and contraceptive method, regardless of the presence or absence of symptoms' is confusing and needs clarification. The term screening is only appropriately used when applied to the testing of asymptomatic women with no clinical evidence of infection. When symptomatic women or women with clinical evidence of infection are to be tested, the issue becomes one of investigation rather than screening.

Women at highest risk of chlamydial infection in Canada are sexually active women between 15 and 25 years of age who use non-barrier contraceptive methods.[1] Additional risk facors include intercourse with two or more partners per year, a new partner within the preceding year, low socioeconomic class, intermenstrual bleeding, cervical friability and purulent cervical discharge.[1]

Screening asymptomatic sexually active young women who have new partners and who do not use barrier contraception makes a lot of sense for Canadian family doctors. Screening asymptomatic sexually active older women in stable mutually monogamous relationships arguably makes no sense whatsoever because the probability of a false positive test is high — test predictive values are dependent on disease prevalence. Symptomatic women and asymptomatic women with clinical evidence of infection need investigation and treatment, not screening.

The possibility of a false negative chlamydia test can be virtually eliminated in asymptomatic women with no clinical evidence of infection if a gram stain of endocervical secretions shows fewer than 10 pus cells per high power microscope oil-immersion lens field. If tests are negative for gonorrhea and chlamydia, the presence of 10 or more pus cells suggests either a false negative chlamydia test or undiagnosed ureaplasma infection.[2]

JAMES MCSHERRY

Victoria Family Medical Centre
60 Chesley Avenue
London
Ontario N5Z 2C1
Canada

References

1. Gully P. Chlamydial infection in Canada. *Can Med Assoc J* 1992; **147:** 893-896.
2. Brunham RC, Paavonen J, Stevens CE, *et al.* Mucopurulent cervicitis — the ignored counterpart in women of urethritis in men. *N Engl J Med* 1984; **311:** 1-6.

Glaucoma screening

Sir,

Sheldrick and Sharp's paper on glaucoma screening is a timely reminder of the importance of glaucoma as a cause of visual loss and the need for a national screening programme (December *Journal,* p.561).

In 1992, a postal questionnaire was undertaken of general practitioners within the catchment area of a district general hospital in North Yorkshire, in order to assess their beliefs about glaucoma screening. A total of 99 replies were received from 120 questionnaires (83%) with at least one reply from every practice. A total of 85 respondents (86%) were aware that it is possible to screen for glaucoma, while (14%) were not aware that this is possible.

Three practices reported performing some form of screening, using a variety of techniques, such as direct applanation tonometry, fundoscopy and visual field analysis using oculokinetic perimetry charts. The screening was performed by doctors, all of whom held appointments as clinical assistants in ophthalmology or who had a special interest in the eye. Diabetic patients were screened in all three practices, with other groups at risk being screened either opportunistically (one practice) or in an organized fashion.

The main barriers to screening in the other practices were reported to be lack of equipment, mentioned by 56 respondents (72%), lack of knowledge 33 (42%), a perception that it was someone else's job 32 (41%), and lack of time 31 (40%). When all respondents were asked whom they thought should screen for glaucoma, optometrists were suggested by 83 general practitioners (84%), ophthalmologists by 24%, and general practitioners by 24%. Eighty respondents (81%) said that they would consider setting up a glaucoma screening clinic if it could be shown to be of benefit.

Although there is little in the way of organized screening happening at present, most general practitioners seem willing to consider it. There findings support the view that if a national screening programme based within general practice is set up, issues concerning education and training, equipment and time must be addressed if it is to succeed.

RICHARD D NEAL

Centre for Research in Primary Care
University of Leeds
30/32 Hyde Terrace
Leeds LS2 9LN

Living up to expectations?

Sir,

I was interested to read your editorial about the current state and status of the *Journal* (January *Journal,* p.3) and I applaud the achievements outlined. However, I would like to draw attention to one aspect of the *Journal* which I believe reflects badly on the rest of the contents: ironically, the forum for this point of view, the correspondence columns.

The content, quality and scientific worth of published letters are extremely variable. What form of control is exerted over publication in this section? Your editorial mentions correspondence only in the context of it being somewhere to publish material which cannot justifiably be published elsewhere in the *Journal.* Is this why the correspondence columns occasionally come across as a dustbin for unscientific whimsy?

A letter on cervical screening is a case in point.[1] This letter seemed manifestly unscientific, perpetuated medical mythology and reached an invalid conclusion, as has already been commented upon.[2] Publishing material such as this sits uncomfortably with a journal which 'gives scientific respectability to general practice' and which has a 'responsibility for scientific vigour'.

I accept that the correspondence columns can, and should, provoke controversy and stimulate debate. It should not, however, devalue the rest of the *Journal.*

KEITH HOPCROFT

Laindon Health Centre
Laindon
Basildon
Essex SS15 5TR

References

1. Arnot AM. Cervical screening [letter]. *Br J Gen Pract* 1994; **44:** 478-479.
2. Seamark C. Vaginal examination [letter]. *Br J Gen Pract* 1995; **45:** 110-111.

Sir,

Your editorial (January *Journal,* p.3) raises some interesting questions. By the yardstick of scientific quality the *Journal* makes steady progress and has obviously established itself in its field. If this should simply be its purpose then you are entitled to feel satisfied but I have reservations about whether you are adequately fulfilling the needs of the wider membership of the Royal College of General Practitioners. In its present form, *Connection* magazine is rather forlorn and in addition has left the *Journal* with an even narrower appeal. A plate of roast beef may be succulent and nutritious but without the trimmings is relatively unappealing to all but the starving.

Over the years I have been intrigued by the piles of unopened issues of the *Journal* and *British Medical Journal* of many of my colleagues and have pondered on what stimulates me to remove the wrappers. You may argue you are not prim-

arily in research publishing to dangle carrots to the half-hearted but I find the appetizers of the *British Medical Journal* inevitably lead me on to read the more erudite contents. Has the *Journal* not a serious obligation to attract and stimulate as broad a readership as it can and does this not therefore make the removal of the wrapper one of your major priorities?

Certain journals have that charismatic something which I call a soul. The *British Medical Journal* has it but so far our *Journal* has not quite found it.

Do we perhaps take ourselves too seriously in trying to achieve a heavy weight image and so lose our sense of proportion and humour? The RCGP is no longer an infant crying out to be heard but a mature adult, so the *Journal* should be mirroring this.

I believe it only requires a subtle change in presentation of papers into different categories with additional 'outsider' comments to increase the debate. In this way each issue has something of interest for everyone. At present the impression is that the *Journal* is being run primarily for the benefit of researchers with the needs of the readership some way behind. Editorial comment also leavens the content and this is where flair and imagination can transform any one issue.

The editorial team, however, is to be congratulated on its efforts and should take heart that we armchair critics can only further the cause for improvement with the odd observation and by trying to cross an occasional 't'.

F L P FOUIN

147 North Deeside Road
Milltimber
Aberdeen AB1 0JS

Leicester assessment package

Sir,

In mounting a detailed statistical challenge to our recommendation that the Leicester assessment package can be used for the assessment of consultation competence in general practice (letter, January *Journal*, p.51), Braunholtz overlooks the fact that assessment of professional competence, in whatever sphere, is not an exact science. In reviewing the complexities of assessing teaching competence, which is directly analogous to clinical competence, one of the foremost experts on assessment has stated '... in the end, validity is judgemental. At best it will be a form of content or construct validity, depending on the consensual judgement of "experts". To hunt for validity in statistical procedures is to pursue a will o' the wisp.'[1]

A search for any assessment package must balance what is reasonable and practical alongside acceptable statistical levels of validity and reliability. Braunholtz may not be aware of the work on assessment of clinical competence reported among other places in the proceedings of six Ottawa conferences. From many studies it appeared that over 30 hours of testing time with structured patient simulations would be required to arrive at reliable scores. In the United Kingdom situation, or elsewhere, it is highly unlikely that such methods could be applied, since both candidates and examiners are practising general practitioners. Indeed, other assessing authorities worldwide are faced with similar problems and we challenge Dr Braunholtz to nominate a single assessment process which achieves a coefficient of 0.95 or even 0.90.

Braunholtz criticizes our methodology for not testing the scores arrived at using the Leicester assessment package against a gold standard. He suggests that such a gold standard 'might be approximated by a number of experienced assessors (say four), each assessing a large number of consultations (say 20) several times over a period of months (say three)'. This is, however, merely a judgement, and some would say a reasonable judgement, but others might insist on six or eight assessors using 70 consultations over six months, just to make sure that all eventualities were taken into account.

Furthermore, since assessors in the real world get the opportunity to judge performance on one occasion rather than repeatedly, any experiment to test an assessment instrument should replicate reality. Furthermore, it is simply not feasible to test and retest candidates on 'several well-spaced occasions', as Braunholtz suggests.

While we accept that the numbers used in our reliability study were small, it was a fully crossed design so that there were no empty cells in the variance table, that is, each level of every facet was crossed with each level of every other facet. This was infinitely superior to nested designs in which true measurements of examiner variance cannot be made. True variance was deliberately introduced into the system by having candidates with a range of expertise. There are only two ways of introducing true variance; one is by having a large number of candidates at the same level, and the other is by having a small number of candidates at a wide variety of levels. Since the first was economically and practically impossible, and since we were investigating the measurement characteristics of the Leicester assessment package scale and the examiners, rather than the candidates, we concluded that this was a reasonable way to proceed. Our study was almost unique in having a fully crossed design and the time involved for examiners, candidates and patients was a true reflection of what is possible in the real world of general practice.

We are concerned that Braunholtz seems to have conceptualized passing and failing in terms of a norm-referenced scale. It is surely axiomatic that a test of competence must be criterion referenced and those who have been working in the area do not suggest a pass/fail cut off in terms of standard deviations below the mean but in terms of basic failures in competence. In all the pilot studies and in our reported research, it has always been possible for assessors using the package to identify the small number of candidates whose competence gives cause for concern. This has also been true on the many occasions when the package has been used for regulatory assessment of doctors consulting with real patients in the examinations for the diploma in family practice (Royal College of General Practitioners/ Kuwait).

We are also disappointed that Braunholtz has interpreted our suggestion that 'all assessors should be trained and calibrated before being sanctioned to assess real candidates... ' as being 'hardly a firm basis for recommending the Leicester package', when it was self-evidently meant to relate to the use of any assessment package.

In conclusion, the Leicester assessment package criteria have been shown to be valid by expert consensus[2] and the package as a whole capable of producing reliable results.[3] Until Braunholtz, or anybody else, can cite an assessment package which has been demonstrated to be more valid, reliable and feasible in the real world, we shall continue to feel justified to recommend the use of the Leicester assessment package in formative and summative assessment of clinical competence in general practice.

ROBIN C FRASER
ROBERT K MCKINLEY
HELEN MULHOLLAND

Department of General Practice
University of Leicester
Leicester General Hospital
Leicester LE5 4PW

References

1. Stones E. Assessment of a complex skill: improving teacher education. *Assess Educ* 1994; **1:** 235-251.
2. Fraser RC, McKinley RK, Mulholland H. Consultation competence in general practice: establishing the face validity of prioritized criteria in the Leicester assessment package. *Br J Gen Pract* 1994; **44:** 109-113.
3. Fraser RC, McKinley RK, Mulholland H. Consultation competence in general practice: testing the reliability of the Leicester assessment package. *Br J Gen Pract* 1994; **44:** 293-296.