This week in BMJ

US second class postage paid at Rahway, NJ. Postmaster: send address changes to: BMJ, c/o Mercury Airfreight International Ltd Inc, 2323 Randolph Avenue, Avenel, NJ 07001, USA.
US (direct) subscription $227.00.

Published by the proprietors, the British Medical Association, Tavistock Square, London WC1H 9JR, telephone 071 387 4499 (editorial fax 071 383 6418). Printed by BPCC Magazines (Milton Keynes) Ltd, Milton Keynes. Typesetting by Bedford Typesetters Ltd, Bedford. Registered as a newspaper.

Success rates for assisted conception

It is difficult to provide infertile couples with a reliable indication of expected success from in vitro fertilisation and related assisted conception methods because of selective reporting, varying criteria for success, and lack of birth data. In addition, the number of embryos or eggs transferred per treatment is now limited to three. But there have also been substantial improvements in treatment methods. On p 1465 Hull *et al* give an account of all treatment given during 1988-91 through to the completion of all pregnancies. From these results they were able to obtain a reliable indication of prognosis. In women under 40 years and men with normal sperm, whatever their cause of infertility, the cumulative probability of a successful birth per couple (the take home baby rate) was 70% after six cycles of treatment and 90% after nine cycles (with a lower 95% confidence limit of 72%). These results are as good as normal fecundity, except that cycles of treatment are spread over a greater number of months in time.

How good is postmarketing surveillance?

In Britain company sponsored postmarketing surveillance has been subject to voluntary guidelines since 1987. On p 1470 Waller *et al* present a review of the 31 studies so far conducted under the guidelines. The value of most of the studies was limited by a lack of comparator groups. In addition, most of the studies were prospective and may have included selected patients since some studies had inappropriate exclusion criteria. More than one third of studies had been abandoned before sufficient patients had been recruited. The authors conclude that these studies have so far made little contribution to monitoring of drug safety. The existing guidelines require revision based on this experience, and this process is underway.

Reporting to NCEPOD

National audits such as the National Enquiry into Perioperative Deaths (NCEPOD) require correct and complete reporting of all eligible cases. Local reporters appointed to the inquiry rely on patient information systems as their source, whose accuracy is open to question. On p 1472 Clark *et al* report the accuracy and specificity of identification of deaths eligible for reporting to the national inquiry from manual and computerised routine information systems in two hospital units. If their findings are representative of most other hospitals the authors conclude that routine information systems may miss up to half of the eligible deaths.

Routine ultrasound scanning for fetal abnormalities

Screening and prenatal diagnosis programmes seem to be wanted by the population, and they offer medical and financial advantages. Very little has been published on the effect of ultrasound screening in a general population. Luck (p 1474) analysed routine, detailed, 19 week obstetric scans over a period of four years and found that 85% of all anomalies were detected, including all major renal and central nervous system anomalies and most skeletal and gastrointestinal anomalies. Most parents opted for termination when faced with their child having lethal or severely crippling disorders. Parents with babies who had less severe problems were able to prepare for the future, and appropriate delivery sites were planned where necessary.

Effective medical audit in general practice

Few studies have evaluated the effectiveness of medical audit in British general practice. Two papers (pp 1480, 1484) report the results of a comprehensive and rigorous evaluation of medical audit, particularly the setting of clinical standards, among 10 groups of trainers in 62 general practices by the North of England Study of Standards and Performance in General Practice. Each group set standards for one childhood condition randomly selected from five (acute cough, acute vomiting, bedwetting, itchy rash, and recurrent wheezy chest) and randomly received a clinical standard set by another group, summary data comparing clinical performance with that of all other groups, summary data from only their own group, and nothing (control) for each of the four other study conditions. Comparison of trainers' behaviour before and after the standards were set, based on the medical records of 3500 randomly selected children consulting for any of the study conditions, showed improvements in prescribing practice and follow up. These improvements were confined to trainers setting standards for the condition or were significantly greater for them than for the other trainers. Within a random sample of 9000 children standard setting produced improvement in children with recurrent wheezy chest, resulting in better compliance with drug treament and improved respiratory function. The authors conclude that medical audit, particularly setting standards, can be effective in improving clinical performance and patient outcome in general practice.

Dealing with observer variation

Observer variation—both in the same observer and between different observers—is an important problem in studies which rely on medical judgment to assess patients or endpoints. On p 1491 Brennan and Silman review the statistical techniques available to measure observer variability. The commonest way of accounting for variability is the κ statistic, which provides an overall measure of agreement between observers depending on the prevalence of the item being measured. It does not, however, take account of systematic bias, and analysis of observer variation needs to cover both agreement and bias, the authors argue. As well as describing ways of measuring bias the paper also covers variation when there are more than two variables or when the variables are continuous. Once detected both bias and variation need to be minimised by non-statistical means.