intolerant and diabetic subjects) with a high sensitivity (0·96) and moderate specificity (0·67). With the cut off at 5·9% the sensitivity decreased to 0·61 but the specificity rose to 1. For an area under the curve of 1100 mmol/l/2 h sensitivity was 1 and specificity 0·95.

We agree with McCance and colleagues that determination of the glycated haemoglobin concentration alone may be an acceptable alternative to the more cumbersome oral glucose tolerance test for screening and diagnostic purposes. We recommend using a glycated haemoglobin concentration of 5·3% as the cut off point for screening for diabetes and glucose intolerance in adults.

MARTIN ROUBICEK
ALICIA GONZÁLEZ SANGUINETI
GLORIA VIÑES

Hospital Privado de Comunidad,
7600 Mar del Plata,
Argentina

1 McCance DR, Hanson RL, Charles M-A, Jacobsson LTH, Pettitt DJ, Bennett PH, *et al.* Comparison of tests for glycated haemoglobin and fasting and two hour plasma glucose concentrations as diagnostic methods for diabetes. *BMJ* 1994;308: 1323-8. (21 May.)
2 McCloud WW, Messenger LJ, Schell R, Wagoman F, Yip KF. Unitized cartridge system for the de-centralized measure of hemoglobin A1c. *Biochem Clin* 1990;14(suppl 2):14.
3 National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 1979;28:1039-57.

Methods for estimating glycated haemoglobin differ

EDITOR,—David R McCance and colleagues report that measurement of the glycated haemoglobin or fasting plasma glucose concentration alone might be an acceptable alternative to measurement of the plasma glucose concentration two hours after an oral glucose load for diagnosing diabetes.[1] Although the authors advise caution against placing undue emphasis on the absolute values of variables measured, some practical problems arise regarding the measurement of glycated haemoglobin concentration.

Use of the glycated haemoglobin concentration for estimating overall glycaemic control is widely accepted in the clinical care of diabetic patients. Several methods for determining the concentration have been developed, which yield different normal ranges.[2] Two methods—namely, electrophoresis for haemoglobin A_1 and, later, high pressure liquid chromatography for haemoglobin A_{1c}—were used in the study of McCance and colleagues.[1] To overcome the confusion, not only the actual values measured but also the normal ranges of the glycated haemoglobin assays should be given in reports comparing different studies.[3]

The same method of measuring glycated haemoglobin concentration is unlikely to be used in different countries or even in different diabetes centres. To detect diabetes, therefore, measurement of the glycated haemoglobin concentration is unlikely to be seen as an acceptable alternative to measurement of the plasma glucose concentration, which is simple and standardised. I believe that the oral glucose tolerance test will remain the gold standard, at least in the near future.

GYÖRGY JERMENDY
Head of department

Bajcsy-Zsilinszky Hospital,
Medical Department,
Budapest,
Hungary H-1106

1 McCance DR, Hanson RL, Charles M-A, Jacobsson LTH, Pettitt DJ, Bennett PH, *et al.* Comparison of tests for glycated haemoglobin and fasting and two hour plasma glucose concentrations as diagnostic methods for diabetes. *BMJ* 1994;308: 1323-8. (21 May.)
2 Boucher BJ, Burrin JM, Gould BJ, John PN, Lewis G, Owens C, *et al.* A collaborative study of the measurement of glycosylated haemoglobin by several methods in seven laboratories in the United Kingdom. *Diabetologia* 1983;24:265-71.
3 Peterson CM, Jovanovic L, Raskin P, Goldstein DE. A comparative evaluation of glycosylated haemoglobin assays: feasibility of references and standards. *Diabetologia* 1984;26:214-7.

Authors' reply

EDITOR,—As Kornel Simon points out, our finding that glycated haemoglobin, fasting plasma glucose, and two hour plasma glucose concentrations are equivalent in predicting microvascular complications of diabetes is related to the high correlation among these variables (Spearman correlation coefficients for the subjects shown in figure 1 of our paper: two hour plasma glucose with fasting plasma glucose, $r=0·75$; two hour plasma glucose with glycated haemoglobin, $r=0·69$; fasting plasma glucose with glycated haemoglobin, $r=0·73$). The relations are not linear and vary with the degree of hyperglycaemia.[1] Simon further suggests that the oral glucose tolerance test is superior for diagnosing diabetes because the two hour plasma glucose concentration has the highest sensitivity. The statistics Simon presents from our paper, however, are not sensitivities but the proportions of the study population whose values exceed the cut off points. The two hour plasma glucose concentration itself does not result in a greater proportion; this is a function of the cut off point chosen.

Variation in cut off points also accounts for the apparent discrepancy cited by Alan J Sinclair. In the earlier paper roughly three quarters of subjects with abnormal two hour plasma glucose concentrations had fasting glucose concentrations below the value defined as abnormal by the World Health Organisation.[2] This simply illustrates that the two values defined as abnormal by the WHO are not equivalent. In our current paper cut off points were identified independently on the basis of the relation with microvascular complications or the separation of the components of a bimodal frequency distribution.

As Damian McHugh points out, one cannot extrapolate exactly from one population to another. His calculations for the predictive value of a positive test result, however, are based on prevalences of diabetes of 50% and 2% and on our values of sensitivity and specificity for retinopathy, not diabetes. His estimates are therefore invalid. We would expect the relative merits of these three tests for diagnosing microvascular disease to be similar in other populations, but this hypothesis needs to be tested.

György Jermendy points out that a non-standardised laboratory test may be of little clinical use, but the problem of reproducibility is not limited to glycated haemoglobin.[3] The choice of a diagnostic test may well be determined by local circumstances.

If the two hour oral glucose tolerance test is the gold standard for diagnosing diabetes, as these correspondents seem to assume, then every other test will, of necessity, be inferior. We suggest, however, that a test result should be considered to be diagnostic of diabetes not because it exceeds an arbitrarily selected value but because it indicates a high risk of specific microvascular complications. Glycated haemoglobin and fasting glucose concentrations are equivalent to the two hour oral glucose tolerance test in this respect and thus may be equally suitable for diagnosing diabetes.

DAVID R McCANCE
Consultant physician

Sir George E Clark Metabolic Unit,
Regional Centre for Endocrinology and Diabetes,
Royal Victoria Hospital,
Belfast BT12 6BA

ROBERT L HANSON, Medical staff fellow
M ALAINE CHARLES, Visiting associate
LENNART T H JACOBSSON, Visiting associate
DAVID J PETTITT, Assistant chief of section
PETER H BENNETT, Branch chief
WILLIAM C KNOWLER, Section chief

Diabetes and Arthritis Epidemiology Section,
National Institute of Diabetes and Digestive and Kidney Diseases and National Institute of Arthritis and Musculoskeletal and Skin Diseases,
Phoenix,
Arizona 85014, USA

1 Hanson RL, Nelson RG, McCance DR, Beart JA, Charles MA, Pettitt DJ, *et al.* Comparison of screening tests for non-insulin-dependent diabetes mellitus. *Arch Intern Med* 1993;153: 2133-40.
2 Harris MI, Hadden WC, Knowler WC, Bennett PH. Prevalence of diabetes and impaired glucose tolerance and plasma glucose levels in US population aged 20-74 yr. *Diabetes* 1987;36: 523-34.
3 McDonald GW, Fisher GF, Burnham C. Reproducibility of the oral glucose tolerance test. *Diabetes* 1965;14:473-80.

Promoting research into peer review

No quick fixes

EDITOR,—Anyone who has been involved at either end of the peer review process will be aware of problems associated with the system. Richard Smith's editorial talks of the failings of a costly process and reflects some of the concerns that seem to have prompted recent proposals to streamline grant reviewing in both the National Institutes of Health in the United States and the Engineering and Physical Sciences Research Council in Britain. Other grant awarding institutions seem likely to follow.

The editorial emphasises the need for further interventionist research into peer review. Interestingly, one recent interventionist study of peer review of academic articles concluded that the process was indeed unreliable,[2] yet provoked a discussion with opinions ranging from agreement[3] to rejection[4] and threw up a number of suggestions for reform of the process.

In the end, it is up to the reader to form an opinion of an article. None the less, any editor will surely want to ensure that the readers are not overburdened by chaff. In doing so, journals can maintain a high "impact factor," but even this is not an infallible measure of scientific merit. Innovative research may, for example, take time to attract interest and citations, and peer review may tend to favour conservative opinions. The whole process of scientific publishing could be described in terms of such tensions and conflicts, and research into peer review will doubtless draw them out. I wouldn't expect it to provide any quick fixes, however, although I like Emiliani's suggestion for the absolute review system (ARS), in which authors review their own work, on the presumption that they are the ones most familiar with it.[5]

SIMON A T REDFERN
Lecturer

Department of Earth Sciences,
University of Cambridge,
Cambridge CB2 3EQ

1 Smith R. Promoting research into peer review. *BMJ* 1994;309: 143-4. (16 July.)
2 Ernst E, Saradeth T, Resch KL. Drawbacks of peer review. *Nature* 1993;363:296.
3 Bradshaw RH, Bubier NE. Reform options for peer review. *Nature* 1993;364:183.
4 Green SA. Reform options for peer review. *Nature* 1993;364:184.
5 Emiliani C. Reform options for peer review. *Nature* 1993;364:184.

Identify referees' institutions

EDITOR,—I wish to make a suggestion regarding some of the problems inherent in the peer review process.[1] The anonymous reviewer who rejects a manuscript will always incite the failed author's enmity, sometimes with justification. A solution to bolster the credibility of the peer review process would be to make available the name of the host institution and the relevant department of each referee without specifying that person by name so that he or she would still be partially anonymous. The dates on which a manuscript was sent to the referee and returned to the journals' editorial office should also be available to the authors if requested.

Implementation of these two steps could, if applied widely, all but eliminate accusations of tardiness, lack of integrity, or bias on the part of an

unknown referee as any evidence of these could then be referred back to the referee's institution. The bench test of a piece of research is its shelf life: will it still be of interest or relevant in 10 years' time? This test could equally well apply to negative findings. Many important discoveries throughout history have initially been considered to be unacceptable on religious or cultural grounds. As the ultimate peer reviewer is the wider medical audience, the medical referee should perhaps just act as the referee without acting simultaneously as the goalkeeper, leaving readers and subsequent events to be the best judge of a paper's worth.

D N POLLER
Senior registrar

Gloucestershire Royal Hospital NHS Trust,
Gloucester GL1 3NN

1 Smith R. Promoting research into peer review. *BMJ* 1994;309: 143-4. (16 July.)

Referees should provide references

EDITOR,—I applaud the *BMJ*'s efforts to improve the quality of the peer review process.[1] A small improvement would come if referees were under the same obligation as authors to provide references in support of statements drawn from the literature. A critical review should be a catalyst to renewed and improved effort. This is not the effect if an author is unable to trace the source of a reviewer's criticism; instead, the temptation is to question the motives or abilities of the reviewer and even to question the probity of the journal he or she represents. This should also apply to reviews of applications for research grants, when a vague, unsupported statement saying "something like this has been done before" will result in almost certain rejection.

If a reviewer is genuinely in a position to give an expert opinion the addition of references should mean little extra work. If a bibliographic database is used[2] there is even no need for extra typing.

PETER FURNESS
Senior lecturer

Department of Pathology,
Leicester General Hospital NHS Trust,
Leicester LE5 4PW

1 Smith R. Promoting research into peer review. *BMJ* 1994;309: 143-4. (16 July.)
2 Jones RG. Personal computer software for handling references from CD-ROM and mainframe sources for scientific and medical reports. *BMJ* 1993;307:180-4.

The role of letters in reviewing research

EDITOR,—R S Bhopal and Alison Tonks have pointed out that the potential of material on the correspondence pages remains underdeveloped and undervalued.[1] An editorial by Charlton published in *Anaesthesia* included some letters to the editor without making any reference to those letters or their authors.[2] Charlton had been advised by the journal's editor that "it would not be normal to refer to the correspondence when referencing an editorial about a topic of major interest" (personal communication.) It is perplexing that recommendations published in correspondence can be used but their authors not credited. It is high time for the editors of the leading international scientific journals to formulate guidance for authors of articles and editorials so that they cite published correspondence when they use the message contained in such letters.

Letters to the editor that are in the form of suggestions or recommendations that do not necessarily comment on a published article should also be considered. Bhopal and Tonks also stated that if literature searches of published reports are to include relevant letters, corrections, and other comments, then a system will need to be developed systematically and reliably to link papers with other relevant material; indexing of all letters to original research must be the first step. In my opinion such indexing should not be restricted to letters responding to original research: it should include other letters of wide interest. Implementing the suggestions put forth in such letters to the editor might be beyond the scope of an individual, demanding participation and commitment of institutions. Some authors might not be backed by fully equipped laboratories and other infrastructure to scientifically validate their ideas. Letters to the editor are one outlet for airing those ideas. These need encouragement, not reproach.

GOPI N KALLA
Anaesthesiologist

PO Box 101, KFMH,
Khamis Mushayt,
Saudi Arabia

1 Bhopal RS, Tonks A. The role of letters in reviewing research. *BMJ* 1994;308:1582-3. (18 June.)
2 Charlton JE. Checklists and patient safety. *Anaesthesia* 1990;45: 425-6.

Statistics notes

Defining sensitivity and specificity

EDITOR,—In their statistics note on the sensitivity and specificity of diagnostic tests Douglas G Altman and J Martin Bland's idiosyncratic use of the term "true positive" is unconventional and unhelpful and will confuse readers.[1] Since 1947, when Yerushalmy introduced the terms sensitivity and specificity to aid understanding of the utility of diagnostic tests,[2] respected authorities on both sides of the Atlantic have used the term true positives to indicate those cases in which the disease is present and the diagnostic test gives a positive result.[3-5] To use the term to mean all cases of the disease regardless of the test result is a redundancy, and to use it in this way when defining sensitivity and specificity further obfuscates what for several students and physicians (and perhaps statisticians?) is already confusing.

GORDON L DICKIE
Associate professor

Department of Family Medicine,
University of Western Ontario,
London,
Canada

1 Altman DG, Bland JM. Diagnostic tests 1: sensitivity and specificity. *BMJ* 1994;308:1552. (11 June.)
2 Yerushalmy J. Statistical problems in assessing methods of medical diagnosis, with special reference to x-ray techniques. *Public Health Rep* 1947;62:1432-49.
3 Rose G, Barker DJP. *Epidemiology for the uninitiated.* London: BMA, 1979.
4 Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine.* Toronto: Little, Brown, 1991.
5 Armitage P. *Statistical methods in medical research.* New York: Wiley, 1977:433.

Regression towards the mean

EDITOR,—J Martin Bland and Douglas G Altman's note on regression towards the mean fails to convey the process underlying the phenomenon except in mathematical terms.[1] I hope that the following account may make the relation between regression towards the mean and the correlation coefficient r more understandable.

Assume that there is perfect correlation between a child's height and the mid-parental height ($r=1$). Then one can predict, without error, the child's height from knowledge of his or her parents' height and vice versa. If the mid-parental height is above the mean then so will the child's be, by an equal amount. There is no regression to the mean.

On the other hand, assume that a person's height is completely unpredictable from knowledge of his or her parents' height. There is no correlation between parental and child height ($r=0$). Tall parents will have had this unpredictable effect act in the direction of making them tall, but their children will have an equal likelihood of being tall or small. The mean height of the children of tall parents will be the mean height of all children. There is complete regression towards the mean.

In reality r lies between 1 and 0; there is a predictable component relating parental height to child height and, in addition, an unpredictable component. If a child's parents are tall then the likelihood is that the chance effect acted to increase the parents' heights above what would have been predicted. These chance effects are not heritable (by definition), and the child's height, on average, is therefore closer to the mean height of all children than the parents' heights are to the mean height of all parents. The larger the unpredictable component relative to the predictable component the smaller is r and the more likely it is that a person's height deviates from the mean because of chance effects. These unpredictable effects will not be seen in the relative whose height is closer to the mean. Therefore the smaller is r the greater is the regression to the mean.

SIMON FLEMINGER
Senior lecturer in psychiatry

Academic Department of Psychiatry,
London Hospital Medical College,
London E1 2AD

1 Bland MJ, Altman DG. Regression towards the mean. *BMJ* 1994;308:1499. (4 June.)

Authors' reply

EDITOR,—We agree with Gordon L Dickie that in the context of diagnostic tests the term "true positives" is usually used to mean those people with the disease in whom the diagnostic test gives a positive result. We think that our meaning was clear and hope that not too many readers were dismayed by our non-standard terminology.

Regression towards the mean is a difficult concept. We hope that *BMJ* readers will find Simon Fleminger's description helpful.

DOUGLAS G ALTMAN
Head

Medical Statistics Laboratory,
Imperial Cancer Research Fund,
London WC2A 3PX

J M BLAND
Reader in medical statistics

Department of Public Health Sciences,
St George's Medical School,
London SW17 0RE

Patients are unwilling to enter clinical trials

EDITOR,—T C B Dehn is right to draw attention to the problems of obtaining informed consent from patients entering randomised trials, particularly when one option entails additional, potentially toxic treatment, such as chemotherapy.[1] But Dehn's letter also illustrates the problem of avoiding bias when describing a trial: Dehn admits that five of six patients refused to enter a trial comparing preoperative chemotherapy with surgery alone for oesophageal cancer because of anxiety that they might get chemotherapy. I wonder if, had the patients been seen by an oncologist, a similar number might not have refused because they did not want to miss out on the "beneficial" effects of the chemotherapy.

A colleague and I are both committed to a particular trial testing the value of prophylactic cranial irradiation in small cell lung cancer. We both think that we are honest in our description of the hazards and benefits, and, while many of my patients refuse because they do not want the risk of toxicity, many of hers refuse because they do not