Statistics in Question

SHEILA M GORE

ASSESSING CLINICAL TRIALS—WHY RANDOMISE?

The principle of randomisation has been challenged for not being expedient[1]: randomised clinical trials impose a discipline on patient selection and accrual and make ethical considerations explicit and multicentre collaboration necessary if the duration of study is to be limited. What has not been shown is that the principle is ill-founded.[2 3] This article discusses reasons for randomisation and reasons against other schemes of assignment. The next two articles explain simple randomisation—that is, the equivalent to tossing a coin—and restricted randomisation procedures. Randomised clinical trials remain the reliable method for making specific comparisons between treatments.

(15) *What are the purposes of randomisation?*

—to safeguard against selection bias

—as insurance, in the long run, against accidental bias

—as the sinew of statistical tests

COMMENT

Firstly, randomisation is the only safeguard against selection bias: the doctor does not know the treatment assignment for the next patient, and the assignment is maximally unpredictable. There can be neither unwitting nor deliberate selection because there is no identifiable pattern in random allocation. When better results are obtained in one treatment group than in others this may be explained (*a*) by the patient groups not being comparable initially, (*b*) by the superiority of a particular treatment, or (*c*) by a marginal effect of treatment that was enhanced because the distribution of patients to that treatment group was slightly favoured. To eliminate *a* and *c* the investigator must at least be able to defend the method of assigning treatment against the accusation of selection bias. The following example (see figure) cannot be so defended—a comparison of subsequent seizure rates in three groups of children: 87 children whose doctors decided to treat them with long-term anticonvulsant drugs; 188 children who were not given prophylactic treatment; and 229 children for whom follow-up was allowed to

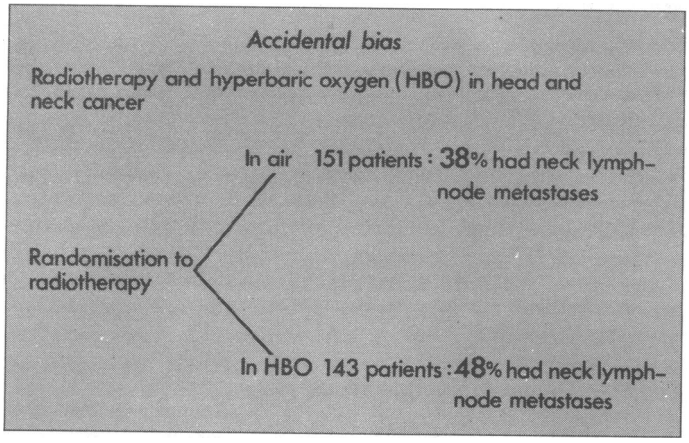
Selection bias

Is the question: *Can recurrent febrile convulsions be prevented?* answerable by comparing subsequent seizure rates in

Group 1	87 children for whom, in a paediatrician's judgment, long-term anticonvulsant treatment was appropriate
Group 2	188 children who were given no prophylactic treatment but were followed up adequately
Group 3	229 children who received no prophylactic treatment and were lost to follow-up

lapse (none of these was given prophylactic treatment). The reason is that allocation of the children to these groups was determined by the paediatrician's assessment and by the stringency of follow-up.

Secondly, randomisation is an insurance in the long run against substantial accidental bias between treatment groups with respect to some important patient variable. The guarantee applies only to large trials; if fewer than 200 patients are randomised a chance imbalance between treatment groups will need careful analysis. In the first Medical Research Council trial of radiotherapy and hyperbaric oxygen for patients with head and neck cancer,[4] 294 patients were randomised to radiotherapy in air or in hyperbaric oxygen. Despite the moderate trial size the treatment groups unfortunately differed in the proportion of patients who had lymph-node metastases in the neck. The hyperbaric oxygen group had the higher proportion of patients —48%—compared with 38% of the 151 patients who had been randomised to treatment in air. A diverse patient population presents a matrix of insufferable complexity.[5] Although randomisation is the surest way through the maze, investigators should still always check that the groups as randomised do not differ with respect to characteristics that are assessed before treatment begins. Nineteen of 28 between-patient trials published in the *Lancet* from July to December 1977 reported making a check



on how comparable the treatment groups were initially; there was some imbalance in respect of prognostic factors in eight of the 19 studies that made a check. Investigators should not be misled into believing that because randomisation guarantees against accidental bias in the long run that it does so in every instance.

Thirdly, the logical foundation for many statistical tests is the premise that each patient could have received any one of the treatments being compared.

In the survey of 38 clinical trials from the *Lancet* referred to in a previous article[6] no trial was uncontrolled, but three trials relied solely on historical controls. A majority of prospective trials were randomised (22 out of 35), but in only one trial was the randomisation procedure described. Methods of randomisation can be learned only if authors refer to them. Mosteller, Gilbert, and McPeek[7] surveyed controlled randomised trials in myeloma and leukaemia, gastrointestinal cancer, and breast cancer. Only one-third of the 132 articles reported the method of randomisation. Mosteller *et al* suggested that authors should briefly describe how randomisation was performed because only then can the reader properly evaluate the trial. They remarked, "When the randomisation leaks, the trial's guarantee of lack of bias runs down the drain."

(16) *Comment critically on the following non-random assignment schemes: (i) by hospital or surgical team; (ii) by assigning treatments alternately; (iii) by hospital number; (iv) by date of birth.*

(i) liable to serious bias: not only treatment, but also patients' environment and medical supervision differ

(ii) by identifying the treatment for the previous patient doctors can deduce the next assignment

(iii) no safeguard against selection bias

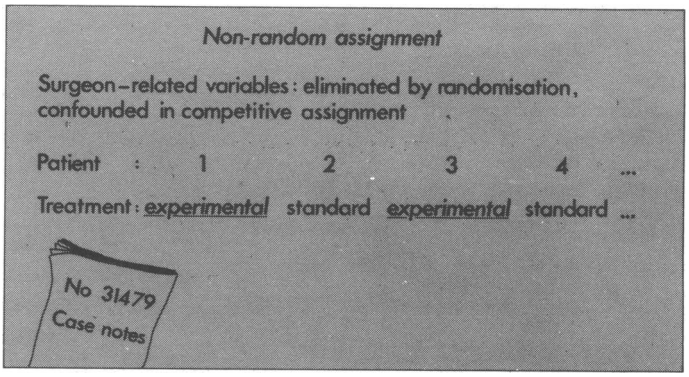
(iv) assignment by date of birth, because it is systematic, invites selection bias;

correlation between morbidity and month of birth

COMMENT

(i) Competitive assignment—that is, when patients referred from different hospitals or by different clinicians receive different treatments—is liable to serious bias. Not only the treatments but also the patients' environment and medical supervision differ,[8] and the latter as much as the former could account for any observed response difference. When the Medical Research Council set up a multicentre trial to compare wound infection rate after major orthopaedic surgery between patients operated on in ultra-clean air and in other theatres, a condition of the trial was that both systems were available in all the trial centres and that the surgical teams operated in both ultra-clean air and in other theatres in random sequence so that, provided the randomisation scheme was adhered to, neither centre nor surgeon would be a source of bias.

Comparison of radical operations in rectal cancer requires that suitable patients are assigned at random to type of operation and that surgical skill is taken into account. A conservative operation is technically more difficult—if it were performed only by specialist teams on selectively referred patients then comparing their results with the outcome of the radical operation by other surgeons could not lead to an unambiguous recommendation of either procedure.



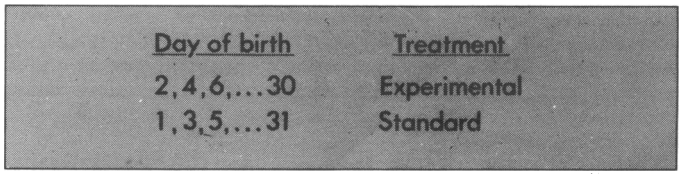
(ii) Alternate assignment of treatments is subject to selection bias if the doctor can identify the treatment for the last patient who entered the trial and so deduce the next assignment. The knowledge may make him defer registration of a particular patient until the treatment he prefers is due for assignment instead of following the unbiased procedure—which is never to register a patient for whom he has a definite treatment preference. Better to limit the population of trial patients than to introduce bias.

Even if the trial treatments are indistinguishable, selection bias can still operate if the doctor knows that the assignment scheme is alternation. All he has to do is guess correctly the treatment for one patient and he can then deduce the treatment for every subsequent patient.

(iii) Assignment by hospital number is unsatisfactory again because there is no safeguard against selection bias. One of the first things that the doctor notices when he looks at a patient's case notes is the hospital number, which is the key to treatment assignment for that patient. The question: "Has forewarning of the treatment group influenced the doctors' decision about whether or not to admit patients to the trial?" will always be asked, and the doubt will linger.

Stead[9] reported that the last digit of the hospital number was used in 1953-5 to assign patients with tuberculosis to three treatments in the ratio 30:30:40. The first and third treatments included streptomycin. Six thousand and nine patients were registered in the trial. Their actual distribution—32% (1898), 26% (1607), 42% (2504)—was significantly different from the target ($p < 0.001$); the regimen that did not include streptomycin was avoided. Randomisation would have eliminated this subjective element from the selection of treatment.

(iv) Date of birth is almost as obvious a detail in the patient's



case notes as is hospital number, and so, because it is systematic, assignment by date of birth invites selection bias. There is a second objection—namely, that there may be a correlation between morbidity and month of birth.[11] If so, then baseline characteristics as well as treatment differ when patients are allocated by date of birth. It is unnecessary to take this risk. Random assignment may be easily achieved.

Prophylaxis of postoperative leg vein thrombosis[10]

Month of birth	Treatment	No of patients
Jan – Apr	Stimulator treatment	88
May – Aug	Low-dose subcutaneous heparin calcium	85
Sept – Dec	Leg exercises only (controls)	122

References

[1] Gehan E, Freireich E. Non-randomised controls in cancer clinical trial *N Engl J Med* 1974;**290**:198-203.
[2] Byar DP. Why data bases should not replace randomized clinical trial *Biometrics* 1980;**36**:337-42.
[3] Meier P. Statistics and medical experimentation. *Biometrics* 1975;**31** 511-29.
[4] Henk JM, Kunkler PB, Smith CW. Radiotherapy and hyperbaric oxyge in head and neck cancer. *Lancet* 1977;ii:101-3.
[5] Black D. The paradox of medical care. *J R Coll Physicians London* 197 **13**:57-65.
[6] Gore SM. Assessing clinical trials—design I. *Br Med J* 1981;**282**:1780-1.
[7] Mosteller F, Gilbert JP, McPeek B. Reporting standards and researc strategies for controlled trials. Agenda for the editor. *Controlled Clinic Trials* 1980;**1**:37-58.
[8] Irwin TT. Surgeon-related variables. *Lancet* 1978;ii:890.
[9] Stead WW. A suggested change in the method of randomization c patients in therapeutic trials. *Transactions of the 16th conference on th chemotherapy of tuberculosis*. St Louis: Veterans Administration, 1957.
[10] Rosenberg HL, Evans M, Pollock AV. Prophylaxis of postoperative le vein thrombosis by low dose subcutaneous heparin or preoperative ca muscle stimulation: a controlled clinical trial. *Br Med J* 1975;i:649-5
[11] Lewis JA. Randomization. *Br Med J* 1975;iii:41.

Sheila M Gore, MA, is a statistician in the MRC Biostatistics Unit Medical Research Council Centre, Hills Road, Cambridge CB2 2QH

No reprints will be available from the author.

A 60-year-old man has suddenly developed attacks of blanching of the fingers lasting 30 minutes or so when walking in the cold. What might be the cause and what advice should be given?

The late onset of Raynaud's syndrome in a man certainly deserves some simple investigation. He should have a full blood count and ESR, and possibly antinuclear factor, to exclude an arteritis. Diabetes and endocarditis also need to be excluded. It would be worth taking *x*-rays of his chest and neck to exclude cervical rib and neoplasms at the thoracic outlet. At the end of the day, however, artheroma of the subclavian, brachial, or digital arteries is the most likely cause. A bruit may be audible over the subclavian or brachial artery. If symptoms are very troublesome an arch aortogram might be justified to localise the atheroma, which might occasionally be amenable to surgery. Simple advice on avoiding precipitating factors usually suffices. In resistant cases reserpine or methyldopa have been advocated.

What is the best method of preventing discoloration of the teeth in the middle-aged and elderly? Is capping of the front teeth justifiable in teeth already discoloured?

The many causes of discoloration of teeth in the middle-aged and elderly may be divided into "extrinsic" and "intrinsic." Extrinsic discolorations are stains deposited on the surface of the teeth, often from nicotine, tea or coffee, or combinations of these, and may be removed without difficulty by the dentist or his hygienist. These surface stains, however, also penetrate into the fine cracks that tend to develop in enamel with age, and these are not easily removed. When oral hygiene is poor dental plaque builds up and can become discoloured and also calcified. When this calcification occurs beneath the gum margin it may be darker in colour, and appears to give the gum and tooth a stained appearance. The common use of fluoridated toothpastes with a lower polishing efficiency than some non-fluoridated toothpastes may also contribute to the build-up of extrinsic stains. In addition there is some evidence that some fluoridated toothpastes may contribute to the surface staining. With aging, the enamel covering may become worn, resulting in the exposure of the underlying dentine, which is darker in colour and less transluscent than enamel. Dentine often stains heavily, to a greater depth than enamel, and is more difficult to clean. Again, the loss of attachment of the soft tissue from the tooth as a consequence of periodontal disease exposes the root surface dentine, which is a different colour and tends to stain.

Changes that occur within the tooth cause "intrinsic" staining. Teeth do darken with age, but these changes are slight in the absenc of other, usually pathological, factors. The commonest of these i probably dental caries, while the darkening of the tooth-tissue adjacen to a restoration may be due to caries adjacent to the restoration or th passage of metalic ions from an amalgam into the adjacent toot substance. Tooth-coloured restorations also pick up surface stain after some time in the mouth, since they are made of resins with fille particles, and may become unsightly. When a single tooth become less transluscent and darker than the adjacent teeth this is often du to the death of the pulp or nerve of the tooth. This must be treate by removing the dead pulp to prevent chronic or even acute infectio affecting the bone surrounding the tooth and to deal with the aestheti problem. Only a dentist can advice in any particular case.

"Capping" is normally referred to as crowning and should b avoided if possible as the cause of discoloration can often be deal with by other methods. Though many crowns are made, failures fron one or more causes are not uncommon. It is not an easy procedure t carry out successfully, but when performed competently a crown ma last a life time and be indistinguishable from a natural tooth. Ba crowning looks unnatural, causes periodontal disease (disease of th gums), kills the nerve of the tooth, interferes with the occlusion (th way the teeth in each jaw relate to each other), and often falls off o breaks. If a single tooth is discoloured and other methods are no indicated to improve the appearances, crowning may not present to many problems. When several teeth are discoloured—for instance, th six upper anterior teeth—then crowning should not be undertake lightly, as this treatment really does need to be carefully planned an carried out by an expert. For example, if the discoloration results fron attrition, then the problems are considerable, for the whole occlusio is affected and many posterior teeth would need crowning, as well a the six anterior teeth to stablise the occlusion, to prevent furthe attrition to other teeth, and to allow a successful aesthetic result to b achieved.

Crowns are expensive as time must be allowed for crowns that d not fit as well as they should, are wrongly contoured or coloured, o have incorrect markings. They are individually made and need to b constructed of porcelain, or porcelain fused to metal, usually gold Resin or plastic crowns should be considered only as temporary, o at very best "semi-permanent" as the materials do not stand up t wear. The colour and arrangement of the anterior teeth can be o great importance to a patient and may be blamed by many for variou problems—the inability to mix with company, a reluctance to smil or laugh in public, or even the breakdown of personal relations Consequently, although contraindicated on all other grounds crowning may be necessary on psychological ones.