



## UvA-DARE (Digital Academic Repository)

### Simple presentation of test accuracy may lead to inflated disease probabilities (letter)

Bachmann, L.M.; Steurer, J.; ter Riet, G.

#### Publication date

2003

#### Published in

BMJ : British medical journal

[Link to publication](#)

#### Citation for published version (APA):

Bachmann, L. M., Steurer, J., & ter Riet, G. (2003). Simple presentation of test accuracy may lead to inflated disease probabilities (letter). *BMJ : British medical journal*, 326, 393-394.

#### General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



## Indiscriminate investigations have adverse effects

Julian H Barth and Richard G Jones

*BMJ* 2003;326:393-  
doi:10.1136/bmj.326.7385.393

---

Updated information and services can be found at:  
<http://bmj.com/cgi/content/full/326/7385/393>

---

*These include:*

### References

This article cites 5 articles, 2 of which can be accessed free at:  
<http://bmj.com/cgi/content/full/326/7385/393#BIBL>

### Rapid responses

3 rapid responses have been posted to this article, which you can access for free at:  
<http://bmj.com/cgi/content/full/326/7385/393#responses>

You can respond to this article at:  
<http://bmj.com/cgi/eletter-submit/326/7385/393>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

---

### Notes

---

To order reprints of this article go to:  
<http://www.bmjournals.com/cgi/reprintform>

To subscribe to *BMJ* go to:  
<http://bmj.bmjournals.com/subscriptions/subscribe.shtml>

## Indiscriminate investigations have adverse effects

EDITOR—The application of evidence based medicine is leading to better treatments by thorough evaluation of treatments based on analyses of risks and benefits. These balance the beneficial clinical gains against the adverse pharmacological and medical effects, using information derived from randomised controlled trials and cost effectiveness studies. In contrast, no such critical approach has been taken for diagnostic tests nor have the consequences and adverse effects of inappropriate investigations been explored. The debate around diagnostic tests has centred largely on minimising the unit costs of the delivery of tests in the light of the enormous increase in the demand for investigations without an obvious and proportionate improvement in health status.<sup>1</sup>

The case report by Krishnan et al highlights an adverse effect of an inappropriate investigation in a woman with hypothyroid induced ascites.<sup>2</sup> The published literature is clear that ascites, and any serous effusion of any aetiology, is associated with raised CA125 concentration.<sup>3</sup> Yet despite this evidence, the interpretation of a false positive result triggered a number of adverse effects and consequences—namely, a clinical consultation by an oncologist, computed tomography of the abdomen, diagnostic laparoscopy, mammography, and oral gastroduodenoscopy. These inappropriate secondary investigations carry considerable physical, emotional, and financial cost.

What can we do to improve the appropriate use of laboratory and radiological investigations? Previous attempts at educating clinical staff have shown only short lived improvements.<sup>4</sup> We need better solutions because there is a vicious amplification cycle in which increases in investigations are mirrored by increases in operative procedures,<sup>5</sup> justified on the basis of the investigations which themselves generate investigations. This increase in test volume increases the probability of error and harm to patients. The discipline of evidence based diagnostics may not exist because we do not know what questions to ask in relation to investigation strategies or because there are no hard end points (such as death or cure) to judge success as in pharmacological studies. That should not be an excuse to ignore a significant problem. Where the definition of a disease is made by laboratory and radiological investigations, it is mandatory

that the error rate and interferences in the tests are recognised.

**Julian H Barth** *consultant in chemical pathology and metabolic medicine*  
j.h.barth@leeds.ac.uk

**Richard G Jones** *senior lecturer and honorary consultant in chemical pathology*  
Department of Clinical Biochemistry and Immunology, Leeds Teaching Hospitals NHS Trust, Leeds General Infirmary, Leeds LS1 3EX

- 1 Winkles R, Dinant G-J. Rational, cost effective use of investigations in clinical practice. *BMJ* 2002;321:783-5.
- 2 Krishnan STM, Philipose Z, Rayman G. Hypothyroidism mimicking intra-abdominal malignancy. *BMJ* 2002;325:946-7.
- 3 Ferrer J, Villarino MA, Encabo G, Felip E, Bermejo B, Vila S, et al. Diagnostic utility of CYFRA 21-1, carcinoembryonic antigen, CA 125, neuron specific enolase, and squamous cell antigen level determinations in the serum and pleural fluid of patients with pleural effusions. *Cancer* 1999;86:1488-95.
- 4 Fraser CG, Woodford FP. Strategies to modify the test-requesting patterns of clinicians. *Ann Clin Biochem* 1987;24:223-31.
- 5 Verrilli D, Welch HG. The impact of diagnostic testing on therapeutic interventions. *J Am Med Assoc* 1996;275:1197-8.

## Simple presentation of test accuracy may lead to inflated disease probabilities

EDITOR—We found that conveying information on the accuracy of tests in non-technical language improved doctors' ability to estimate disease probabilities accurately.<sup>1</sup> We investigated whether doctors might mis-

use such non-technical presentation when considering the probability of endometrial cancer in a patient with positive results on transvaginal ultrasonography.

We presented 263 general practitioners in Switzerland with a pre-test probability of 10%, information that the patient was aged 65, and a positive transvaginal ultrasound result. Ninety two general practitioners (group 1) received no information on the test's accuracy; 92 (group 2) were told that the sensitivity of the test was 80% and specificity 60%; and 79 (group 3) were told that a positive result is obtained twice as frequently in women with endometrial cancer as in those without the disease, reflecting a likelihood ratio of 2. The last two statements are numerically equivalent since the likelihood ratio equals sensitivity/(1-specificity).

The table shows that the degree of over-estimation of diagnostic accuracy varied with the presentational format. As we found previously,<sup>1</sup> almost half of the doctors did not change their probability estimates after they were provided with the patient's age.

We also found that the non-technical format resulted in 25 of the 79 general practitioners in group 3 (32% (95% confidence interval 22% to 43%)) multiplying their pre-test probability by exactly 2. This is theoretically incorrect since, for example, a likelihood ratio of 2 changes a pre-test probability of 40% to 57% only, not to 80%, which requires a likelihood ratio of 6. Unfortunately, in our study, this mistake helped those respondents who did not change their pre-test probability after being given the patient's age to get close to the correct value,

Distributions of attributed likelihood ratios in three groups given different summaries of information on diagnostic accuracy

Group	Median attributed likelihood ratio (25th centile, 75th centile)	Comparison between groups	P value*
<b>Original analysis</b>			
1 (n=92)	9 (3, 69)	1 v 2	0.0193
2 (n=92)	6 (2, 22)	1 v 3	0.0003
3 (n=79)	3 (2, 9)	2 v 3	0.2284
		1 v (2+3)	0.0006
		All	0.0013
<b>Stricter analysis</b>			
1 (n=92)	9 (3, 69)	1 v 2	0.0193
2 (n=92)	6 (2, 22)	1 v 3	0.1636
3 (n=54)	9 (3, 17)	2 v 3	0.5682
		1 v (2+3)	0.0216
		All	0.0599

\*Kruskal-Wallis test.

Formula to convert pre-test probability ( $P_1$ ) into post-test probability ( $P_2$ ):

Pre-test odds × likelihood ratio = post-test odds, where pre-test odds =  $P_1/(1-P_1)$  and  $P_2$  = post-test odds/(1+post-test odds).

Group 1 received no information on the test's accuracy; group 2 were told that the sensitivity of the test was 80% and specificity 60%; group 3 were told that a positive result is obtained twice as frequently in women with endometrial cancer as those without the disease.

Actual likelihood ratio associated with the test result was 2.25.

changing 10% into 20%, corresponding to an attributed likelihood ratio of 2.25.

The table also shows the results after omission of these 25 doctors. The provision of some form of quantitative information still seems advantageous (contrast group 1 *v* groups 2 and 3;  $P=0.0216$ ). However, all comparisons including group 3 are affected by this stricter analysis.

Framing the diagnostic information in the user friendly way that we used for the likelihood ratio may invite doctors to use simple arithmetic and might lead to grossly inflated inferences when pre-test probabilities are high or likelihood ratios are larger.

**Lucas M Bachmann** *research fellow*  
lucas.bachmann@evimed.ch

**Johann Steurer** *director*  
Horten Centre, Zurich University, PO Box Nord,  
CH-8091 Zurich, Switzerland

**Gerben ter Riet** *senior research fellow*  
Department of General Practice, Academic Medical  
Centre, University of Amsterdam, Room J3-354,  
Meibergdreef 15, NL-1105 AZ Amsterdam,  
Netherlands

1 Steurer J, Fischer JE, Bachmann LM, Koller M, ter Riet G. Communicating accuracy of tests to general practitioners: a controlled study. *BMJ* 2002;324:824-6.

## Effect of computerised evidence based guidelines

### Computer support is complex intervention

**EDITOR**—Eccles et al's rigorous approach to the evaluation of a computerised decision support system for the management of angina and asthma accounted for many of the flaws in previous trials of computer support.<sup>1</sup> They were no doubt disappointed that no effect was seen, probably due to low usage of the system.

Although not discussed in the paper, a possible explanation for this is that, given the comparatively high use of computers required for inclusion in the trial, the practices already used simpler computerised templates to promote collection of process of care data. Practitioners may therefore have perceived little further to be gained by using the more detailed decision support system, particularly if it did not allow easy switching between the guideline and the clinical system.

The study by Eccles et al shows the complexity of interventions in primary care that incorporate computerised decision support systems. This complexity needs to be fully accounted for in designing and evaluating such interventions.<sup>2</sup> Even with an apparently well developed piece of software, the trial assumed that offering brief training to a minority of practitioners in each practice would be sufficient for it to be incorporated into the increasingly complex care provided in routine general practice consultations.

Trials of computer support in primary care need to acknowledge this complexity by embedding use of the software in a carefully specified model of care. For the high quality management of chronic disease, this model

will probably require subspecialisation within a general practice, as proposed in the new general practitioner contract.<sup>3</sup>

Providing focused training to key people in a practice and supporting subspecialisation through computer decision support may be a more appropriate approach to chronic disease management in primary care. Future trials of computer support must consider not only the technical features of the software but also the model of service it is supporting and hence the training requirements of potential users. Theoretically derived measures that predict use of the software by practitioners in these trials could provide further important data on the potential role of decision support in clinical practice. Only then can one truly give computer decision support a fair trial.

**Jon D Emery** *Cancer Research UK clinician scientist*  
Department of Public Health and Primary Care,  
University of Cambridge, Institute of Public Health,  
Cambridge CB2 2SR  
jde10@medschl.cam.ac.uk

Competing interests: None declared.

1 Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, et al. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ* 2002;325:941-4. (26 October.)

2 Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercoc P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;321:694-6.

3 Marshall M, Roland M. The new contract renaissance or requirement for general practice? *Br J Gen Pract* 2002;52:531-2.

### Challenge should not be abandoned

**EDITOR**—As a coauthor of the trial of COGENT, a clinical decision support system, I would like to correct any misunderstanding this paper may have caused.<sup>1,2</sup> I head a large centre for health informatics in the United Kingdom and lead the development of the Prodigy clinical decision support system.

The COGENT trial of two computerised guidelines found no differences in a range of measures of the process and outcomes of care, primarily because the system was not used. But these findings should not be extrapolated to other decision support systems.

Readers to whom I have talked have assumed that COGENT guidance and software was based on the current Prodigy system. COGENT used evidence based guidelines from the north of England on the management of asthma and angina and software based on ideas from early Prodigy software. Constraints in the COGENT trial did not allow the software to be tested in practice before the intervention period or the guidance to be reworded for easier comprehension. Major shortcomings were soon apparent, but these problems could not be addressed because the trial method did not accommodate the usual process of software development and guidance formatting. With hindsight, a randomised controlled trial of a new technology (such as a clinical decision support system) should not be have been undertaken until the technology had been shown to be usable and to be regularly used.<sup>3</sup>

Which way forward? In an increasingly complex world, clinicians overloaded with information need computerised decision support systems if their practice is to be evidence based. The challenge of developing and integrating such systems into clinical workflow should not be abandoned. Not to invest in such systems would be as inappropriate as suggesting that the British army should give up its rifles because of their current technical problems.

**Ian N Purves** *professor of health informatics*  
Sowerby Centre for Health Informatics at  
Newcastle, University of Newcastle on Tyne,  
Newcastle on Tyne NE2 4AB  
ian.purves@ncl.ac.uk

Competing interests: INP is a grant holder, Prodigy contract (Department of Health).

1 Eccles M, Grimshaw J, Steen N, Parkin D, Purves I, McColl E, et al. The design and analysis of a randomised controlled trial to evaluate computerised decision support in primary care: the COGENT Study. *Fam Pract* 2000;17: 180-6.

2 Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, et al. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ* 2002;325:941-4. (26 October.)

3 Heathfield H, Pity D, Hanka R. Evaluating information technology in health care: barriers and challenges. *BMJ* 1998;316:1959-61.

### It is good to be honest and say that systems were not used

**EDITOR**—The paper by Eccles et al possesses academic integrity, which is widely lacking in computing research.<sup>1</sup>

I was the main researcher for the first two phases of the Prodigy project and believe that this project has much to teach the Prodigy team. One of the first detailed reports I wrote on Prodigy in 1998 indicated that Prodigy was actually used very little, about seven times a week, and most of the time (88%) users requested to bypass the system ([www.robinbt2.free-online.co.uk/virtualclassroom/chap13/report1.pdf](http://www.robinbt2.free-online.co.uk/virtualclassroom/chap13/report1.pdf)). I am very heartened to see that this type of information is being disseminated rather than suppressed, as was the case with the report I produced.

**Robin Beaumont** *independent health informatics consultant*  
Faculty of Medical Informatics, Royal College of Surgeons of Edinburgh, Edinburgh EH8 9DW  
robin@ieg-net.co.uk

Competing interests: None declared.

1 Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, et al. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ* 2002;325:941-4. (26 October.)

### Opportunity was missed

**EDITOR**—Eccles et al performed a methodologically sound study of a poorly developed intervention.<sup>1</sup> They define a computerised support system as a system that compares patients' characteristics with a knowledge base and then guides a health provider by offering patient specific and situation specific advice.<sup>1</sup>

The intervention developed and tested in their study does not seem to meet these



criteria. It did not depend on patient specific information but entry of a more general Read code. It did not contain a reminder to initiate review of patient care or arrange follow up. How far treatment recommendations depended on the patient's individual clinical review rather than issuing more generic recommendations for treatment is also unclear.<sup>2,3</sup>

The Prodigy system, the intervention around which this study was based, is an electronic version of a paper guideline that is triggered by entry of a prespecified Read code. By making this the only way in which to enter the computerised guideline the investigators ensured a low level of use during the study. General practitioners are unlikely to continue to enter the same Read code at every consultation as it would mean that each participating patient would have multiple duplicate entries of the same Read code in their electronic record.

By excluding any sort of reminder function in their system,<sup>2</sup> the investigators have not accounted for a barrier in managing chronic diseases—registration, recall, and regular review of patients. Analysis of factors that operate in managing angina and asthma should have uncovered such barriers before the start of this study.<sup>4</sup>

Other details about the use of the computerised guideline require clarification. What is the definitive number of patients randomised and followed up in each practice for each intervention? What is the number (percentage) of patients in whom the computer guideline went past the first screen? What is the number (percentage) for whom a complete record entry was made? The authors make no comment on the differential use of the electronic guidelines between the two computer suppliers.

This study reinforces the fact that passive diffusion of guidelines, in electronic or paper format, is an ineffective way to implement best practice.<sup>4</sup> Paying insufficient attention to how a computer interface operates has produced low levels of usage and made the evaluation less useful than it might have been.<sup>2,5</sup> Future studies should take into account the different functions of computer based clinical decision support systems,<sup>5</sup> rather than simply generate suggestions to alter prescribing practice.<sup>2</sup>

**Tom Fahey** *professor of primary care medicine*  
t.p.fahey@dundee.ac.uk

**Liz Mitchell** *lecturer*  
**Frank Sullivan** *professor of research and development in primary care*  
Tayside Centre of General Practice, University of Dundee, Dundee DD2 4AD

**Alan Montgomery** *lecturer in primary care health services research*  
Division of Primary Health Care, University of Bristol, Bristol BS6 6JL

**Peter Gregor** *senior lecturer in applied computing*  
Department of Applied Computing, University of Dundee, Dundee DD1 4HN

Competing interests: None declared.

1 Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, et al. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ* 2002;325:941-4. (26 October.)

2 Randolph A, Haynes RB, Wyatt JC, Cook DJ, Guyatt GH. Users' guides to the medical literature. XVIII. How to use an article evaluating the clinical impact of a computer-based clinical decision support system. *JAMA* 1999;282:67-74.  
3 Hunt D, Haynes RB, Hanna S, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes. *JAMA* 1998;280:1339-46.  
4 Anonymous. Getting evidence into practice. *Effect Health Care Bull* 1999;5:1-16.  
5 Tai S, Nazareth I, Donegan C, Haines A. Evaluation of general practice computer templates: lessons from a pilot randomized controlled trial. *Methods Inf Med* 1999;38:177-81.

### Effect may be function of incentive

**EDITOR**—In their paper Grimshaw et al showed that having guidelines available does not result in people using them.<sup>1</sup> Analogously, Eccles et al showed that having a decision support system available does not lead to people using it.<sup>2</sup> Benson advocated incentives are needed before healthcare workers start using computers.<sup>3</sup>

In contrast to Eccles et al, van Wijk et al showed effects from a guideline decision support system.<sup>4,5</sup> The general practitioners in these studies had incentives to use the tool, whereas such incentives were missing in Eccles et al's design. We believe that authors of papers describing an evaluation of a decision support system should in the future explicitly discuss incentives for and barriers to using these systems.

**Jacobus T van Wyk** *research fellow*  
**Peter W Moorman** *assistant professor*  
**Marc A M van Wijk** *assistant professor*  
Department of Medical Informatics, Erasmus MC, University Medical Centre Rotterdam, PO Box 1738, NL-3000 DR Rotterdam, Netherlands

1 Grimshaw J, Russell I. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993;342:1317-22.  
2 Eccles M, McColl E, Steen N, Rousseau N, Grimshaw J, Parkin D, et al. Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *BMJ* 2002;325:941-4. (26 October.)  
3 Benson T. Why general practitioners use computers and hospital doctors do not—part 2: scalability. *BMJ* 2002;325:1090-3.  
4 van Wijk MA, van der Lei J, Mosseveld M, Bohnen AM, van Bemmel JH. Assessment of decision support for blood test ordering in primary care. A randomized trial. *Ann Intern Med* 2001;134:274-81.  
5 van Wijk MA, der Lei J, Mosseveld M, Bohnen AM, van Bemmel JH. Compliance of general practitioners with a guideline-based decision support system for ordering blood tests. *Clin Chem* 2002;48:55-60.

### Authors' reply

**EDITOR**—We agree that complex interventions should ideally be developed through an iterative process.<sup>1</sup> Exceptions to this include evaluating a preformed intervention that would not otherwise be rigorously evaluated. This applied at the outset of our study, although our intervention drew heavily on the iterative development of Prodigy software. We conducted an integrated process evaluation to understand the results better. This will appear in the *BMJ* shortly.

The NHS has invested large amounts of money in information technology, sometimes for little or no benefit. The evaluation of information technology is complex and multifaceted, but a computerised decision support system can be evaluated as a health technology. Although formative evaluation may be an important element of software development, until someone comes up with better methods of producing unbiased

estimates of effectiveness and efficiency we maintain that all health technologies should be considered evaluable in randomised controlled trials.

Important methodological issues exist about the timing and duration of such evaluations, and we agree with Purves that they should be performed on stable systems. Given the cyclical nature of software development and the self belief and enthusiasm of developers, such points must be pre-specified and enacted to avoid self perpetuating iterative cycles of development with the constant promise of jam tomorrow.

Our description of the system that we evaluated is accurate, and none of the authors dissented from it up to the point of publication.

Data were collected from November 1997 to September 2000, with the intervention running during the last 15 months of this time period. The trial was paused for six months while the software team worked on improvements. The rates of presentation of patients we reported equated to opportunities for the system to be used between twice a day and every other day. Moreover, by the start of the intervention period, Prodigy software had become available and was delivered to trial practices alongside the study software. Our feedback from practices indicated that at least some asked for the Prodigy software to be turned off. This echoes Beaumont's letter and implies that increasing the number of guidelines offered may not be the remedy that Purves suggests.

Two correspondents identified the importance of the issue of training. Contrary to Purves's letter, two people from each practice were invited to a one day training session and the software was installed within 10 weeks by the computer supplier of two thirds of the trial practices. For the second supplier this interval was almost double, owing to unforeseeable commercial considerations in the company. We acknowledged the importance of training while suggesting that what happened was representative of the real world of primary care. We still believe this to be true but support Emery's and Purves's call for better training in service settings.

Fahey et al say that the low levels of use of the system were partly due to requiring the entry of a single Read code and lack of responsiveness to patient specific information. Initially the system could be triggered automatically by a range of specified Read codes in the patient record. It could also be triggered by a clinician entering Read codes selected by the practice and was therefore not a passive method of dissemination. But this was changed in response to requests from the study practices. The automatic triggering was removed and a customisable Read code entry method was used for the final eight months of the intervention. Thus the system did rely on patient specific information.

Emery said that we may have had a ceiling effect due to practices currently using computerised templates. This seems unlikely because only 26% of practices already had

computerised guidelines or protocols for angina and 46% for asthma.

Within Emery's suggestion of specified models of care we see the risk that clinicians and patients in primary care will be constrained to consult in ways that computers can cope with, rather than addressing the challenge of the integration of computers into patient centered consultations.

**Martin P Eccles** *professor of clinical effectiveness*  
martin.eccles@ncl.ac.uk

**Elaine McColl** *national primary care career scientist*  
**Nick Steen** *statistician*

**Nikki Rousseau** *research associate*  
Centre for Health Services Research, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4AA

**Jeremy Grimshaw** *professor*  
Health Services Research Unit, University of Aberdeen, Aberdeen AB25 2ZD

**David Parkin** *professor of health economics*  
Department of Economics, City University, London EC1V 0HB

Competing interests: None declared.

1 Medical Research Council. *A framework for development and evaluation of RCTs for complex interventions to improve health*. London: MRC, 2000.

## Novartis was not in breach of code for "inventing" disease

EDITOR—Ferriman's news item is incorrect on at least two counts.<sup>1</sup>

Firstly, it is not true to state that the authority had imposed no penalty on the company for issuing misleading literature. Novartis, like all companies ruled in breach of the Association of the British Pharmaceutical Industry's code of practice for the pharmaceutical industry, had to undertake that the use of all relevant material and activity would cease forthwith and that it would take all possible steps to avoid a similar breach of the code in the future.

Novartis voluntarily withdrew the material before the Prescription Medicines Code of Practice Authority had been notified of Dr Robert Flowerdew's concerns and well in advance of Novartis being required by the authority to withdraw all relevant material. Four times a year the authority publishes detailed case reports in the *Code of Practice Review*. The review is widely circulated by the authority and is available on request. There is also some secondary publication of the reports. Publicity is seen as a major sanction. I anticipate that the report on this case will be published in this month's edition of the *Code of Practice Review*.

Secondly, the impression from the heading to the article, as above, is misleading. Novartis was not ruled in breach of the code for inventing a disease. Novartis was ruled in breach of the code for giving a misleading impression of the effect of Starlix on cardiovascular mortality and risk as detailed in the main body of the article.

**Heather Simmonds** *director*  
Prescription Medicines Code of Practice Authority, London SW1YA 2DY

1 Ferriman A. Novartis breached code after doctors say it "invented" a disease. *BMJ* 2002;325:1379. (14 December.)

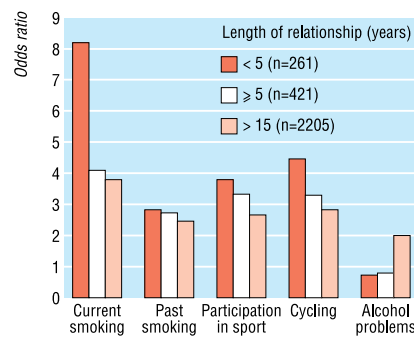
## Assortative mating may explain spouses' risk of same disease

EDITOR—Hippisley-Cox et al observed significant similarities for disease between spouses in a large sample of 8386 couples recruited through general practice.<sup>1</sup> They think that shared environmental factors may cause these similarities but reject assortative mating as an explanation.

In a sample from the Netherlands twin register we could not replicate their spouse similarities for asthma, depression, diabetes, and cardiovascular disease, possibly because of our smaller sample size of 2152 spouse pairs.<sup>2</sup> When we examined health behaviour in a larger sample we found good associations between spouses for smoking, alcohol problems, and exercise behaviour, even after controlling for age and body mass index of both spouses.

The duration of the relationship influenced these associations between spouses (figure). Except for alcohol problems, spouse similarities in health behaviour decreased as the duration of the relationship increased. This implies that assortment for these factors is based on similarity at the time dating began<sup>3</sup> and highlights the importance of determining similarities in disease status at the time of dating, as suggested by Hippisley-Cox et al.

Assortative mating may further be based on social factors and personality traits. In our sample we found significant correlations between spouses for educational attainment, an indicator of socioeconomic status, which is also related to disease development. These correlations increased as the duration of the relationship increased ( $r=0.292$ ,  $r=0.356$ ,  $r=0.587$  for  $<5$  years,  $\geq 5$  years, and  $>15$  years, respectively), possibly owing to convergence of phenotypes of the spouses or to a higher divorce rate in dissimilar pairs.<sup>3</sup> Significant correlations between spouses were also found for inhibition, a personality trait associated with increased risk behaviour, but these correlations were unaffected by the duration of the relationship ( $r=0.386$ ,  $r=0.334$ ,



Odds of smoking (current smoking and ever smoked), participating in exercise (sport and cycling), and alcohol problems in women whose partner had the same lifestyle compared with women whose partner did not. Logistic regression analyses were adjusted for age and body mass index of both men and women

$r=0.373$  for  $<5$  years,  $\geq 5$  years, and  $>15$  years, respectively).

These results show that different mechanisms underlie similarities between spouses for health behaviour, social factors, and personality traits. The fact that similarities between spouses were found for this wide range of variables indicates, however, that assortative mating should not be hastily dismissed as a cause for spouse similarities in disease.

Any association between spouses does not exclude genetic effects. Hippisley-Cox et al assume that because spouses are unrelated, genes do not influence the association. But the similarity of spouses may be an example of an active genotype-environment correlation which occurs when a particular genotype is associated with the selection or creation of a particular environmental circumstance.<sup>4</sup>

**Gonneke Willemsen** *lecturer, epidemiology*  
ahm.willemsen@psy.vu.nl

**Jacqueline M Vink** *researcher, genetic epidemiology*  
**Dorret I Boomsma** *professor, genetic epidemiology*  
Department of Biological Psychology, Vrije Universiteit, NL-1081 BT Amsterdam, Netherlands

1 Hippisley-Cox J, Coupland C, Pringle M, Crown N, Hammersley V. Married couples' risk of same disease: Cross sectional study. *BMJ* 2002;325:1-5. (21 September.)

2 Boomsma DI, Vink JM, van Beijsterveldt CEM, de Geus EJC, Beem AL, Mulder EJC, et al. Netherlands twin register: a focus on longitudinal research. *Twin Res* (in press).

3 Price RA, Vandenberg SG. Spouse similarity in American and Swedish couples. *Behav Genet* 1980;10:59-71.

4 Rowe DC. Assessing genotype-environment interactions and correlations in the postgenomic era. In: Plomin R, DeFries JC, Craig IW, McGuffin P, eds. *Behavioral genetics in the postgenomic era*. Washington, DC: American Psychological Association, 2002:71-86.

## Drug misuse should always be considered in young people with impaired consciousness

EDITOR—We agree with Ikeda et al that the absence of systolic hypertension may provide some discriminatory power towards exclusion of brain lesions, be they ischaemic, haemorrhagic, or space occupying in nature.<sup>1</sup> However, we disagree with them that neurological examination of patients with impaired consciousness is often a waste of time and resources and can delay diagnosis.

Although hypertension may be an important potential marker, a careful neurological examination of the papillary response, reflexes, and funduscopy is an important part of the assessment of any patient with impaired consciousness.<sup>2</sup>

Furthermore, we would like to raise concern over the idea that impaired consciousness in conjunction with systolic hypertension implies that a brain lesion is present. This may be true for older people (the mean age in the Ikeda study was 65 years), but in our experience, impaired consciousness with systolic hypertension in younger people ( $<30$  years) implies ingestion of sympathomimetic drugs—for example, ecstasy, amphetamine, cocaine.<sup>3 4</sup>

Hypertension secondary to ingestion of sympathomimetic drugs requires urgent correction (usually with intravenous nitrates) to prevent secondary complications such as intracerebral haemorrhage, renal failure,



and myocardial ischaemia<sup>3-5</sup>; if, as is said by Ikeda et al, the hypertension and impaired consciousness are assumed to be related to a brain lesion, the delays in obtaining imaging investigations could lead to delays in instituting potentially life saving treatment.

The possibility of illicit drug ingestion should be considered in any young, hypertensive patient presenting to an emergency department with reduced consciousness, so that appropriate management can be started without delay.

**Kim Whelan** *registrar in toxicology*  
kim.whelan@gstt.sthames.nhs.uk

**Alison Jones** *consultant physician*

**Paul Dargan** *registrar in toxicology*  
National Poisons Information Service, Guy's and St Thomas' NHS Trust, London SE14 5ER

- Ikeda M, Matsunaga T, Irabu N, Yoshida S. Using vital signs to diagnose impaired consciousness: cross sectional observational study. *BMJ* 2002;325:800-4. (12 October.)
- Myburgh JA, Oh TE. Disorders of consciousness. In: Oh TE, ed. *Intensive care manual*. Oxford: Butterworth-Heinemann 1997:375-80.
- Ferdinand KC. Substance abuse and hypertension. *J Clin Hypertens* 2000;2:37-40.
- Mangiardi JR, Daras M, Geller ME, Weizner I, Tuchman AJ. Cocaine-related intracranial haemorrhage. Report of nine cases and review. *Acta Neurol Scand* 1988;77:177-80.
- Woodrow PG, Harden P, Turney JH. Acute renal failure due to accelerated hypertension following ingestion of 3,4-methylenedioxymethamphetamine (ecstasy). *Nephrol Dial Transplant* 1995;10:399-400.

## Unit of analysis errors should be clarified in meta-analyses

EDITOR—Weingarten et al present a comprehensive study in what is a complex area of research.<sup>1</sup> We were, however, unclear whether any of the included primary studies had unit of analysis errors and how the authors dealt with such studies in their meta-analysis.

Unit of analysis errors occur in cluster randomised trials when individual patients' data are analysed as if there was no clustering in the provider, practice, or units randomised to the intervention groups (patients' data are analysed as independent observations).<sup>2</sup> Standard statistical methods that do not account for cluster effects in cluster randomised trial data result in the overestimation of the significance of an intervention (artificially extreme P values and overly narrow confidence intervals).<sup>2</sup> Correspondingly, the inclusion of studies with unit of analysis errors in a meta-analysis will give greater weight to the results of such studies.<sup>3</sup>

The table of included studies reported by Weingarten et al indicated that the unit of analysis differed from the unit of randomisation in 22 cluster randomised trials, but it was not clear from the report how often unit of analysis errors occurred in these studies or how the authors dealt with studies with such errors in the meta-analysis. Methods exist for re-analysing studies with such errors.

We recently completed a systematic review of guideline dissemination and implementation strategies; 51 out of 110 cluster randomised trials had unit of analysis errors, and reanalysis was possible in only one study. Poor reporting of cluster randomised trials has led to a proposed

extension to the CONSORT statement, which is currently under discussion.<sup>4</sup> Systematic reviews of studies with unit of analysis errors should clearly state how they handled such studies in a review.

**Ruth E Thomas** *research fellow*  
r.e.thomas@abdn.ac.uk

**Craig R Ramsay** *senior statistician*  
Health Services Research Unit, University of Aberdeen, Aberdeen AB25 2ZD

**Laura McAuley** *review group coordinator*  
Cochrane Effective Practice and Organisation of Care Group, Cochrane Effective Practice and Organisation of Care Group, Institute of Population Health, University of Ottawa, 1 Stewart Street, Ottawa, ON, Canada K1N 6N5

**Jeremy M Grimshaw** *director, clinical epidemiology programme*  
Institute of Population Health, University of Ottawa, Ottawa Health Research Institute, 1053 Carling Avenue, Ottawa, ON, Canada K1Y 4E9

- Weingarten SR, Henning JM, Badamgarav E, Knight K, Hasselblad V, et al. Interventions used in disease management programmes for patients with chronic illness—what ones work? Meta-analysis of published reports. *BMJ* 2002;325:925. (26 October.)
- Whiting-O'Keefe QE, Henke C, Simborg DW. Choosing the correct unit of analysis in medical care experiments. *Med Care* 1984;22:1101-14.
- Donner A, Piaggio G, Villar J. Statistical methods for the meta-analysis of cluster randomised trials. *Stat Methods Med Res* 2001;10:235-338.
- Elbourne DR, Campbell MK. Extending the CONSORT statement to cluster randomised trials: for discussion. *Stat Med* 2001;20:489-96.

## GPs can separate oncological wheat from chaff

EDITOR—Summerton's editorial on the problem of identifying symptoms potentially indicating an underlying cancer in primary care calls for research based theory.<sup>1</sup> Only community based studies will help general practitioners to decide on the importance of a symptom or physical sign reported by their patients as hospital series are unrepresentative.

Although selection bias is part of the problem, it may be comparatively minor as most patients with cancer are treated in secondary care, even though their disease is diagnosed in primary care. Perhaps more important is bias from the development of symptoms in the time from the first symptom appearing to presentation in primary care and finally hospital.

General practitioners are already quite successful in sifting out the wheat from the chaff (although the comparative rarity of cancer in primary care indicates that a better farming analogy would be finding a needle in the haystack). For example, an estimated positive predictive value for colorectal cancer of rectal bleeding in the community is 0.1%.<sup>2,3</sup> Once it is reported to the doctor this rises to 2-3%,<sup>3,4</sup> and in patients referred for investigation it is 5.2%.<sup>5</sup>

**William T Hamilton** *lead researcher*  
12 Barnfield Hill, Exeter EX1 1SR  
wt.hamilton@btopenworld.com

**Alison P Round** *public health consultant*  
Dean Clarke House, Exeter EX1 1PQ

**Deborah Sharp** *professor*

**Tim Peters** *professor*  
Division of Primary Care, University of Bristol, Bristol BS6 6JL

- Summerton N. Symptoms of possible oncological significance: separating the wheat from the chaff. *BMJ* 2002;325:1254-5. (30 November.)
- Fijten G, Muris J, Starmans R, Knottnerus JA, Blijham G, TF K. The incidence and outcome of rectal bleeding in general practice. *Fam Pract* 1993;10:283-7.
- Fijten G, Blijham G, Knottnerus JA. Occurrence and clinical significance of overt blood loss per rectum in the general population and in medical practice. *Br J Gen Pract* 1994;44:320-5.
- Fijten G, Starmans R, Muris J, Schouten H, Blijham G, Knottnerus JA. Predictive value of signs and symptoms for colorectal cancer in patients with rectal bleeding in general practice. *Fam Pract* 1995;12:279-86.
- Selvachandran S, Hodder R, Ballal M, Jones P, Cade D. Prediction of colorectal cancer by a patient consultation questionnaire and scoring system: a prospective study. *Lancet* 2002;360:278-83.

## Look before you flush

EDITOR—Moayyedi and Ford described recent advances in gastroenterology.<sup>1</sup> The national programme for early detection of colon cancer uses the following statement to raise public awareness for early detection of rectal bleeding: "Look at it before you flush it"

It is really helpful to make people aware of looking for early signs of colon cancer, but I have noticed some important things that might hinder this national programme.

The lavatory disinfectants now sold in supermarkets are mostly blue in colour and change water blue, which makes looking for any blood quite difficult. I suggest that we stop selling colouring agents and replace them by colourless ones or even use reagents that turn a certain colour in the presence of minor blood amounts. Can we?

**Mourad Ibrahim Habib** *clinical research fellow*  
St James's University Hospitals, Leeds LS9 7TF  
mouradhabib@hotmail.com

- Moayyedi P, Ford A. Recent advances in gastroenterology. *BMJ* 2002;325:1399-402. (14 December.)

## Correction

*Open letter to Tony Blair: Call to prevent escalating violence*

An editorial error occurred in this open letter to Tony Blair (p 220, 25 January). By adding "the" to the authorship line we implied that the letter had been signed by all staff, students, and alumni of the London School of Hygiene and Tropical Medicine. The authorship line should have read: "On behalf of 500 staff, students, and alumni of the London School of Hygiene and Tropical Medicine, and in collaboration with Medact" [not "On behalf of the staff, students, and alumni..." as published].

**bmj.com**

Letters appearing here are an edited selection of rapid responses originally posted on [bmj.com](http://bmj.com)

We ask for all letters to the editor to be submitted as rapid responses via [bmj.com](http://bmj.com). For advice see: [bmj.com/rapidresponses](http://bmj.com/rapidresponses)