Analysing indicators of performance, satisfaction, or safety using empirical logit transformation

Sarah Stevens,[1,2] Jose M Valderas,[3] Tim Doran,[4] Rafael Perera,[1,2] Evangelos Kontopantelis[2,5]

[1]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
[2]National Institute for Health Research School for Primary Care Research, Oxford, UK
[3]APEx Collaboration for Academic Primary Care, Institute for Health Services Research, University of Exeter Medical School, University of Exeter, Exeter, UK
[4]Department of Health Sciences, University of York, UK
[5]Centre for Health Informatics, Institute of Population Health, University of Manchester, Manchester M13 9GB, UK

Correspondence to: E Kontopantelis e.kontopantelis@manchester.ac.uk

Cite this as: *BMJ* 2016;352:i1114
http://dx.doi.org/10.1136/bmj.i1114

Accepted: 20 January 2016

Performance, satisfaction, and safety indicators are commonly measured on a percentage scale. Such indicators are often subject to ceiling or floor effects and performance may be inherently non-linear. For example, improving from 85% to 95% might be more difficult and need more effort than improving from 55% to 65%. As such, analysis of these indicators is not always straightforward and standard linear analysis could be problematic. We present the most common approach to dealing with this problem: a logit transformation of the score, following which standard linear analysis can be conducted on the transformed score. We also demonstrate how estimates can be back-transformed to percentages for easier communication of findings. In this paper, we discuss the benefits of this method, use algebra to describe the relevant steps in the transformation process, provide guidance on interpretation, and provide a tool for analysis.

SUMMARY BOX

- Performance, satisfaction, or safety indicators in healthcare are commonly measured on a percentage scale
- Standard linear analysis could be problematic owing to ceiling or floor effects or non-linearity
- A logit transformation of the score is the most common solution
- Estimates can be back-transformed to percentages for a more intuitive interpretation

In recent years, efforts to improve the quality and safety of healthcare have resulted in the introduction of systems for monitoring the performance of healthcare providers and the satisfaction and safety of patients. New quality and performance indicators have been created, to which financial and reputational rewards for providers are often attached. Although performance indicators are measured for each patient, they are often only reported in aggregate form (eg, at the practice or hospital level). Therefore, an indicator that begins as a binary outcome (that is, the target is either met or not met for each patient),[1] becomes a proportion (that is, the percentage of patients for whom the quality target is met). Such summary indicators are usually analysed by linear models. This is appropriate in many scenarios where the scores retain linear properties, for example, in the analysis of referral rates and their predictors.[2]

However, for aggregate analyses of performance indicators, two particular problems can emerge. Firstly, it is common for individual indicators within a set to vary in intrinsic difficulty (eg, recording blood pressure is easier than controlling blood pressure) or vary in the size of associated incentives. This frequently results in healthcare providers achieving targets for 100% of patients for easier indicators[3] and, less often, for 0% of patients for more difficult indicators, or for indicators with smaller incentives. Maximum (100%) and minimum (0%) scores are more common when patient groups are small (the problem of small denominators). These "ceiling" and "floor" effects can cause problems in analyses of data at the patient level, but also make the use of aggregate performance scores in linear models problematic. This is a particular problem for prediction modelling (eg, in interrupted time series designs)[4] where predictions might fall outside the 0-100% range.

Secondly, there is inherent non-linearity in performance indicators, because the effort required by a health worker is not uniform across patients. For example, some patients might attend clinic appointments infrequently while others might be persistent in refusing a measurement or treatment. Similarly, satisfaction is subjective and different levels of effort are needed to satisfy different patients, whereas in terms of safety, risk management is inexact and some patients might be more difficult to manage clinically. Therefore, it is generally more difficult to achieve an improvement from 85% to 95%, than from 55% to 65%. Analogously, an improvement from 0% to 10% should pose very little difficulty. Box 1 presents some examples of performance, satisfaction, and safety[5] indicators.

One potential solution to these issues is for researchers to dichotomise the indicator by classifying healthcare providers simply as high or low achievers (in terms of performance, satisfaction, or safety), based on a specified threshold of achievement. For example, assume that a healthcare provider has met a target (=1) if the relevant performance score is over 85%, and not met the target (=0) otherwise. Analyses are then possible by use of logistic models, and odds ratios would be used to quantify effects. However, odds ratios are intuitively difficult to conceptualise and are frequently interpreted as

BMJ: first published as 10.1136/bmj.i1114 on 10 March 2016. Downloaded from https://www.bmj.com/ on 9 January 2025 by guest. Protected by copyright.