hypothesis testing mean that the probability of an inferential error is known if one is rejecting the null hypothesis but not if one is accepting it.² The probability of a type II error in equivalence trials is unknown and can be high with small sample sizes. The conclusion that a difference exists is therefore made on a clearer basis in placebo controlled trials. Equivalence trials are rarely large enough to give the same degree of certainty, and the practical difficulties of obtaining larger sample sizes are more inhibiting than Rothman suggests.

The effect of active treatment is greater in equivalence trials than in placebo controlled trials. For example, in antidepressant trials, placebo controlled trials show a lower efficacy of active drugs than do studies that use no control or an active drug.³ The context of the trial is important, and the placebo component of the response to treatment seems to be greater in equivalence trials than in placebo controlled trials. Evaluating a new drug only in equivalence trials can lead to an inflated estimate of its effect size.

Blinding may not be as critical in equivalence trials as in placebo controlled trials. Unblinding does occur in clinical trials, leading to the potential for bias.⁴ Subjects who have never received any similar treatment are likely to have fewer cues to help them determine whether they are receiving the new or standard treatment in an equivalence trial, but side effects or other cues may help them determine whether they are receiving an active or a placebo drug. Such factors are likely to increase the chance of finding that two active treatments are equivalent, as the degree of unblinding correlates with measured efficacy in a clinical trial.⁵ Restricting the evaluation of a new drug to equivalence studies does not give the opportunity to estimate the bias introduced through unblinding.

Too much of the effectiveness of clinical practice is due to the placebo effect for placebo controlled trials to be abandoned. Placebos are required, if only to provide more data about the effectiveness of active treatment. If blind assessments could be achieved in clinical trials and scientific rigour did lead to the accumulation of medical knowledge then more confidence could be placed in equivalence trials.

D B DOUBLE
Consultant psychiatrist

Southeast Sheffield Mental Health Services,
Sheffield S12 4QN

- 1 Rothman KJ. Placebo mania. *BMJ* 1996;313:3-4. (6 July.)
- 2 Rozeboom AW. The fallacy of the null-hypothesis significance test. *Psychol Bull* 1960;57:416-28.
- 3 Wechsler H, Grosser GH, Greenblatt M. Research evaluating antidepressant medications on hospitalized mental patients: a survey of published reports during a five-year period. *J Nerv Ment Dis* 1965;141:231-9.
- 4 Fisher S, Greenberg RP. How sound is the double-blind design for evaluating psychotropic drugs? *J Nerv Ment Dis* 1993;181:345-50.
- 5 Double DB. Unblinding in trials of the withdrawal of anticholinergic agents in patients maintained on neuroleptics. *J Nerv Ment Dis* 1995;183:599-602.

High quality placebos should be used

EDITOR,—In his editorial Kenneth J Rothman states that "as medical knowledge accumulates, the number of placebo trials should fall" and that "placebo control should no longer be part of the gold standard."¹ I disagree with the latter statement. The main purposes of the placebo are to protect evaluations of outcome from (i) the effects of experimenter bias, by enabling blind assessment; and (ii) the influence of patients' expectations, by controlling for non-specific aspects of health care. Such control is particularly important with treatments that depend on subjective assessment of outcome.

While active drugs given in tablet form under double blind conditions may be indistinguishable from each other, there are many other medical procedures that are very different in appearance and mode of administration, for which it would prove difficult to maintain blindness in a comparative trial. Also, the use of a standard treatment in comparative trials would not necessarily control for the same set of non-specific factors as a placebo, which matches the treatment in every way except for the main component believed to be responsible for therapeutic effects—a component control placebo.² Indeed, many active treatments produce unique side effects, which may act as a cue to the condition to which participants have been assigned in a clinical trial, thereby leading to unblinding.³ Evidence suggests that the inclusion of an active as opposed to an inert placebo in such trials would serve to strengthen the double blind paradigm.⁴

Neither should it be assumed that a standard treatment is necessarily more effective than an appropriate placebo rated as being equally as credible as active treatment. A number of orthodox treatments that are currently in use are of unknown efficacy and would not gain ethical clearance if it were sought today. Furthermore, the use of old treatments in comparative trials would serve little in delineating the mechanism by which a new treatment, if it proves effective, achieves its effects.

Therefore, in the hierarchy of evidence of effectiveness, a randomised controlled trial that uses a component control, active, or credible placebo should feature higher than a comparison with an active drug or medical procedure of unknown effectiveness. The gold standard could be refined by the use of these high quality placebos in place of the standardised "sugar pill" placebo.

AMANDA GEORGIUO
Lecturer in public health

Division of Public Health,
Nuffield Institute for Health,
University of Leeds,
Leeds LS2 9PL

- 1 Rothman KJ. Placebo mania. *BMJ* 1996;313:3-4. (6 July.)
- 2 Horvath P. Placebos and common factors in two decades of psychotherapy research. *Psychol Bull* 1988;104:214-25.
- 3 Greenberg RS, Bornstein RF, Zborowski MJ, Fisher S, Greenberg MD. A meta-analysis of fluoxetine outcome in the treatment of depression. *J Nerv Ment Dis* 1994;182:547-51.
- 4 Fisher S, Greenberg RP. How sound is the double-blind design for evaluating psychotropic drugs? *J Nerv Ment Dis* 1993;181:345-50.

Use of anabolic steroids has been reported by 9% of men attending gymnasiums

EDITOR,—F H Lloyd and colleagues report several cases of subfertility in male body builders who misused anabolic steroids.¹ I would draw the authors' attention to a report commissioned by the Departments of Health for England, Scotland, and Wales, entitled *Anabolic Steroid Use in Great Britain: An Exploratory Investigation*.²

The research for the report investigated use of anabolic steroids in 21 gymnasiums in England, Scotland, and Wales and found that 119 (9.1%) of the 1310 male respondents to the questionnaire and eight (2.3%) of the 349 female respondents had taken anabolic steroids. The youngest user was aged 16. The prevalence of use of anabolic steroids in the gymnasiums ranged from zero (in three gymnasiums) to 46% (28 of 61 respondents). The response rate to the questionnaire was 59% (1677/2834). These findings were confirmed by a similar study by Lenehan *et al*, who investigated the prevalence of use of anabolic steroids in gymnasiums in Merseyside.³

Lloyd and colleagues express concern about male subfertility and raise several important points, including users' lack of knowledge of possible side effects, the high doses taken, and polypharmacy. These concerns were confirmed in in depth interviews with 110 users of anabolic steroids (97 men and 13 women).² Fifty four (56%) of the men reported having experienced testicular atrophy, and eight (62%) of the women reported menstrual irregularities, which suggests that subfertility is also a problem among women. Only 36 of the users had told their general practitioner that they used anabolic steroids. Polypharmacy was common: 55 of the interviewees took three or more different drugs during their "on cycle" of anabolic steroids, and 88 injected them. Use of several other drugs, including human chorionic gonadotrophin, oestrogen antagonists, growth hormone, thyroid hormones, and nalbuphine hydrochloride, was also reported. The findings also suggested that use of other substances is not uncommon: in the previous six months 20 of the interviewees had taken amphetamines, 26 cannabis, and four cocaine. Forty eight reported that dealers were their main sources of anabolic steroids and 28 that friends were; dealers and friends were also the main sources of information about anabolic steroids, along with the "underground" handbooks on steroids.

PIRKKO KORKIA

Senior lecturer in sport and exercise science

University of Luton,
Luton LU1 3JU

- 1 Lloyd FH, Powell P, Murdoch AP. Anabolic steroid abuse by body builders and male subfertility. *BMJ* 1996;313:100-1. (13 July.)
- 2 Korkia P, Stimson GV. *Anabolic steroid use in Great Britain: an exploratory investigation. A report to the Department of Health, the Welsh Office and the Chief Scientist Office, Scottish Home and Health Department.* London: HMSO, 1993.
- 3 Lenehan P, Bellis M, McVeigh J. *Anabolic steroid use in the north west of England. A summary.* Liverpool: Drugs and Sport Information Service, 1996.

Usefulness of urine dipstick tests

False negative results may occur in the absence of antibiotics, ketones, and glucose

EDITOR,—Several studies,^{1,2} including that by J H Beer and colleagues,³ have highlighted the fact that there is still debate over the diagnostic accuracy of rapid dipstick tests. Beer and colleagues point out that the presence of certain antibiotics in urine samples may cause false positive results. Other workers have pointed out that false negative results may be due to antibiotics, glucose, or oxalic acid. Bonnardeaux *et al* concluded that results of dipstick tests were an unreliable predictor of results of microscopy and that false negative results can result from the presence of ketones or glucose.² Hurlbut and Littenburg concluded that "a negative urine dipstick test cannot exclude the diagnosis of urinary tract infection."¹

In a recent study of 1705 clinical urine specimens performed in our laboratory we compared automated (Clintec 200+, Ames) and manual reading of dipsticks (Multistix 8SG in both cases) with quantitative leucocyte determinations made with microchamber slides and a standardised urine analysis system (Kova disposable slides, Hycor Biomedical). We also tested each urine specimen with a variable orifice particle analyser (Questor, Micro-Med). To distinguish positive from negative samples we used a cut off point of 70 white cells/ml for all tests except analysis with the variable orifice particle analyser (for which we used a cut off of 100 white cells/ml). To overcome variability among the test methods a specimen had to give a positive result

by all three test methods (microchamber, standardised urine analysis system, and variable orifice particle analyser) for us to regard the result as positive, or a negative result by all three test methods for us to regard the result as negative. We then compared these results with the combined results of the manual and automated dipstick determinations.

Altogether, 1286, specimens gave negative results in all three control tests, of which 133 gave false positive results on dipstick testing. Eighteen of these false positive results occurred in patients taking antibiotics. More importantly, 194 specimens gave positive results in all three control tests, of which 17 gave false negative results on dipstick testing. Eleven of these false negative results occurred in patients taking antibiotics. Five specimens were positive for ketones and two for glucose. There were various combinations of positivity for antibiotics, ketones, and protein. Three samples did not contain any glucose, ketones, or antibiotics to account for the false negativity result.

If dipsticks are used for screening samples then some false positive results may be acceptable, but the same is not true of false negative results. Although we can confirm that the presence of antibiotics, ketones, and glucose can confound results of dipstick testing, we also found incorrect results in specimens that did not contain these substances.

PETER WILSON Consultant	FLORENCE V CLARKE Clinical scientist
RONALD R CUTLER Clinical scientist	JO MERRETT Medical laboratory scientific officer
PETER JENKS Lecturer	

Newham District Microbiology Laboratories,
St Andrew's Hospital,
London E3 3NT

- Hurlbut TA, Littenburg B. The diagnostic accuracy of rapid dipstick tests to predict urinary tract infections. *Am J Clin Pathol* 1991;96:582-8.
- Bonnardeaux A, Sommerville P, Kaye M. A study on the reliability of dipstick urinalysis. *Clin Nephrol* 1994;41:167-72.
- Beer JH, Vogt A, Neftel K, Cottagnoud P. False positive results for leucocytes in urine dipstick test with common antibiotics. *BMJ* 1996;313:25. (6 July.)

Packaging may lead to false positive results

EDITOR,—J H Beer and colleagues report false positive results for leucocyte esterase on urine analysis in patients concurrently using certain antibiotics.¹ We have encountered false positive results for leucocyte esterase and nitrite on urine analysis.

With Boehringer Mannheim, we are developing a kit for parents to use to test their child's urine for infection after they present their sick infant to primary care. By overcoming some difficulties encountered in primary care (for example, time constraints, availability of equipment)² this test should help general practitioners to investigate for urinary tract infection. Initial results indicate that the test is acceptable to parents, but repackaging the sticks for the kit (in polyethylene with silica gel desiccant) has led to results that are false positive when compared with negative results of microscopy carried out in the laboratory. Tests for nitrite gave positive results in all nine samples studied, and tests for leucocyte esterase gave positive results in six (at 10-25 cells/ul by semiquantitative assay). In one sample a test for leucocyte esterase gave a result indicating 75 cells/ul but microscopy showed 0-5 white cells per high powered field. In contrast, sticks from the manufacturer's original packaging correctly identified 12 sterile urine samples as being negative for both nitrite and leucocyte esterase in the children's outpatient department.

Thus our pilot trial indicated that the packaging led to false positive results. This is probably not insurmountable, and the problem is not as critical

as false negative results would be. Clinical efficacy requires a high sensitivity to exclude urinary tract infection in children presenting with often non-specific symptoms and to obviate further microbiological investigation. If the specificity was also high, a positive result on urine analysis would enable a presumptive diagnosis of urinary tract infection when the presentation was unclear.

The clinical context of such tests, and the prior probability of developing urinary tract infection, is also important.³ A high clinical probability means that a negative result of a dipstick test cannot exclude the diagnosis, as in the febrile leucopenic patients described by Beer and colleagues, and microscopy and culture of urine are mandatory.⁴ When the clinical presentation is undifferentiated or occurs in an area of low prevalence, as is often the case in primary care, then urine analysis may help in distinguishing urinary tract infection from other diagnoses⁵ and identifying those patients who need urine microscopy and culture.

ADRIAN EDWARDS
Clinical fellow
STEPHEN GRANIER
Registrar in general practice

Department of General Practice,
University of Wales College of Medicine,
Llanedeyrn Health Centre,
Llanedeyrn,
Cardiff CF3 7PN

JUDITH VAN DER VOORT
Paediatric registrar

Department of Child Health,
Cardiff Royal Infirmary,
Cardiff CF2 1SZ

- Beer JH, Vogt A, Neftel K, Cottagnoud P. False positive results for leucocytes in urine dipstick test with common antibiotics. *BMJ* 1996;313:25. (6 July.)
- Van der Voort J, Edwards A, Roberts R, Verrier Jones K. The struggle to diagnose urinary tract infection in children under two in primary care. *Fam Pract* (in press).
- Hurlbut TA, Littenburg B. The diagnostic accuracy of rapid dipstick tests to predict urinary tract infection. *Am J Clin Pathol* 1991;96:582-8.
- Lohr JA. Use of routine urinalysis in making a presumptive diagnosis of urinary tract infection in children. *Pediatr Infect Dis J* 1991;10:646-50.
- Lohr JA, Portilla MG, Geuder TG, Dunn ML, Dudley SM. Making a presumptive diagnosis of urinary tract infection by using a urinalysis performed in an on-site laboratory. *J Pediatr* 1993;122:22-5.

Unfair book reviews?

Authors of book hit back

EDITOR,—We were disappointed by David W Denning's review of our book *The Infectious Diseases Manual*.¹ We would have welcomed and learnt from valid criticism, however adverse, but this review contains so many errors of fact that we feel obliged to reply. In particular, we regard the allegation that the book contains multiple errors in drug doses as extremely serious, and we reject it entirely.

Specifically, far from being "rather gentle," the doses of aminoglycoside that we recommend are at the top of the range currently recommended in the *British National Formulary* and the manufacturers' datasheets. We do not recommend once daily treatment as this was not advocated in those sources at the time we wrote the book. We do, however, give guidelines for once daily treatment on p 315. Similarly, the doses given for flu-cytosine and itraconazole are those recommended in the manufacturers' datasheets. Contrary to the assertion that our dose is "dangerously low," we do not recommend any dose of itraconazole for invasive aspergillosis, because we regard this as an option under assessment.

We take exception to our advice on management of neutropenic fever being called "entrenched dogma." This section repeatedly acknowledges the varying policies used in different centres, gives several alternatives for almost every therapeutic decision, and refers readers to

their local infection specialist. Could this section be less dogmatic without being indecisive? Denning's comments seem to be based on a cursory examination of the text, because the points he makes about the use of ceftazidime and vancomycin appear in the body of our text, adjacent to the algorithm about which he complains. Nowhere do we imply that use of vancomycin is "de rigueur"; our wording is, "Some centres include vancomycin from the start; others add it if there is no response at 48h."

We consider that aspergillosis and candidiasis are dealt with at some length in the book. Both receive more attention than tularaemia, not less as Denning implies, and they are certainly described in a degree of detail appropriate for our target readers.

We disagree with several other points in the review, such as the inclusion of dose recommendations for infrequently used antibiotics, but we accept these as matters of opinion on which we can agree to differ.

DAVID WILKS
Consultant physician

City Hospital,
Edinburgh EH10 5SB

MARK FARRINGTON
Consultant microbiologist
DAVID RUBENSTEIN
Consultant physician

Addenbrooke's Hospital,
Cambridge CB2 2QQ

- Denning DW. The infectious diseases manual [book review]. *BMJ* 1996;312:1488. (8 June.)

Reviewer's reply

EDITOR,—David Wilks and colleagues defend their book vigorously—and so they should, as it has much to commend it, as I pointed out in my review. The drug doses they quote are based on the *British National Formulary* and manufacturers' datasheets, but unfortunately some of these are out of date or simply incorrect. Take the recommendation that a patient with a creatinine clearance of <10 ml/min or receiving dialysis should be given amphotericin B 36 hourly. This is simply wrong, just as a maximum dose of 50 mg is.¹ Giving fluconazole at 48 hour or 72 hour intervals for invasive candidiasis in patients with mild or moderate renal failure is exceptionally parsimonious, given an attributable mortality of 35% in this disease² and negligible toxicity at higher doses.³

The problem with the section on neutropenic fever is the huge disparity between the text (with which I broadly agree) and the algorithm, which is greatly at odds with the text. I am sure that the authors will iron out this problem in the next edition.

The coverage of systemic fungal infections in the book is simply inadequate. The common clinical features of these diseases are barely described, let alone the rarer manifestations. Mucormycosis does not appear in the index, and yet 50-100 cases occur annually in Britain, only half of which yield positive results on culture.⁴ A recent necropsy study from Germany showed that 7% of all patients died with or of a fungal infection in the 1990s.⁵ How many cases of brucellosis, chancroid, tularaemia, plague, anthrax, diphtheria, melioidosis, botulism, and onchocerciasis, which are all described in greater depth than invasive aspergillosis or mucormycosis, occur in Britain annually?

The authors set themselves a substantial task, which broadly they have been successful in completing. They should not let my criticisms weigh on them too heavily.

DAVID W DENNING
Senior lecturer in infectious diseases

University of Manchester,
Hope Hospital,
Salford M6 8HD