Reporting bias inflates the reputation of medical treatments: A comparison of outcomes in clinical trials and online product reviews ✩

Mícheál de Barra

Health Psychology Group, Institute of Applied Health Sciences, University of Aberdeen, 2nd Floor Health Sciences Building, Foresterhill Campus, Ashgrove Road, Aberdeen, AB25 2ZD Scotland, United Kingdom

A B S T R A C T

Objectives: People often hold unduly positive expectations about the outcomes of medicines and other healthcare products. Here the following explanation is tested: people who have a positive outcome tend to tell more people about their disease/treatment than people with poor or average outcomes. Akin to the file drawer problem in science, this systematically and positively distorts the information available to others.

Method: If people with good treatment outcomes are more inclined to tell others, then they should also be more inclined to write online medical product reviews. Therefore, average treatment outcomes in these reviews should be more positive than those found in randomised controlled trials (RCTs). Data on duration of treatment and outcome (i.e., weight/cholesterol change) were extracted from user-generated health product reviews on Amazon.com and compared to RCT data for the same treatments using ANOVA. The sample included 1675 reviews of cholesterol reduction (Benecol, CholestOff) and weight loss (Orlistat) treatments and the primary outcome was cholesterol change (Bencol and CholestOff) or weight change (Orlistat).

Results: In three independent tests, average outcomes reported in the reviews were substantially more positive than the outcomes reported in the medical literature ($\eta^2 = 0.01$ to 0.06; $p = 0.04$ to 0.001). For example, average cholesterol change following use of Benecol is -14 mg/dl in RCTs and -45 mg/dl in online reviews.

Conclusions: People with good treatment outcomes are more inclined to share information about their treatment, which distorts the information available to others. People who rely on word of mouth reputation, electronic or real life, are likely to develop unduly positive expectations.

1. Introduction

People often use medical treatments that are unlikely to have a direct therapeutic benefit (Ernst and Singh, 2006; Evans et al., 2010). Within conventional medicine, this is described as medical overuse or over-treatment and it is common in both prescribed medications and procedures and, importantly for the present work, in over-the-counter medicine use (Busfield, 2015). Moreover, medical systems such as herbal, alternative, complementary, Aru-vedic, and Chinese medicine remain popular despite offering few treatments with demonstrable therapeutic benefits (though it is

possible that some users experience benefits besides those typically emphasised in biomedical science). For example, over 100 million Europeans currently use traditional and complimentary medicine treatments (WHO, 2013), few of which are supported by scientific evidence. For the purposes of this paper, the term 'medical overuse' is used to designate the use of a conventional or alternative therapy that would not have been used had the patient had full knowledge of outcome probabilities. The primary focus of this work is the causes of medical overuse in regards to non-prescription medicines; for a review of the economic, health and environmental consequences of medical overuse, see Thomas and Depledge (2015).

The financial interest of pharmaceutical and other healthcare industries is undoubtedly an important driver of medical overuse (Busfield, 2015; Thomas and Depledge, 2015). However, an additional or interacting cause is that patients often hold preferences

for treatments that are likely to have little therapeutic effect. These preferences are particularly important in the context of over-the-counter treatments, but much recent work suggests that patients' expectations are important drivers of professional behaviour (Coenen et al., 2006; Covvey et al., 2014) and policy decisions (Taylor and Bury, 2007). There is strong evidence that patients hold unduly positive expectations about treatment outcomes. A recent systematic review found that estimations of benefit were unduly high in at least 63% of the samples studied (Hoffmann and Del Mar 2015). Benefits were underestimated for just 3% of outcomes. These positive expectations about treatment outcomes are likely to drive medical overuse. Many basic psychological theories like subjective expected utility theory (Edwards, 1954) and social cognitive theory (Bandura, 1986) emphasise that positive expectations about the outcome of a given behaviour make that behaviour more likely. Moreover, there is empirical evidence that people who believe that a medicine will have positive effects are more likely to use that medicine (Horne et al., 2013).

A recent review of medical reasoning by Lilienfeld et al. (2014) catalogues the processes that predispose people to incorrectly attribute positive change to a treatment. For instance, health often improves due to regression-to-the-mean, because the disease is self-limiting, or because of placebo effects, and these positive changes can be incorrectly attributed to the medical treatment. These processes may explain why people come to believe that a treatment works *after one has tried it*, but as Hoffmann and Del Mar (2015) demonstrate, people often hold false beliefs about likely treatment outcomes prior to use.

An alternative explanation of unduly positive outcome expectations was proposed by de Barra et al. (2014). In contrast with the mechanisms listed in Lilienfeld et al. (2014), this theory does not derive its explanatory power from peoples' biased reasoning or faulty logic, but instead from features of the health communication process. Inaccurate health beliefs emerge, they argue, because a non-representative subset of treatment outcomes is communicated from person to person. The remainder of this paper explores the assumptions of, and predictions derived from, this theory.

When people use a medical treatment, there is generally a broad range of outcomes, with some people improving and others deteriorating. A subset of the people who use a given treatment will communicate their experience to other people. For example, they might tell friends and colleagues that they lost weight after using a weight loss drug or that their cholesterol count has unexpectedly increased since they started using statins. Exposure to this kind of health information is likely to influence the recipient's health beliefs and health behaviour (see below). Health beliefs based on the outcome experience of small samples of people are likely to be error prone, with individuals developing overly positive or overly negative treatment expectations. However, what de Barra et al. (2014) additionally suggest is that the subset of individuals who actively communicate information about their treatment/outcome is not representative of the total outcome distribution. Rather, they propose that people who have positive outcomes are more inclined to share information than people with negative or neutral outcomes. If there is a positive correlation between outcome positivity and probability of information sharing, then the information circulating about the treatment will be systematically and positively biased because people with poorer outcomes will appear to be relatively rare. Such a communication pattern could account for the unduly positive treatment outcome expectations discussed above. Note that a similar under-representation of poor to middling reviews has been documented in the marketing literature, where it is termed the *under-reporting bias* (Anderson, 1998; Hu et al., 2006). However, the under-reporting of negative outcomes might equivalently be described as an over-reporting of positive outcomes.

Thus, the term *reporting bias* will be used here, where 'bias' is meant in the statistical sense (i.e., a biased sample) rather than in the psychological sense (i.e., a deviation from some normative standard of reasoning, as in Tversky and Kahneman, 1974).

The notion that exposure to other people's health outcomes might influence beliefs and subsequent health behaviour is consistent with a range of models of health behaviour. Social cognition models, like the health belief model (Rosenstock, 1966), the theory of planned behaviour (Ajzen, 1985), and descendent theories, assume that people choose to adopt a particular behaviour, in part, because of a belief that that behaviour will lead to a preferred outcome with an acceptably high probability. The crucial issue of how these beliefs are formed is not precisely described by these theories, but it seems safe to assume that observation of another person's outcome is an influential event. Imagine, for example, that we encounter a person who speaks highly of a cholesterol reduction drug they have used. This encounter might influence key health behaviour determinants like (a) the subjective probability that using this cholesterol reduction drug will have the desired outcome, (b) how much this outcome — cholesterol reduction — would positively improve our well-being, (c) the degree to which we see the health behaviour, taking a pill twice daily in perpetuity, as achievable for us (our perceived behavioural control), or (d) it might influence our perception of the social norm, that is the degree to which we see ignoring high cholesterol as socially acceptable. That we are tuned to learn from other's outcomes is also consistent with observational learning theory (Fryling et al., 2011) and several empirical studies (Betsch et al., 2011; Gregory et al., 2011; Winterbottom et al., 2008) as well as the sampling framework within cognitive psychology (Fiedler and Juslin, 2006).

A central prediction from de Barra et al.'s *reporting bias* theory is that the reputation of the treatment (i.e., the average outcome among people who choose to describe their treatment to others) should be substantially more positive than the average outcome as measured in a clinical trial of representative patients. Online medical product reviews provide a good arena in which to test this prediction because this form of information sharing leaves a lasting digital trace that can be quantified and analysed. Moreover, the psychological and contextual factors that lead people to share information about medical products appears to be similar in offline and online contexts (King et al., 2014), suggesting that findings from the online domain may generalise to the offline domain.

1.1. Research questions

de Barra et al. (2014) demonstrated that the reputation of several alternative medicines, measured by averaging the outcomes in multiple Amazon reviews, is more positive than one might expect based on clinical trials. The present work replicates and extends this finding in several ways. First, the previous work focused on alternative/unorthodox treatments (the Atkins diet, herbal fertility treatments). One possible explanation for this finding is that people who have average/negative outcomes after using an unorthodox treatment may be unlikely to share their experiences because they are ashamed to have made a poor medical decision or because their experiences provide little new information to a broadly sceptical audience. If so, then the explanatory scope of the theory presented here is quite narrow and it cannot explain why outcome expectancies for conventional, commonly used medicines are unduly positive. The first and main objective of this study, research question one (RQ1), is to test the generality of the reporting bias by examining if the reputation of conventional treatments (i.e., treatments widely supported by national health agencies and commonly prescribed by medical doctors) is also