone decided on who got what treatment; and 2) investigators could not tell what treatment groups patients had been assigned to until they had been admitted to the study. This first known randomised controlled trial was published in the BMJ in 1948 and to this day that particular study design remains the ‘gold standard’ for evaluating interventions. The reason for this is simple: randomisation is the only means we have for ensuring that all known and unknown factors influencing clinical outcomes are equally distributed between groups receiving alternative forms of treatment and thus for ensuring the comparison of like with like.



Of course, other methods were introduced at various times that have also contributed to reducing bias in treatment evaluation. Placebos, for instance, had been used in Russia as early as the 1800s to distinguish between the results of active treatment and the non-specific effects of “dummy” treatments. The method of “blinding” i.e. preventing providers, patients or those assessing the outcome of treatment from knowing the actual treatment being administered was known to have been used from the late 18th century onwards. But none of these strategies were as important as randomisation.

### Dealing with the play of chance

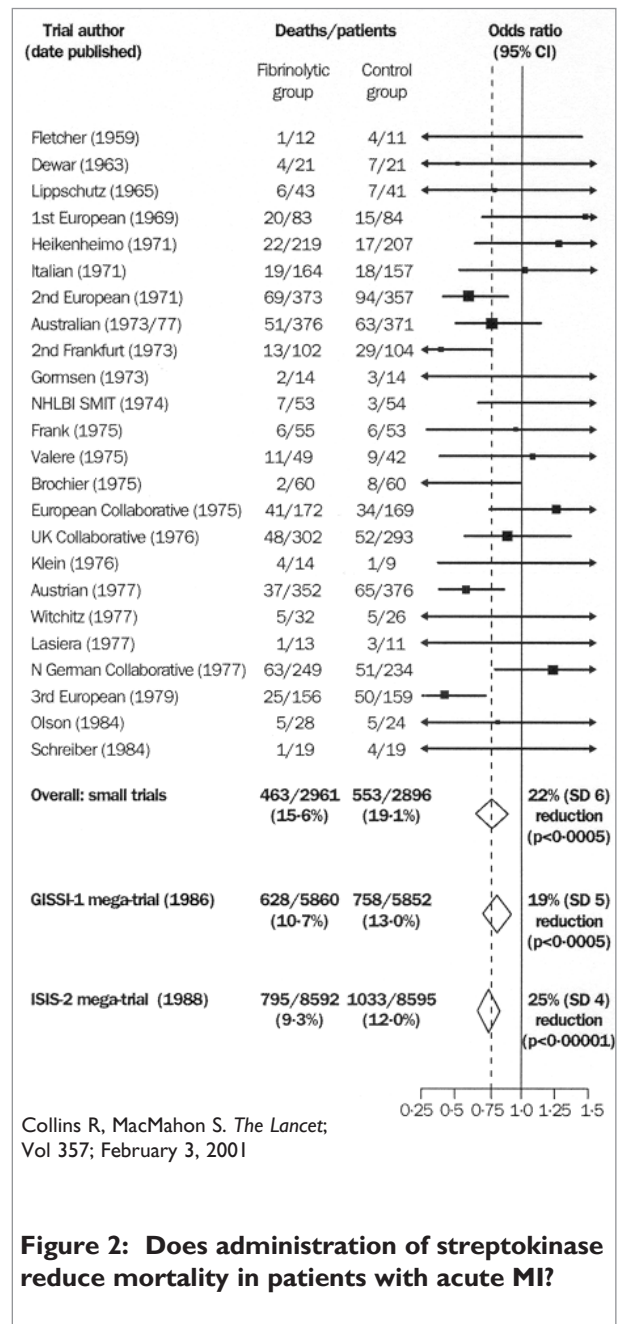
Those evaluating the effects of treatment found that, in addition to various forms of bias, a further source of error could bedevil the interpretation or research viz. random variation (a.k.a. the “play of chance”). It was discovered that, even if the same study was repeated over and over again, the findings ended up being different every time purely due to chance.

Figure 2 illustrates this problem. The figure shows the results of several trials evaluating the effect of streptokinase (a clot-buster drug) on the risk of death in people newly diagnosed with a heart attack (Collins and MacMahon 2001). What you see in the figure is that some studies show that the treatment is beneficial (the little block is the left of the solid vertical line) while other studies find that treatment is harmful (the little block is to the right of the solid vertical line). What is the practitioner to make of these findings? How are they to be interpreted? It turns out that there is no real difference between these results; the variation you see is due to the play of chance which is a common occurrence in small studies.

Over time it was found that the problem of chance variation could be overcome by improving the precision of study findings by either conducting a large enough study or by combining the results of a number of smaller studies if they are sufficiently similar (known as meta-analysis). The diamonds in figure 2 illustrate this. They show that either one of the two large studies or a meta-analysis of all the smaller studies provides unequivocal evidence that streptokinase reduces mortality.

### Using scientific methods to sum up existing evidence

Reliable evidence about whether a treatment works will very rarely come from a single trial. What is usually needed is a synthesis of all studies that have evaluated the effects of the treatment. But this not a trivial exer



**Figure 2: Does administration of streptokinase reduce mortality in patients with acute MI?**

cise for a variety of reasons including the following: 1) access to research is often biased, because studies with positive or dramatic results are more likely to be published than negative studies, bibliographic databases do not provide comprehensive coverage of the literature, and studies may be selectively cited due to personal prejudices or commercial interests; 2) poor-quality studies are published that draw incorrect conclusions; and 3) many studies are too small to be conclusive (Volmink et al. 2004).

In response to these obstacles a scientific methodology for conducting research syntheses, known as the



systematic review, has evolved. A systematic review can be defined as

*“a review in which bias has been reduced by the systematic identification, appraisal, synthesis, and, if relevant, statistical aggregation of all relevant studies on a specific topic according to a predetermined and explicit method”*

(Moher et al. 1999).

The Cochrane Collaboration, a non-profit organisation, provides a global infrastructure for preparing and maintaining systematic reviews of the effects of health care interventions (<http://www.cochrane.org/>). Formed at a meeting of 90 people in Oxford in 1993, the Cochrane Collaboration now involves about 15 000 people from almost 100 countries. Cochrane reviews are subjected to a rigorous peer review, following which they are published electronically in The Cochrane Library (<http://www.thecochranelibrary.com/>).

Locally, the Cochrane Collaboration is represented by the South African Cochrane Centre (<http://www.mrc.ac.za/cochrane/cochrane.htm>). This Centre has existed as an intra-mural Unit of the Medical Research Council for nearly 10 years. As the only Cochrane Centre in Africa, it provides training and support to people across sub-Saharan Africa who wish to conduct Cochrane reviews. The Centre's primary focus is to facilitate the preparation of up-to-date systematic reviews of interventions for high priority health conditions in Africa, especially HIV and tuberculosis.

## CONFRONTING THE ILLUSION OF CERTAINTY – EXPERIENCES IN THE SCIENTIFIC ERA

In the final part of this talk I would like to show, using a few recent examples, how the methods I have briefly described earlier can be used to confront certainty regarding the effects of health care interventions. As I do so, it would be worth keeping in mind the four “idols” of Sir Francis Bacon, which you may recall are:

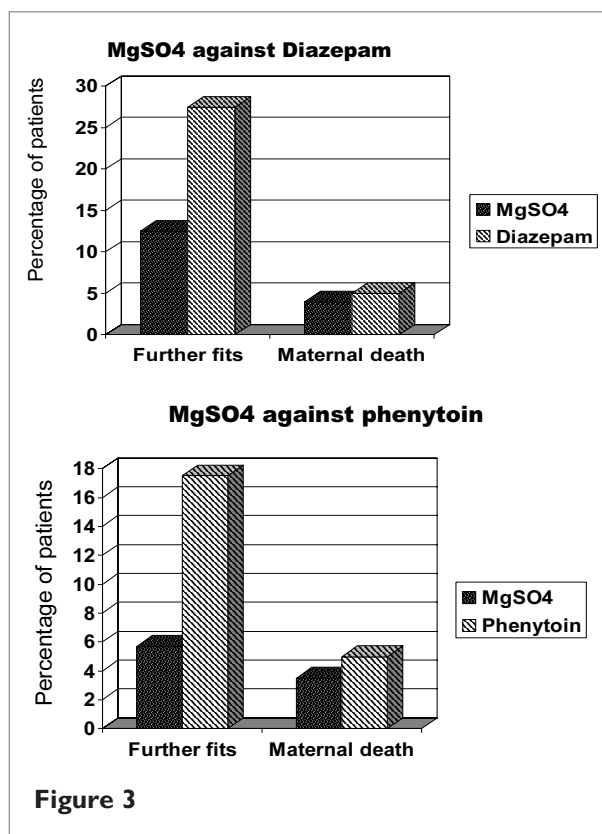
- 1) “Idols of the Tribe” (common to the human race);
- 2) “Idols of the Den” (peculiar to the individual);
- 3) “Idols of the Marketplace” (misuse of language or ‘spin’); and
- 4) “Idols of the Theatre” (abuse of authority).

*Example 1: Certainty based on clinical experience*  
Eclampsia is a serious complication in pregnant women characterised by high blood pressure, proteinuria and

convulsions. Every year an estimated 50 000 women die of eclampsia and 99% of these deaths occur in developing countries (Duley 1992). The pathophysiology of the condition is not yet well understood.

In 1925 an American obstetrician reported a case series of 20 women with eclampsia in which magnesium sulphate was considered to have improved clinical outcomes. Magnesium sulphate was subsequently adopted by some obstetricians for treating eclampsia, but over the years (also based on clinical experience) other treatments, such as phenytoin, diazepam or a ‘lytic cocktail’, were introduced. For many years clinicians fiercely contested each other for the therapeutic high ground, with each being more sure than others of the correctness of their approach. A large, multicentre, randomised trial published in the Lancet in 1995 finally settled the debate (Eclampsia Trial Collaborative Group 1995).

As can be seen in figure 3 the Collaborative Eclampsia Trial found that women receiving magnesium sulphate had a 52% lower risk of recurrent convulsions than those receiving diazepam and also had a lower risk of death. Women allocated to treatment with magnesium sulphate also had a 67% lower risk of recurrent convulsions and a lower risk of death than those allocated to phenytoin. Magnesium sulphate can now be recommended as the anticonvulsant of choice for



women with eclampsia – decades too late for thousands of women who had died as a result of receiving inferior treatment.

An editorial reflecting on this experience and drawing attention to the need for reliable evidence summed up the situation rather elegantly:

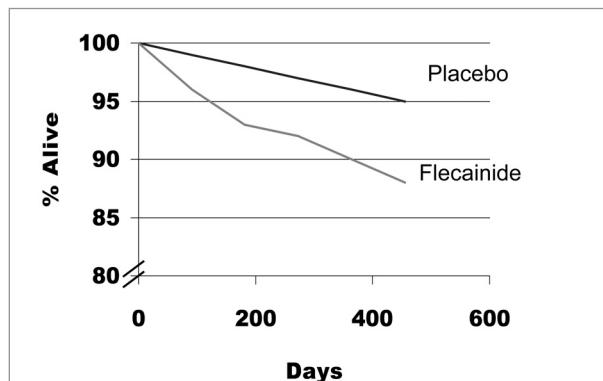
*“For 70 years, the proponents of various drugs and drug cocktails have hurled disdainful abuse at each other from separate mountain tops, secure in the knowledge that no strong evidence existed that could undermine any one of their multitude of conflicting opinions”*

(Chalmers and Grant 1996).

### Example 2: Certainty based on understanding of pathophysiology

It has long been known that there is a high rate of death in the early period following a heart attack, with the highest risk in the first 24 hours. People who develop abnormal heart rhythms during this time are four times more likely to die than those with regular heart beats. Over the years studies in pigs and in humans found that abnormal beats could be suppressed by a group of drugs called class I anti-arrhythmic drugs. Based on these studies the drugs were routinely used as prophylactic treatment in patients with heart attack.

When randomised trials were conducted years later, it was found that patients receiving anti-arrhythmics had



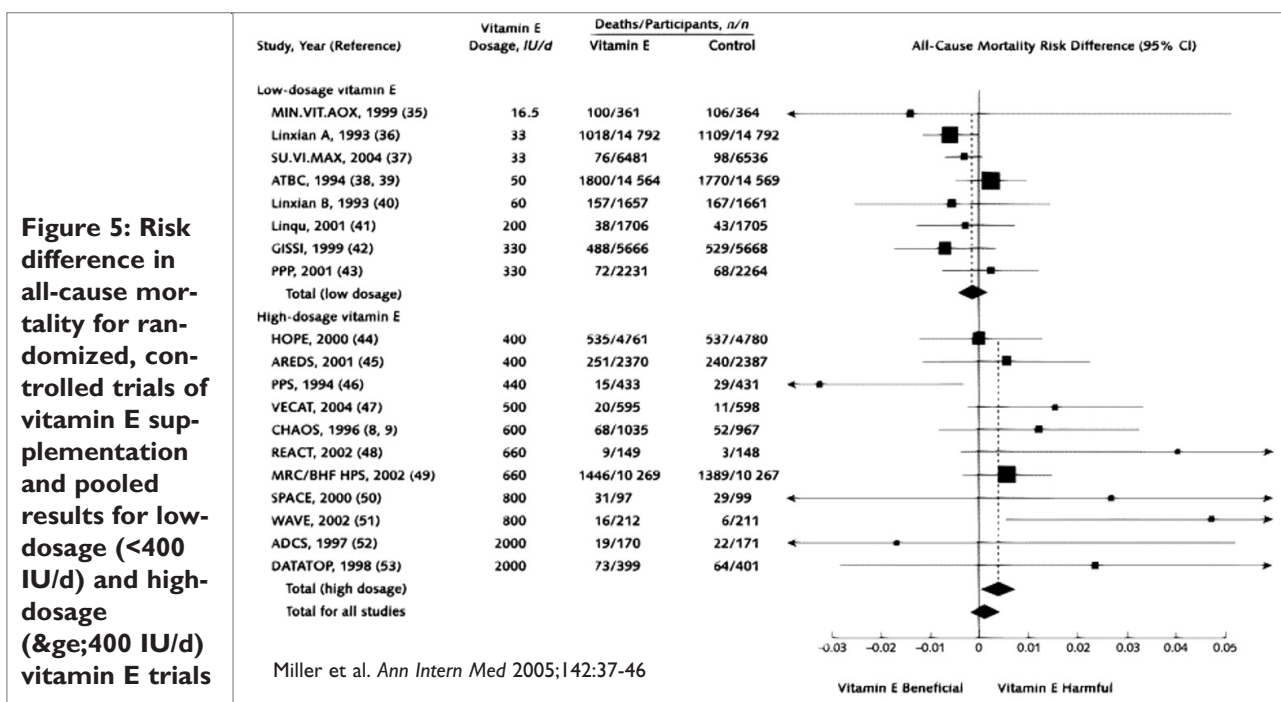
**Figure 4: The Cardiac Arrhythmia Suppression Trial**

beautifully controlled heart rhythms, but to everyone's surprise they were, nonetheless, more likely to die than those patients receiving a placebo (Figure 4) (Echt et al. 1991). It has been estimated that between 20 000 and 70 000 people died as a result of this treatment during the late 1980s in the United States alone (Moore 1995).

This example is a powerful reminder that theoretical knowledge alone does not automatically lead to the introduction of effective treatment.

### Example 3: Certainty based on the influence of the health care industry

It is widely believed that nutritional supplements are good for one's health. These supplements are widely promoted as “antioxidants” to be taken as a preventive

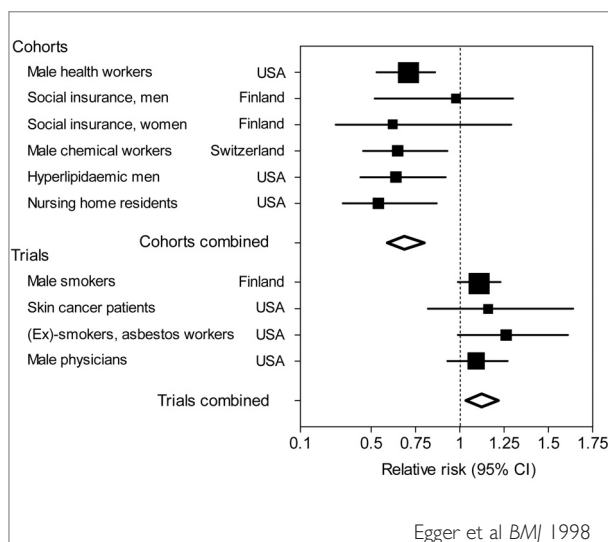


treatment for a variety of chronic diseases. Dietary supplements usually consist of a mixture of vitamins C and E, carotenoids and other phytochemicals and plant extracts. Currently a third of all adults in the US and half of all those older than 55 years report taking supplements daily. In 2003 sales of supplements accounted for \$19 billion in the US alone. Not everyone is convinced of the value of dietary supplements, but even some sceptics hold the view that supplements “won’t hurt and they might help”. But what does the evidence show?

Vitamin E is the most widely used product and is taken daily as specific supplementation (usually as 400 IU of alpha-tocopherol). A recently published systematic review evaluated the evidence from 19 RCTs with almost 136 000 patients – 9 tested vitamin E alone and 10 tested vitamin E combined with other vitamins and minerals (Figure 5). The review found that overall vitamin E had no effect on survival. Of considerable concern was the finding that death rates were higher in people taking high-dose vitamins (>400 IU/d) compared with controls (Miller et al. 2005).

The review is not unusual in finding that nutritional supplements may be harmful. In a similar synthesis of randomised trials it was found that the risk of cardiovascular death was higher in those taking supplements containing beta-carotene compared with those receiving a placebo (Egger et al. 1998) (Figure 6).

Should healthy people taking nutritional supplements be worried about this evidence? Probably. Will people stop taking them? Not if big business can help it!



**Figure 6: Beta-carotene supplementation and cardiovascular mortality**

*Example 4: Certainty based on the opinion of health authorities (policy makers)*

My final example relates to DOTS (directly observed therapy, short-course), a 5-point strategy for delivering care to patients with tuberculosis. The DOTS strategy is, in the main, a sensible approach for improving the management of people with tuberculosis. However, it includes one intervention (directly observed therapy or DOT) that can be considered controversial. The WHO believes that a policy in which patients are watched while they swallow their tablets (directly observed therapy) is crucial for ensuring successful treatment. In 1997 the World Health Organisation released a press statement in which the Director-General claimed that the policy of DOTS was the “biggest health care breakthrough of this decade” and the “single most important development in the fight [against tuberculosis]....since Robert Koch discovered the TB bacillus in 1882”. He further claimed that there was “overwhelming evidence” from studies in USA, Asia and Africa that directly observed therapy was effective. I decided to respond to these assertions by conducting a Cochrane review with Paul Garner in Liverpool as a research partner. We found that, although some adherence promoting interventions had been evaluated in randomised trials, no trials of directly observed therapy (DOT) had ever been done (Volmink and Garner 1997).

We also carried out a systematic review of observational studies reporting on TB programmes in which DOT had been used (Volmink et al. 2000). This review included 32 studies and found that in all settings DOT was only one of several interventions introduced so that any potential effect of DOT will have been highly confounded by the effects of co-interventions.

Since the publication of our first Cochrane review on this topic nine years ago, I am pleased to say 10 RCTs of DOT including nearly 4 000 participants have been conducted. What can we conclude from this evidence? I quote from our most recent Cochrane review update:

*“The results of randomised controlled trials conducted in low-, middle- and high-income countries provide no assurance that directly observed therapy compared with self-administered treatment has any quantitatively important effect on cure or treatment completion in people receiving treatment for tuberculosis”*

(Volmink and Garner 2006).

Is it reasonable to expect that the WHO will respond favourably to this evidence? Of course not!

## CONCLUSION

I would like to conclude with a few caveats. In this talk I have drawn attention to the dangers of basing judgments about the effects of treatment on personal testimony. This is not to denigrate personal experience as a form of knowledge; it can be extremely valuable. There are even situations in which personal experience or clinical observation can provide conclusive evidence, e.g. where the effect of a treatment is dramatic. However, for practical purposes it is safer to regard personal testimony as an entry point for deciding on whether treatment might be beneficial rather than viewing it as proof.

Furthermore, in my comments regarding theory-based knowledge I did not mean to be dismissive of theory. Many of the advances in modern medicine have resulted from improved understanding of pathophysiology. However, it is important to keep in mind that theoretical knowledge can be wrong and must not be expected to lead automatically to effective treatment.

I hope I've succeeded in making the case for the value of empirical methods in the evaluation of health care interventions. However, if I've created the impression that clinical epidemiology is the "Holy Grail" that will open the door to certain knowledge then I apologise. I must remind you once again of Franklin's Law which says: "In this world nothing is certain except death and taxes." Certainty belongs to the realm of dogma only; in making decisions of any kind there is only probability. The most clinical epidemiology can hope to achieve is helping us estimate the likelihood of an intervention being effective. It is only one form of knowledge, albeit a more objective form than personal experience or theory. But it is certainly fallible, which is why I end by pointing to the need for humility, borrowing the words of Xenophanes, Greek Philosopher of the 6th century BC:

*"Through seeking we may learn and know things better. But as for certain truth no man hath known it, for all is but a woven web of guesses."*

## REFERENCES

- Bernard CL, cited in Wulff HR and Gotzsche PC 2000. *Rational Diagnosis and Treatment. Evidence-based Clinical Decision-making*. Blackwell Science Ltd. Oxford
- Chalmers I, Grant A 1996. Salutory lessons from the Collaborative Eclampsia Trial. *Evidence-based Medicine* January/February 1996
- Collins R, MacMahon S 2001. Reliable assessment of the effects of treatment on mortality and major morbidity, I clinical trials. *Lancet* Feb 3; 357(9253): 373-80
- Duley L 1992. Maternal mortality associated with hypertensive disorders of pregnancy in Africa, Asia, Latin America and the Caribbean. *Br J Obstet Gynaecol*; 99: 547-43
- Echt DS, Liebson PR, Mitchell LB 1991. Mortality and morbidity in patients receiving encainide, flecainide or placebo: The Cardiac Arrhythmia Suppression Trial. *NEJM* 324:781-8
- Eclampsia Trial Collaborative Group 1995. Which anti-convulsant for women with eclampsia? Evidence from the Collaborative Eclampsia Trial. *Lancet* 345; 1455-63
- Editorial. The Clyster Craze, *Time Magazine*, 1 Jul 1946
- Egger M, Schneider M, Davey Smith G 1998. Spurious precision? Meta-analysis of observational studies. *BMJ* 316: 140-144
- Ernst E 1997. Colonic irrigation and the theory of autointoxication: A triumph of ignorance over science. *J Clin Gastroenterol* 24:196-198
- Gigerenzer G 2002. *Reckoning with risk. Learning to live with uncertainty*. Penguin Books, London
- Miller ER III, Pastor-Barriuso R, Dalal D, Riemersma RA, Appel LJ, Guallar E 2005. Meta-Analysis: High-Dosage Vitamin E Supplementation May Increase All-Cause Mortality. *Ann Intern Med* 142: 37-46
- Milne I, Chalmers I 2004. Documenting the evidence: the case of scurvy. *Bulletin World Health Organisation* 82: 791-6
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF 1999. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 354: 1896-900
- Moore TJ 1995. *Deadly Medicine*. Simon and Schuster, New York
- Trohler U 2000. *To improve the evidence of medicine. The 18th century British origins of a critical approach*. Royal Physicians of Edinburgh
- Vickers B (Ed) 2002. *Francis Bacon: The Major Works*. Oxford University Press. Oxford 2002
- Volmink J, Garner P 1997. Systematic review of randomised controlled trials of strategies to promote adherence to tuberculosis treatment. *BMJ* Nov 29; 315(7120): 1403-6
- Volmink J, Matchaba P, Garner P 2000. Directly observed therapy and treatment adherence. *Lancet* Apr 15; 355(9212): 1345-50
- Volmink J, Garner P 2006. Directly observed therapy for treating tuberculosis. *Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD003343. DOI: 10.1002/14651858.CD003343.pub2
- Volmink J, Siegfried N, Robertson K, Gulmezoglu AM 2004. Research synthesis and dissemination as a bridge to knowledge management: the Cochrane Collaboration. *Bulletin of the World Health Organisation* 82: 778-783