Assuming that prevalence ($x$%) is a predictor of achievement, the achievement score ($p$) for a practice on the logit scale can be expressed though a simple regression as:

$\text{Logit}(p) = a + \beta x$, where $a$ is the intercept and $\beta$ is the regression coefficient quantifying the association between $x$ and $p$.

Achievement on the logit scale for the same practice, if prevalence increases by 1%, becomes: $\text{Logit}(p_{new}) = a + \beta(x + 1)$

The difference gives the change in achievement (on the logit scale) for a 1% increase in prevalence (beta values in a linear regression model): $\text{Logit}(p_{new}) - \text{Logit}(p) = \beta$

Achievement on the logit scale is related to a particular value of percentage achievement (as defined by the logit transformation):

$$\text{Logit}(p) = \ln\left(\frac{p}{1-p}\right) \quad (1)$$

Similarly, $$\text{Logit}(p_{new}) = \ln\left(\frac{p_{new}}{1-p_{new}}\right)$$

To work backwards to get a change in percentage achievement, we assume that percentage achievement at prevalence ($x$+1)% is equal to the percentage achievement at prevalence $x$% plus a constant c, the change in percentage achievement per 1% increase in prevalence.

$$\left.\begin{array}{r} p_{new} = p + c \\ \text{Logit}(p_{new}) - \text{Logit}(p) = \beta \end{array}\right\}^{(1)} \Rightarrow \ln\left(\frac{p+c}{1-(p+c)}\right) - \ln\left(\frac{p}{1-p}\right) = \beta$$

Solving for c, we obtain:

$$c = \left[exp\,(-\beta)\left(\frac{1-p}{p}\right) + 1\right]^{-1} - p$$

To obtain $c$, we need to assume an anchor value for $p$, and the average achievement score across clinical units or practices is as good an assumption as any.

**Fig 3** | **Back-transformation explained**

at improving performance, satisfaction, or safety in a low achieving setting only.

We would therefore argue that choice of an anchor score is largely at the discretion of researchers, who should use the research aims to inform their choice. However, the mean or median scores should be suitable in most scenarios and a priori justification would be needed for alternative anchor choices. Another approach could be to present back-transformed results obtained using several different anchor scores to stimulate discussion around this issue, although attention needs to be given to the interpretation of the group of results. It should also be noted that transformed scores, like percentage scores, do not account for the difficulty in meeting a specific indicator and that investigators should be careful with comparisons across indicators of varying difficulty levels. In these cases, the anchor score can be chosen to reflect the inherent difficulty for an indicator, although the relation between the anchor score and difficulty is not intuitive.

To aid researchers with use of these methods, we have made available an Excel workbook with the transformation and back-transformation formulas, given different anchor scores (available from the corresponding author on request or from his personal website (www.statanalysis.co.uk/files/logit_transformation.xlsx)).

**Table 1** | **Quantification of back-transformed effects of a 1% increase in diabetes prevalence, at various anchors**

| Increase in prevalence (%) | Effect of predictor (prevalence) on logit scale (95% CI) | Anchor achievement score ($p$) | Back-transformed effect of predictor on achievement score (absolute difference $c$) (95% CI) |
|---|---|---|---|
| 1 | −0.031 (−0.041 to −0.021) | 0.9500 | −0.0015 (−0.0020 to −0.0010) |
| | | 0.9245 (median) | −0.0022 (−0.0029 to −0.0015) |
| | | 0.7500 | −0.0059 (−0.0078 to −0.0040) |
| | | 0.5000 | −0.0077 (−0.0102 to −0.0052) |
| | | 0.2500 | −0.0058 (−0.0076 to −0.0039) |
| | | 0.0500 | −0.0015 (−0.0019 to −0.0010) |

## Discussion

We have demonstrated the use of empirical logit transformation for the analysis of performance, satisfaction, or safety indicators that are subject to ceiling or floor effects. We have argued the benefits of this method, algebraically described the processes, provided guidance on interpretation, and have made available a simple tool to aid researchers in using the method. These methods have broad applicability in health services research, but can also be applied in other settings, for example, citizen satisfaction with urban services[16] or hotel websites.[17]

We thank Jill Stokes, researcher at the University of Manchester, who provided the prescription safety examples; and Ben Amies, a GP registrar, who provided feedback on the paper.

Contributorship: SS wrote the manuscript, with help from EK. JMV, TD, and RP critically edited the manuscript. SS is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. In this article, we present examples on applications of these methods to a range of research questions and studies as published in major clinical journals, including *The BMJ*. SS is an early career statistician who has recently been involved with analysing performance measures in a primary care setting. TD is a clinical researcher with interests in quality of care and experience in applying these methods. JMV is a clinician with expertise on the measurement of quality of care and in psychometric methods. RP is a medical statistician whose research program focuses, not exclusively, on monitoring in primary care. EK is a biostatistician and health services researcher who has used the reported methods to answer research questions pertaining to incentivisation in primary care.

Funding: UK Medical Research Council Health eResearch Centre grant MR/K006665/1 supported the time and facilities of EK.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Kontopantelis E, Reeves D, Valderas JM, Campbell S, Doran T. Recorded quality of primary care for patients with diabetes in England before and after the introduction of a financial incentive scheme: a longitudinal observational study. *BMJ Qual Saf* 2013;22:53-64. doi:10.1136/bmjqs-2012-001033.

2 Hippisley-Cox J, Hardy C, Pringle M, Fielding K, Carlisle R, Chilvers C. The effect of deprivation on variations in general practitioners' referral rates: a cross sectional study of computerised data on new medical and surgical outpatient referrals in Nottinghamshire. *BMJ* 1997;314:1458-61. doi:10.1136/bmj.314.7092.1458.

3 Doran T, Fullwood C, Gravelle H, et al. Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med* 2006;355:375-84. doi:10.1056/NEJMsa055505.

4 Kontopantelis E, Doran T, Springate DA, Buchan I, Reeves D. Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. *BMJ* 2015;350:h2750. doi:10.1136/bmj.h2750.

5 Stocks SJ, Kontopantelis E, Akbarov A, Rodgers S, Avery AJ, Ashcroft DM. Examining variations in prescribing safety in UK general practice: cross sectional study using the Clinical Practice Research Datalink. *BMJ* 2015;351:h5501. doi:10.1136/bmj.h5501.

6 Davies HTO, Crombie IK, Tavakoli M. When can odds ratios mislead?*BMJ* 1998;316:989-91. doi:10.1136/bmj.316.7136.989.

7 Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. Effects of pay for performance on the quality of primary care in England. *N Engl J Med* 2009;361:368-78. doi:10.1056/NEJMsa0807651.