

Any 'research' is as good as researchers conducting it

Said Shahtahmasebi, PhD

The Good Life Research Centre Trust, Christchurch, New Zealand.

Keywords: human behaviour, dynamics, social processes

Received: 18/2/2014; Revised: 10/3/2014; Accepted: 25/3/2014

INTRODUCTION

Over recent decades there has been much emphasis on evidence-based practice to justify a new policy or changes in existing policy. However, a question arises: what constitutes evidence? In terms of human behaviour, evidence is the means by which we choose to justify an action or a decision. Frequently, an elective model for decision making is employed where decisions are made and supportive evidence is sought selectively subsequent to policy implementation (e.g. see (Short, 1997)). Certainly, under alternative models, hypothesis testing and randomized clinical trials (RCTs) do not constitute evidence on their own to justify a change to an existing policy or a new policy. Although RCTs play a role in the pharmaceutical industry, they have very limited use when investigating issues related to human behaviour. Whether an elective model or other types of model are applied, it is our uncritical use of the 'evidence' that is the weakness and ultimately the cause of an unsound decision.

Human behaviour, like deterministic physical models, will respond to an input, but human reaction unlike the output of deterministic models is not always measurable or quantifiable exactly. For example, measuring health status through a survey questionnaire may lead to the informant's health being described as one of 'excellent', 'good', 'allright', 'not good', 'poor'. Such measurements or description of a variable will be influenced by the informant's own general wellbeing at the time of the interview, as well as his/her personal and social traits.

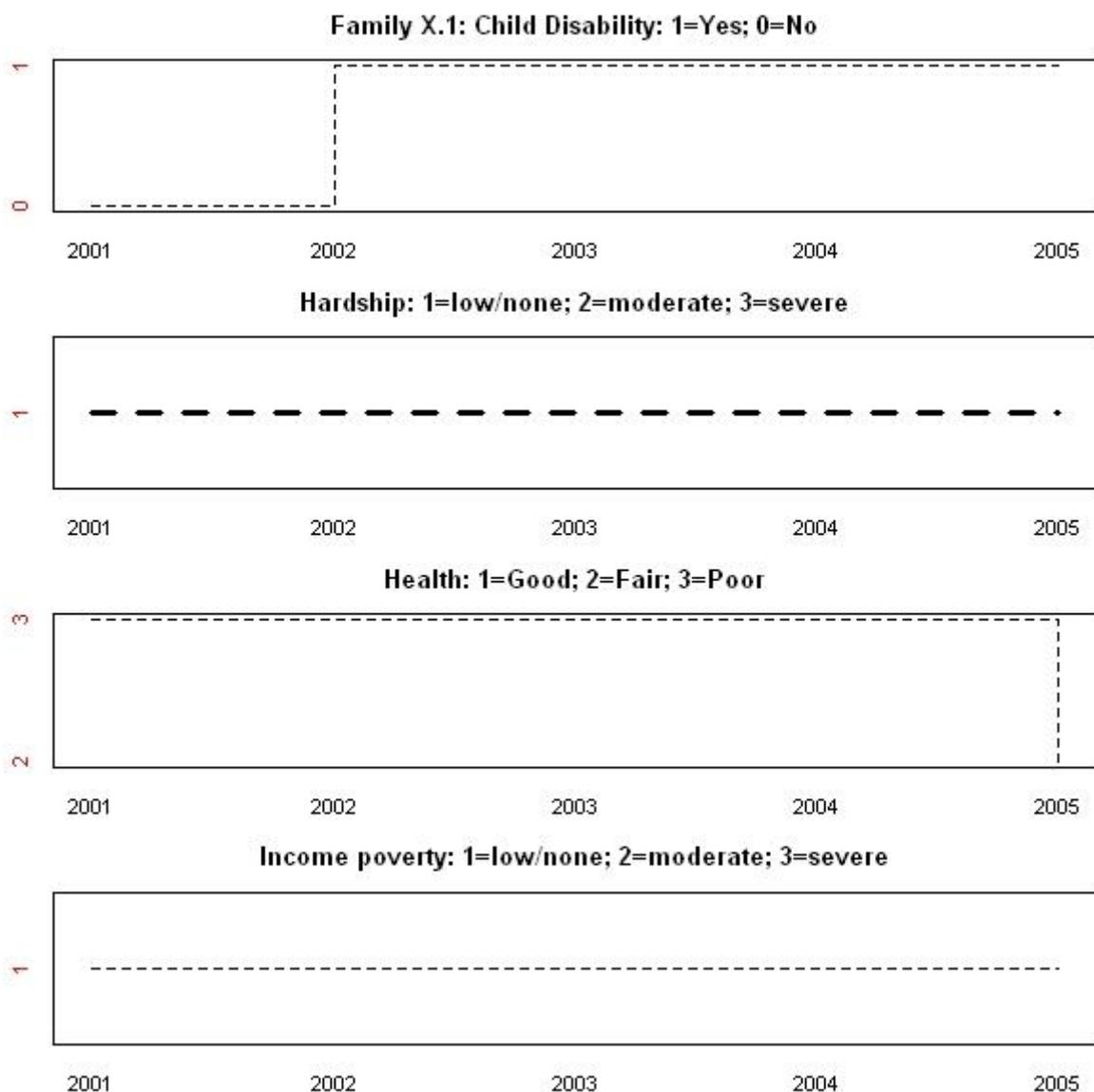
It is plausible that informants may assess their health status by giving consideration to other factors e.g. personal expectation and the feel good factor. For example, elderly people may describe their poor health as allright (for age) (Wenger, 1984). The complexities arise when we attempt to infer cause and effect from the health status variable to another variable such as morale, employment status, smoking, longevity or loneliness. Cross-sectionally, the health effect may appear statistically significant, however, longitudinally, changes in health status may not lead to commensurate changes in morale i.e. over time an individuals' state of health changes but levels morale remain unchanged. Therefore, it is important to control for temporal dependencies.

The following diagrams (figures 1 and 2), from a study relating child disability in the family to poverty, state of health, hardship, and income (Shahtahmasebi, Emerson, Berridge, & Lancaster, 2010, 2011), will illustrate the point more clearly. For instance, figure 1 shows that there was no child disability in the family in 2001, this changed in 2002 and remained so for the rest of the project window. However, this change does not appear to coincide with a change in family hardship, income poverty or health. The only change appears to be an improvement in health from poor to fair three years after the incidence of child disability.

On the other hand the family's hardship shown in figure 2 appears to increase to moderate hardship with the occurrence of disability in the family. A change in health status is reported

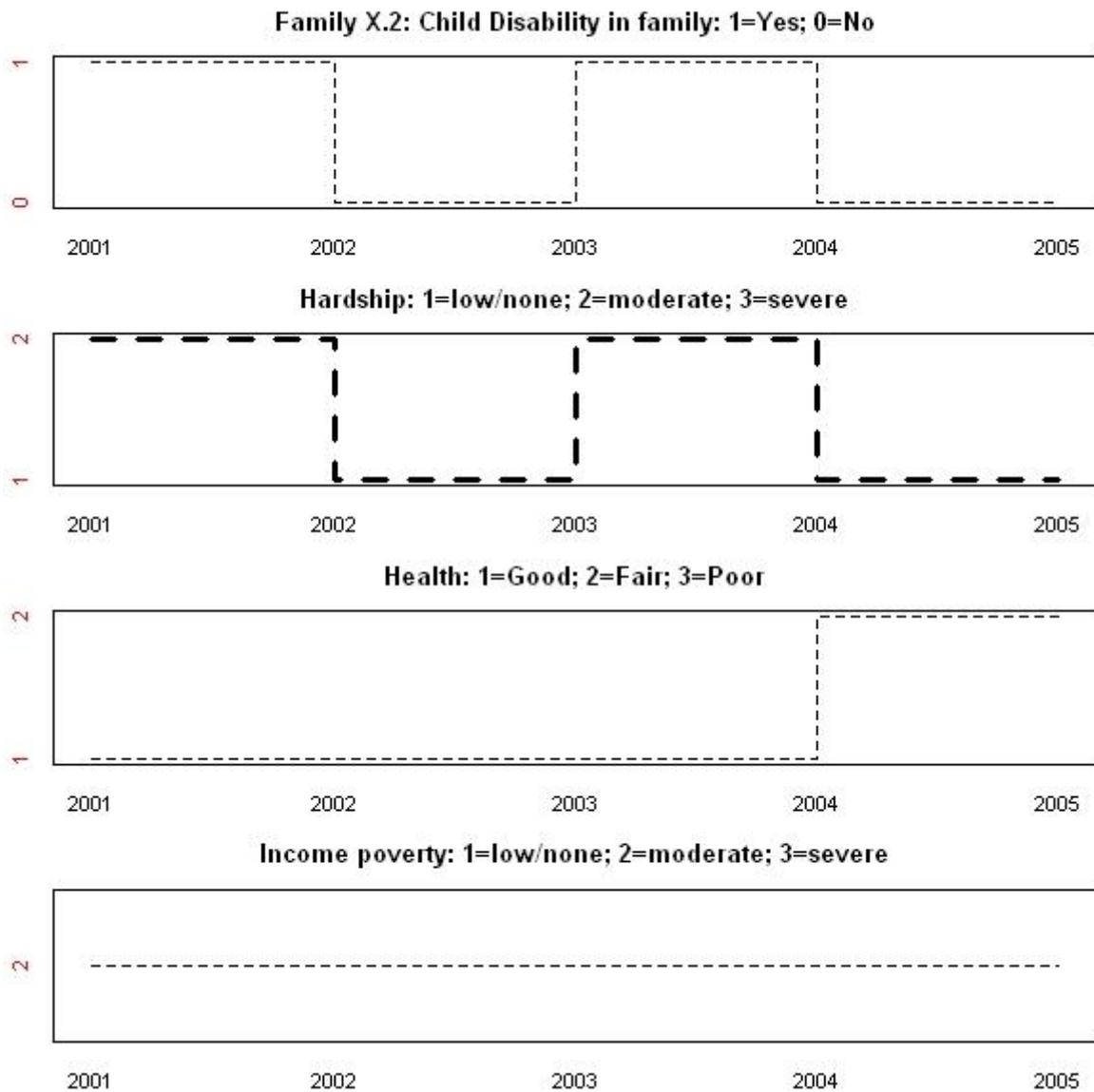
in 2004 from ‘good’ to ‘fair’ and no reported change in income hardship. The fact that prior to change the family were already in severe hardship, poor health and moderate ‘income poverty’ raises several issues: (a) it is unsafe to make statements about a causal relationship if the family’s past history is unknown; (b) there may be families who persistently remain in one state e.g. severe hardship or no hardship (known as stayers); (c) and those families who move between states e.g. good health to poor health and back to good health (known as movers); (d) some families may take a long time to react to a change in circumstance, known as duration dependent e.g. the notion of cumulative inertia or cognitive dissonance; (e) the period of adjustment will vary between families and by type of disability.

Figure 1. Step diagram of disability outcome with other outcomes.



Source: Families and Children Study (FACS), www.data-archive.ac.uk.

Figure 2. Step diagram of disability outcome with other outcomes.



Source: Families and Children Study (FACS), www.data-archive.ac.uk.

Even when the analytical methodology has been appropriately chosen our interpretation of results may be erroneous and misleading (Shahtahmasebi, 2003b, 2004). In other words, there are other variables that are part of human traits such as frailty, personality, the good feel factor that cannot be measured and thus are omitted from study and analysis. We must always bear in mind that we are dealing with human behaviour which implies complex temporal dependencies, feedback effect, and omitted heterogeneity. Ignoring these complexities will adversely impact design, methodology, results, and conclusions.

Some examples from the media

The consequences of an uncritical approach to decision making is all around us:-

In 2013 the Christchurch Press published an article on how a group of researchers and media in Vienna got together and ceased the media reporting of suicides involving the

subway and directly linked a drop in the suicide rate with the method of reporting. The article then claimed this as the most conclusive evidence to date for not reporting suicide in the media. However, an examination of the Viennese suicide trend suggest otherwise (see <http://www.internetandpsychiatry.com/joomla/home-page/editorials-and-commentaries/921-suicide-prevention.html>). In other words, the suicide trend suggests that the downturn began two years before the joining together of researchers and media, and that the reported massive 75% drop in suicide was in fact a drop in the use of subway as a preferred method of suicide.

In the UK the BBC report on the MMR controversy (<http://news.bbc.co.uk/1/hi/health/1808826.stm>) stated that some studies linked the MMR vaccine to autism whilst others offered evidence to the contrary. The BBC also reported that on the one hand ‘Alcohol makes your brain grow: Drinking alcohol boosts the growth of new nerve cells in the brain, research suggests.’ (<http://news.bbc.co.uk/go/em/-/1/hi/health/4496727.stm>), and on the other hand ‘Alcohol link to bowel cancer risk: A daily pint of beer or a large glass of wine raises the risk of bowel cancer by about 10%, research suggests.’ (<http://news.bbc.co.uk/go/em/-/1/hi/health/6921998.stm>).

Whilst, in New Zealand the media took delight in reporting and quoting research from the US, on how alcohol can improve memory (e.g. see <http://www.stuff.co.nz/dominion-post/archive/national-news/44709/Drinking-to-forget-may-backfire>)! Given the effects that alcohol has on living cells and on human health such reported outcomes are counter intuitive at best. Yet, the emphasis is often on the word ‘research’ thus resulting in an uncritical and unquestioning acceptance of research outcomes. Indeed, for every claim from one group of scientists there would have been a counter claim. In suicide research Beautrais (Beautrais, 1996, 2001) claimed that depression and mental illness were the cause of suicide, Khan et al (Khan, Warner, & Brown, 2000) claimed that antidepressants did not reduce suicide but could increase the risk of suicide while Hall et al (Hall et al., 2003) claimed that antidepressants reduced suicides. These studies have failed to address methodological issues relating to design, data collection and analysis thus resulting in misleading conclusions (Shahtahmasebi, 2003a, 2005). We believe that such research and reporting has considerably attenuated the credence and prestige that research once had in the mind of the public.

Still in New Zealand, the Medical Journal of New Zealand claimed depression is a common, serious and significant illness and linked it to suicide and recommends medication [<http://www.nzma.org.nz/journal/117-1206/1200/>]. It is not surprising to hear that young people have been prescribed antidepressants including preschool children. But more alarming is the prescribing of antidepressants to some children under a year old [http://www.nzherald.co.nz/section/1/story.cfm?c_id=1&objectid=10462684].

The results from the above mentioned studies cannot be relied on due to flaws in methodology leading to misconclusions. For example, Beautrais (Beautrais, Joyce, & Mulder, 1994) introduced major bias in her study in several ways. First, major bias was introduced through collecting data about the mental wellbeing of suicide cases from relatives and friends after the event of suicide. Second, additional bias was introduced through a measurement tool. Third, the analytical methodology was inappropriate and failed to account and control for these bias.

Critical thinking

Publication does not equate accuracy. Publication bias is quite real, with an increasing number of authors reporting poor and scientifically compromised publications in reputable journals e.g. see (Altman, 1994; Fang, Steen, & Casadevall, 2012; Leszczynski, 2013; Smith, 2014; Wise, 2009) ranging from poor study design, inappropriate methodology or statistical methods used, invalid interpretation, plagiarism. According to one report the rate of publication retraction has increased ten-fold since 1975 (Fang et al., 2012).

Furthermore, consider publication bias in conjunction with the current practice by established journals of rejecting about 90% of submitted papers without review. Editors would rather prioritise RCTs, longitudinal studies, or meta-analysis. But the inappropriateness of methodology, e.g. using cross-sectional or conventional methods to analyse longitudinal data is often missed out.

The current debate about flawed research and publications appears to blame the peer reviewing system of expert journals. In this context and with the recent advent of “open access” journals – where authors rather than reader pay for publication - it would be even easier to shift the blame completely onto such journals for publishing flawed research (Hawkes, 2013).

Journal editors exacerbate publication bias through their selection bias of not reviewing 90% of submitted papers especially methodology papers, therefore contributing to the complexities of an already complex problem. The peer review system has a major role to play but until it has been depoliticized and become more expert and fair, it would be more beneficial to adopt a self-critical approach in our research and utilizing literature. In particular where seeking evidence to influence the process of decision making. One of the more recurrent failures in research is the appropriateness of methodology for analysis.

Statistical methodology

The emphasis in data analysis must be on distinguishing systematic effects from random variation which tend to obscure any pattern.

Many conventional statistical techniques tend to use hypothesis testing. Within the ‘statistical modelling’ framework, hypothesis testing has a role to play in selecting the most parsimonious model and, in this context, hypothesis testing is a logical part of the comprehensive analysis. This is in contrast to conventional statistical analysis in which hypothesis testing tends to be seen as an independent inferential statement often of meagre substantive value.

Statistical modelling is a comprehensive structured framework for making inference from data. The statistical modelling approach may be summarised as follows (Shahtahmasebi & Berridge, 2009):

- a Model formulation - this step involves the consideration of a well thought out sampling scheme and the type of data in hand. It is guided by substantive theory.
- b Model fitting
- c Model criticism - having ‘best’ estimates for the parameters, the current model can then be tested to see how well it explains the data.

If, as a consequence of (a) and (b), the current model proves satisfactory then proceed with the interpretation of the model using substantive theory else repeat steps (a)-(c).

For the *non-statistician*, grappling with *statistical modelling* is more of a challenge because of its explicit emphasis on *probability modelling*. Always consult a statistician at the planning stage rather than after data collection.

REFERENCES

- Altman, D. G. (1994). The scandal of poor medical research. *BMJ* 1994;308:283.
- Beautrais, A. L. (1996). *Serious suicide attempts in young people: a case control study*. PhD, Christchurch School of Medicine, Christchurch.
- Beautrais, A. L. (2001). Suicides and serious suicide attempts: two populations or one? *Psychol Med*, 31, 837-845.
- Beautrais, A. L., Joyce, P. R., & Mulder, R. T. (1994). The Canterbury suicide project: Aims, overview and progress. *Community Mental Health in New Zealand*, 8(2), 32-39.
- Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences of the United States of America*, doi: 10.1073/pnas.1212247109(<http://www.pnas.org/content/early/2012/09/27/1212247109>).
- Hall, W. D., Mant, A., Mitchell, P. B., Rendle, V. A., Hickie, I. B., & McManus, P. (2003). Association between antidepressant prescribing and suicide in Australia, 1991-2000: trend analysis. *BMJ*, 326.
- Hawkes, N. (2013). Spoof research paper is accepted by 157 journals. *BMJ*, <http://dx.doi.org/10.1136/bmj.f5975>.
- Khan, A., Warner, H. A., & Brown, W. A. (2000). Symptom Reduction and Suicide Risk in Patients Treated With Placebo in Antidepressant Clinical Trials: An Analysis of the Food and Drug Administration Database. *Arch Gen Psychiatry*, 57, 311-317.
- Leszczynski, D. (2013). Opinion: Scientific Peer Review in Crisis: The case of the Danish Cohort. *The Scientist*, <http://www.the-scientist.com/?articles.view/articleNo/34518/title/Opinion--Scientific-Peer-Review-in-Crisis/>.
- Shahtahmasebi, S. (2003a). Suicides by Mentally Ill People. *TheScientificWorldJOURNAL*, 3, 684-693.
- Shahtahmasebi, S. (2003b). Teenage Smoking: some problems in interpreting the evidence. *Int J Adolesc Med Health*, 15(4), 307-320.
- Shahtahmasebi, S. (2004). Longitudinal analysis of teenage smoking: Sabre vs GenStat. *IBC/ASC 2004 International Conference 11-16 July. Cairns, Australia*.
- Shahtahmasebi, S. (2005). Suicides in New Zealand. *TheScientificWorldJOURNAL*, 5, 527-534.
- Shahtahmasebi, S., & Berridge, D. (2009). *Conceptualising behaviour in health and social research: a practical guide to data analysis*. New York: Nova Sci.
- Shahtahmasebi, S., Emerson, E., Berridge, D., & Lancaster, G. (2010). A longitudinal analysis of poverty among families supporting a child with a disability. *Int J Disabil Hum Dev*, 9(1), 65-75.
- Shahtahmasebi, S., Emerson, E., Berridge, D., & Lancaster, G. (2011). Child Disability and the Dynamics of Family Poverty, Hardship and Financial Strain. *Journal of Social Policy*, 40(04), 653 - 675.
- Short, S. (1997). Elective Affinities: Research and Health Policy Development. In H. Gardner (Ed.), *Health Policy in Australia*. Melbourne: Oxford University Press.
- Smith, R. (2014). Most medical research is flawed, says leading medical editor. http://www.mercatornet.com/articles/view/most_medical_research_is_flawed_says_leading_medical_editor.
- Wenger, G. C. (1984). *The Supportive Network - coping with old age*. London: George Allen and Unwin.
- Wise, J. (2009). High proportion of trials published in Chinese medical journals are flawed, study shows. *BMJ* 2009;339:b2729.