

advocate, so when patients say they do not feel well enough to work, the GP is expected to support them. A fine principle perhaps, but clearly much abused. GPs say that advising on fitness to work is an area they are untrained for. Often, however, the issue is not a challenging medical decision; the patient has long since recovered (if there ever was any disease process), and they now have nothing medically important wrong with them. The argument therefore hinges on “ethics” and the distinction between giving patients what they want, or giving them what they and society need. For ethics to create such a conflict at a suggested cost to society of several billion pounds a year suggests a profound ethical muddle.

As the General Medical Council holds the key responsibility for medical ethics in the United Kingdom, this suggests that it holds a trump card for the solution to a substantial percentage of sickness absence in the country. This was not made clear in the Black report,² but it is an area worth exploring further.

Anthony N Williams consultant occupational physician, Working Fit, PO Box 389, Dover CT16 9BF
tonywilliams@workingfit.com

Competing interests: None declared.

- 1 Snashall D. Health of the working age population. *BMJ* 2008;336:682. (29 March.)
- 2 Black C. *Working for a healthier tomorrow*. London: Stationery Office, 2008. www.workingforhealth.gov.uk/documents/working-for-a-healthier-tomorrow-tagged.pdf

Sick note, fit note, no note?

Snashall's description of the current sick note system as an unaudited farce demonstrates a delightful but unusual honesty.¹ *Working for a Healthier Tomorrow* pins considerable faith on moving to a “fit note.” It is unrealistic for general practitioners to be expected not only to assess health but also to understand the detail of job requirements, which may be well outside their personal experience, within a brief consultation.

It's time to demedicalise the process wherever that is possible. Recognition began in 1982 with the extension of self certification from the first three days to the first seven days of absence from work. There is no good reason why this period could not be extended.

A recent small trial in our business, whereby the requirement to submit a sick note was replaced with the offer of early support and assessment by occupational health, was received very positively. The workforce felt a greater degree of trust (rather than the usual suspicion) was being shown towards them by management; and there was even a small reduction in lost working days over the six month trial period.

The problem is substantial. The solution

will not come from tinkering within existing systems.

Martyn J F Davidson group head, Occupational Health, Centrica, Windsor SL4 5GD
martyn.davidson@gmail.com

Competing interests: MD has been invited to take part in discussions with Dame Carol Black's department regarding early interventions in the workplace.

- 1 Snashall D. Health of the working age population. *BMJ* 2008;336:682. (29 March.)

PREDICTING CITATIONS

Validating prediction models

Lokker et al presented an interesting model to predict citation counts for clinical articles.¹ This topic is so important that the paper will probably attract many citations. We want to clarify some of the nomenclature of validation of prediction models, to avoid confusion in future reporting.

The authors randomly divided 1274 articles into a derivation data set of 757 articles for development of a prediction model and a validation dataset for testing of 504 articles, after exclusion of outliers with >150 citations. This procedure is an example of a split sample approach, but the authors refer to it as cross validation. Cross validation would mean that we develop a model in the first part of the data and test it in the second part, and then repeat the procedure with development in the second part and testing in the first.

The authors report that explained variation (R²) decreased from 0.60 at development to 0.56 at validation, and refer to this decrease as shrinkage. Shrinkage is not an appropriate term for this decrease; a better label is optimism.^{2,3} Optimism is the phenomenon that prediction models tend to perform more poorly in new data than in the data where the model was developed; it occurs especially when many predictors are considered in relatively small datasets.⁴

Ironically, a need for shrinkage is well illustrated in figure 2, where the residuals are generally positive for low predictions (which were often too low), and generally negative for high predictions (which were often too high).¹ Shrinkage should be applied to the regression coefficients for more reliable predictions.^{2,4,5} How valid is this model to predict citations? Firstly, the authors did not shrink regression coefficients, which implies that high predictions will be too high and low predictions too low for articles fulfilling the inclusion criteria. Secondly, for a future article we cannot know beforehand whether the article is an outlier, i.e. having more than >150 citations. Exclusion of outliers at validation is artificial and should not have been done; it has inflated the R² of the model. As

always with prediction models, future validation is required and may reveal disappointing performance.

Ewout W Steyerberg professor of medical decision making

Hester F Lingsma junior researcher, Department of Public Health, Erasmus MC, PO Box 2040, 3000 CA Rotterdam, Netherlands

E.Steyerberg@erasmusmc.nl

Competing interests: None declared.

- 1 Lokker C, McKibbin KA, McKinlay RJ, Wilczynski NL, Haynes RB. Prediction of citation counts for clinical articles at two years using data available within three weeks of publication: retrospective cohort study. *BMJ* 2008;336:655-7. (22 March.)
- 2 Harrell FE, Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-87.
- 3 Steyerberg EW, Harrell FE, Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774-81.
- 4 Steyerberg EW, Eijkemans MJ, Harrell FE, Jr, Habbema JD. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making* 2001;21:45-56.
- 5 Copas JB. Regression, prediction and shrinkage. *J R Stat Soc, Ser B* 1983;45:311-54.

DITCHING IMPACT FACTORS

Time for the single researcher impact factor

It is indeed time to consider alternatives for the “impact factor” algorithm.¹ For example, it might be more useful to consider the merits and contributions of all the scientific activities of each single researcher instead of measuring only the impact factor numbers. For example, as reported in a recent debate in *Science* about peer reviewers' responsibilities,² writing and finalising an article is a complex process in which reviewers offer a crucial scientific contribution. One possible solution is to create a new index, the single researcher impact factor, which can take into account the number and quality of traditional publications and other activities such as reviewing manuscripts.

Some experimental versions of this new index are under evaluation. The single researcher centred impact factor will ensure that the evaluation of individual scientific impact in the community will be more accurate and could better motivate researchers to review (without frustration), publish, and share their ideas.

Gianluca Castelnuovo researcher in clinical psychology, Istituto Auxologico Italiano IRCCS, Psychology Research Laboratory and Catholic University of Milan, Milan, Italy
gianluca.castelnuovo@auxologico.it

The other authors are Elena Faccio, Gian P Turchi, Alessandro Salvini, Enrico Molinari, and Antonio Imbasciati.

Competing interests: None declared.

- 1 Hobbs R. Should we ditch impact factors? *BMJ* 2007;334:569.
- 2 Perrin WF. In search of peer reviewers. *Science* 2008;319:32.