STATISTICS IN MEDICINE
Statist. Med. **18**, 3435–3451 (1999)

THE ROLE OF BIOSTATISTICS IN THE PREVENTION, DETECTION AND TREATMENT OF FRAUD IN CLINICAL TRIALS†

MARC BUYSE[1]*, STEPHEN L. GEORGE[2], STEPHEN EVANS[3], NANCY L. GELLER[4], JONAS RANSTAM[5], BRUNO SCHERRER[6], EMMANUEL LESAFFRE[7], GORDON MURRAY[8], LUTZ EDLER[9], JANE HUTTON[10], THEODORE COLTON[11], PETER LACHENBRUCH[12] AND BABU L. VERMA[13]

for the
ISCB SUBCOMMITTEE ON FRAUD

[1] *International Institute for Drug Development, Brussels, and Limburgs Universitair Centrum, Diepenbeek, Belgium*
[2] *Duke University Medical Center, Durham, NC, U.S.A.*
[3] *Medicines Control Agency, London, U.K.*
[4] *Office of Biostatistics Research, National Heart Lung and Blood Institute, Bethesda, MD, U.S.A.*
[5] *University of Lund, Lund, Sweden*
[6] *Institut de Recherche Jouveinal, Fresnes, France*
[7] *Biostatistical Center for Clinical Trials, Universitaire Ziekenhuis St Rafaël, Leuven, Belgium*
[8] *Medical Statistics Unit, University of Edinburgh, Scotland, U.K.*
[9] *Biostatistics Department, German Cancer Research Center, Heidelberg, Germany*
[10] *Department of Mathematics and Statistics, University of Newcastle, Newcastle upon Tyne, U.K.*
[11] *Department of Epidemiology and Biostatistics, Boston University School of Public Health, Boston, MA, U.S.A.*
[12] *Food and Drug Administration, Rockville, MD, U.S.A.*
[13] *Medical College & Hospital, Jhansi, India*

SUMMARY

Recent cases of fraud in clinical trials have attracted considerable media attention, but relatively little reaction from the biostatistical community. In this paper we argue that biostatisticians should be involved in preventing fraud (as well as unintentional errors), detecting it, and quantifying its impact on the outcome of clinical trials. We use the term 'fraud' specifically to refer to *data fabrication* (making up data values) and *falsification* (changing data values). Reported cases of such fraud involve cheating on inclusion criteria so that ineligible patients can enter the trial, and fabricating data so that no requested data are missing. Such types of fraud are partially preventable through a simplification of the eligibility criteria and through a reduction in the amount of data requested. These two measures are feasible and desirable in a surprisingly large number of clinical trials, and neither of them in any way jeopardizes the validity of the trial results. With regards to detection of fraud, a brute force approach has traditionally been used, whereby the

* Correspondence to: Marc Buyse, International Institute for Drug Development (ID^2), 430 avenue Louise B14, B1050 Brussels, Belgium. E-mail: mbuyse@id2.be

† Presented at the International Society for Clinical Biostatistics and the Society for Clinical Trials, Boston, 7–10 July, 1997.

CCC 0277–6715/99/243435–17$17.50

participating centres undergo extensive monitoring involving up to 100 per cent verification of their case records. The cost-effectiveness of this approach seems highly debatable, since one could implement quality control through random sampling schemes, as is done in fields other than clinical medicine. Moreover, there are statistical techniques available (but insufficiently used) to detect 'strange' patterns in the data including, but no limited to, techniques for studying outliers, inliers, overdispersion, underdispersion and correlations or lack thereof. These techniques all rest upon the premise that it is quite difficult to invent plausible data, particularly highly dimensional multivariate data. The multicentric nature of clinical trials also offers an opportunity to check the plausibility of the data submitted by one centre by comparing them with the data from all other centres. Finally, with fraud detected, it is essential to quantify its likely impact upon the outcome of the clinical trial. Many instances of fraud in clinical trials, although morally reprehensible, have a negligible impact on the trial's scientific conclusions.

'Man must learn to simplify, but not to the point of falsification'.
ALDOUS HUXLEY

1. INTRODUCTION

In recent years, the unveiling of a case of fraud in a clinical trial conducted in the U.S.A. and Canada by the National Surgical Adjuvant Breast and Bowel Project (NSABP) has attracted considerable media attention, but surprisingly little reaction from the biostatistical community.[1–12] The statistical profession was largely ignored in the public NSABP debates, and the likely consequences of this clear case of fraud misunderstood. Much of the uproar centred on a gross misrepresentation of the facts, and a lack of explanation of basic principles of clinical research (see Sections 6 and 7).[13] The International Society for Clinical Biostatistics has among its aims 'to promote better understanding of the use and interpretation of biostatistics by the general public, and by national and international organizations and agencies'. It was therefore felt that the Society should support the creation of a subcommittee composed of biostatisticians from various countries, professional backgrounds and institutional affiliations in order to write a position paper on the issue of fraud in clinical research. In this paper, the subcommittee summarizes its views on the appropriate involvement of professional biostatisticians in the prevention, detection and treatment of fraud in clinical trials.

2. BACKGROUND

2.1. Historical perspective

Scientific research has a long history of fraud.[14–17] Over 150 years ago, Charles Babbage, the far-seeing inventor of the calculating machine, established a catalogue of data manipulations, which he called *trimming* (reducing the variance of the data while preserving their mean by deleting extreme observations), *cooking* (reporting only selected observations: '*If a hundred observations are made, the cook must be very unhappy if he cannot pick out fifteen or twenty which will do for serving up*') and *forging* (inventing data).[18,19] Allegations of data tampering have been made against Ptolemy, Galileo, Newton, Dalton, Mendel and Burt, to name just a few.[20] R. A. Fisher's re-analysis of Gregor Mendel's data on peas is a celebrated example of the use of statistical methods to reveal abnormalities; the agreement between the observed frequencies of

certain traits of the peas with the theory was too good to be plausible, which suggested that Mendel or one of his assistants had either manipulated the observations or reported only those results most closely matching theoretical expectations.[21,22] The fraud perpetrated by Cyril Burt was far worse, since it involved the complete fabrication of data on identical twins supposedly separated at birth.[23] Here again the fraud was discovered because of numeric anomalies. The number of identical twins reported by Burt (53 pairs) was too large to be plausible and, while the number of pairs increased from less than 20 to over 50 in a series of Burt's papers, the average correlation of IQ measurements between pairs remained unchanged to the third decimal place![24,25] Over the last decades, prominent cases of fraud in biomedical research involved cancer researcher Marc Straus of the Eastern Cooperative Oncology Group in the seventies, cardiologist John Darsee of Harvard University in the eighties, oncologist Roger Poisson of the NSABP or obstetrician Malcolm Pearce in the nineties.[14,20,26–28] These investigators work for highly respected institutions, yet they are reported to have engaged in what science considers the most serious and unforgivably deviant behaviour; falsifying and fabricating data.[29,30]

2.2. Definitions

Fraud comes in many guises, including some that seem well intended.[9] The boundary between fraud and simple carelessness is often fuzzy, although the former is characterized by a *deliberate* attempt to deceive.[31] The deliberate character of fraud may be very hard to prove in the absence of positive external evidence or confession. Data discrepancies expected as part of the research process, such as transcription errors between the source documents and the case report forms, may potentially be regarded as fraud if they occur in some systematic way or with abnormally high frequency, two circumstances that require a statistical assessment. In many cases, statistical evidence is, however, likely to reveal misunderstandings and unintentional errors rather than fraud.[32,33]

In the U.S.A. the term 'fraud' implies injury or damage to victims, hence the term 'misconduct' might be preferred.[20] However, 'misconduct' also includes practices that fall beyond the scope of this paper, such as plagiarism, conflicts of interest, misuse of funds, and other questionable research practices.[34]

In this paper, we shall use the term 'fraud' specifically to refer to *data fabrication* (making up data values), and *data falsification* (changing or eliminating data values). We are aware that this use of the word is at once far more restrictive than is implied in normal conversation, and less specific than in legal texts, but we prefer the use of this single word for brevity.

3. EXTENT AND INTENT OF FRAUD

The first question that arises with regard to fraud is to quantify its importance. In this section, we discuss the prevalence, perception and intent of fraud, while its possible impact on the results of clinical trials will be deferred to Section 6.

3.1. Prevalence of fraud

Scientific fraud (in the limited sense of data fabrication or falsification) is, in all likelihood, a rare phenomenon, although other misconduct may well be common.[35–38] Fewer than 100 cases of fraud were recognized between 1980 and 1990, an amazingly low figure compared to the number

of scientific research projects conducted in that period.[34] Most authors agree that fraud is also uncommon in clinical trials.[20,39–41] A few cases were uncovered, but they attracted so much media attention that the uncritical observer may have been misled into thinking that the problem was far worse than it actually is.[42–44] While there may be substantial bias in estimating the actual number of cases of fraud (because of those cases that remain unnoticed or unreported), all systematic investigations carried out to uncover fraudulent data found that the proportion of investigators who had actually committed fraud was less than 1 per cent. Miers claimed an 'almost insignificant' prevalence of fraud in research funded by the U.S. National Institutes of Health (NIH).[45] The Cancer and Leukemia Group B (CALGB) reported two cases of fraud detected in 691 on-site audits (0·29 per cent) conducted between 1982 and 1992.[46] The Southwest Oncology Group (SWOG) reported no case of fraud in the audits of 1751 patients conducted between 1983 and 1990.[47] A pharmaceutical company reported one fraudulent investigator among 234 random on-site audits (0·43 per cent) conducted in Europe and South Africa between 1990 and 1994.[48] Against these reassuring statistics lingers the possibility that a large number of cases of fraud may have remained completely unnoticed, and the reported cases only constitute the tip of the iceberg. Although this situation remains hypothetical, there have been reports of fraud being detected and then covered up in trials sponsored by pharmaceutical companies as well as in those performed in academic settings.[27] All in all, we lack reliable data to estimate the true prevalence of fraud, and further prospective investigations in this matter would be very valuable.

3.2. Perception of fraud

Fraud is so much at variance with the ethics of scientific research that *any* amount of it is deemed utterly unacceptable. The Food and Drug Administration (FDA) has for many years had a programme of routine data audits for clinical trials used for new drug approval.[49] Between 1977 and 1988, 11 per cent of the 1955 data audits revealed some serious deficiencies, among which 4 per cent led to a for-cause investigation. These percentages dropped, respectively, to 8 per cent and 1 per cent between 1985 and 1988. In many cases the for-cause investigations revealed sloppiness or incompetence rather than fraud. Hence, as in the studies quoted above, the prevalence of fraud must have been much below 1 per cent. None the less, Shapiro concluded, 'scientific misconduct is common enough in investigational drug trials to be a continuing public concern'.[50] While one must exercise vigilance about fraud in clinical research as much as (or even more than) in other scientific endeavours, it is quite debatable that it should be a 'cause for public concern'.[51] Quantitatively, at least, the opposite seems true: fraud in clinical trials is so rare and, as we discuss below, generally inconsequential, that the public may be far more misguided by studies that are poorly designed, wrongly analysed and inappropriately reported than by fraud.[33,52–58] The concern with fraud in clinical research may in fact be due to common practices that are scientifically unacceptable, such as data dredging, *post hoc* analyses, selective reporting of the most 'interesting' results, non-publication of negative findings etc.[59,60] While these practices may profoundly bias the results of a study as well as their interpretation and dissemination, they do not constitute the focus of the present paper.

3.3. Intent of fraud

The major difference between fraud and mere error lies in the 'intention to cheat' that defines fraud.[61] This difference must, however, be qualified by the nature of the intent, as illustrated by

the following examples. Consider, first, the case of data falsification. Suppose that at some time point the diastolic blood pressure of a patient is read as equal to 96 mmHg. If the value is reported as being equal to 100 mmHg, the discrepancy between the value read and the value reported would constitute a case of data falsification. However the physician who read the diastolic blood pressure may have reported 100 mmHg for simplicity, in recognition of the fact that blood pressure varies substantially in the same patient and that the measurement error is of the order of 5 mmHg anyway. The reporting of a round number that closely approximates the truth would not *per se* be wrong. If, however, a value of 100 mmHg made the patient eligible for the trial while 96 mmHg did not, then the biased reporting would be cause for concern. Worse still, if this biased reporting took place in a non-blinded trial in order to make the control group worse, then the charge of fraud and the need for corrective action would be more than justified. The same arguments hold true for data fabrication. Suppose that the level of neutrophils, a required laboratory examination, were truly missing at the last visit of a patient in a certain trial. Any reported value would therefore have had to be fabricated, perhaps by simply carrying over the value of the previous visit. If this had been done in order to avoid a query from the data management centre for a safety variable of secondary interest, there would be less cause for concern than if neutrophils constituted the primary endpoint of the trial.

The most serious cases of fraud are those in which there is an expectation of gain in terms of prestige, advancement, or money. These cases may involve fabricating complete patients or tampering the data in order to obtain a desirable result. These cases may also be the easiest to detect statistically, especially in multi-centre studies, as we shall see in Section 5.

4. PREVENTION OF FRAUD

In fraud, just as in medicine, an ounce of prevention is worth a pound of cure. It is clearly impossible to eliminate all possibilities of fraud in clinical trials, but it is quite possible to take preventive measures against those that are most likely to happen.

4.1. Data items frequently affected by fraud

It is hardly surprising that some data items are more easily affected by fraud than others. Items that frequently appear prone to error and/or fraud include:

(i) *Eligibility criteria*: data may be 'pushed' a little to make a patient eligible for the trial when in fact that patient does not strictly meet the criteria. Many such examples of fraud may have occurred because eligibility criteria were excessively restrictive.[2,62–65]

(ii) *Repeated measurements*: when the same measurements are requested repeatedly over time (such as, for instance, a battery of laboratory examinations), data may be 'propagated' from the previous visit if the measurements are missing for a particular visit.[20] Such imputation of missing values may be appropriate at the time of analysing the data, not at the time of making the observations.

(iii) *Adverse events*: adverse events are likely to be underreported by some investigators (although such underreporting may reveal lack of interest or differences in interpretation rather than fraud).[66–68]

(iv) *Compliance data*: these data are notoriously unreliable if they are based on the number of medications returned ('pill counts'). Whenever compliance information is deemed important, it is advisable to use objective measurements based on blood or urine tests.[69]

(v) *Patient diaries*: a number of cases of data fabrication have been detected through the color and texture of the pen supposedly used on successive days by the patient, the patient's handwriting, etc. The reliability of information collected in patient diaries can often be questioned.[27,70–71]

4.2. Simplicity

Some types of fraud are preventable through a drastic simplification of randomized clinical trials.[72] Two measures may be particularly effective in this respect: a simplification of the eligibility criteria and a reduction in the amount of data requested. These two measures are feasible and desirable in a surprisingly large number of clinical trials, and neither jeopardizes the validity of the trial results. In trials involving treatments known or expected to induce noticeable toxicities, such as treatments for AIDS or cancer, an accurate and thorough reporting of these toxicities is essential in early trials but may be unnecessary in later phase trials.[73] In trials requiring prolonged observation of each patient, the follow-up can generally be kept as simple and no more frequent than in routine clinical practice.[74] Simplicity is essential in trials conducted with a public health intent, especially when these are large, but it can be justified even in trials conducted as part of a new drug development programme, for there is no regulatory requirement that pivotal trials for drug approval be especially complex.[75] However, the risk of failing to get approval for a new drug along with the fear of potential litigation may dominate all other considerations of cost or efficiency, and as a result clinical trials may end up being excessively complex. The growing number of regulations governing the conduct of clinical trials, even with approved drugs, may also have the unintended consequence of making trials ever more complex.[76] As mentioned above, such complexity may be counterproductive and may pave the way to fraud.

4.3. Allowance for missing data

There is obviously no excuse for making up data, but the temptation will be great for investigators to find ways of avoiding long lists of queries in trials conducted in a fastidious way. It is the responsibility of a competent trial organization to make sure that investigators are not submitted to excessive requests for data clarification. Missing data occur in the real world, and thus they should generally be tolerated in clinical trials (except, of course, for the primary endpoint of the trial). While complete data are undoubtedly better than missing data, attempting to collect too much data, and repeatedly demanding complete data on all patients, may be conducive to fraud, even though it does not exonerate the trial participants from committing it!

5. DETECTION OF FRAUD

5.1. Monitoring

The traditional approach to detect fraud has involved monitoring visits to the clinical centres participating in the trial.[66,77,78] Some such monitoring is obviously needed and useful, for many types of fraud would remain completely undiscovered if it were not for the careful checks carried out during these visits.[27,79] However, monitoring is labour intensive and hence expensive, and it too many fail to pick up fraudulent data.[48,49,80] Moreover, the law of diminishing returns indicates that it is not cost-effective to demand 100 per cent verification of all source data.[81] The

approach used for quality control in industrial or laboratory settings could be used so that the monitoring activities are limited to some random selection of the data, with the possible exception of data pertaining to the primary endpoint of the trial. The random selection could be done at the level of the investigators, the patients or the data items themselves. Using such a random sampling scheme, one could estimate the overall data error rate with prespecified precision, and increase the amount of monitoring if the observed rate exceeded some upper limit.

Regardless of the frequency of monitoring, it is feasible and appropriate, although seldom done in practice, to submit clinical trial data to far more extensive computerized checks than the range and consistency checks usually performed at data entry.[82] We describe some of these checks in the following paragraphs. Most data entry and data management software used for clinical trials performs basic checks, such as identifying outliers, but the other checks typically occur at the end of the study along with final statistical analyses, far too late for corrective action. Statisticians should be involved in developing batteries of statistical checks early on in the course of a trial. Most checks involve standard statistical techniques and graphics and could, therefore, be programmed fairly easily prior to trial commencement.[83]

5.2. Principles of statistical approaches

The principles involved in uncovering fraud through statistical techniques rest on the difficulty of fabricating plausible data, particularly in high dimensions.[84] Univariate observations can always be fabricated to fall close to the mean, although preserving their variance is more of a challenge to the inexperienced.[19] Even the astute cheater who takes care in preserving both the mean and the variance may be tripped up by examination of the kurtosis of the distribution. Multivariate observations must in addition be consistent with the correlation structure between their individual components.[85,86] In general, when data are fabricated to pass certain statistical tests, they are likely to fail on others; Haldane referred to this as 'second order' faking.[87]

Another way of checking fabricated data is based on the fact that humans are poor random number generators.[88] Even informed people seem unable to generate long sequences of numbers that pass simple tests for randomness.[19] Digit preference, especially terminal digit preference, or an excess of round numbers may easily reveal data fabrication.[71,89] Benford's law may also be used to check the randomness of the first digit of all real numbers reported by a single individual (or a single centre). This law establishes that the probability of the first significant digit being equal to D ($D = 1, \ldots, 9$) is approximately given by a logarithmic distribution:[90–94]

$$P(D) \approx \log(D + 1) - \log(D).$$

Hence the frequency of 1's as first digit should be as high as 30 per cent, while that of 9's should be lower than 5 per cent, as shown in Table I.

Benford's law, which runs against intuition, has been used successfully to detect fraud in tax returns.[95] More sophisticated techniques are available to check the randomness of digits in a sequence of data values.[96]

Statistical approaches may also take full advantage of the highly structured nature of clinical trials, which are prospective studies, entirely specified in a written protocol and data collection instrument (the 'case report form'), usually involving several centres, and, when comparative, a randomly assigned treatment.[97] Unusual patterns in the data can be detected by comparing each centre or treatment to the others in terms of the distribution of some variables, either taken in isolation (univariate approach) or jointly (multivariate approach). Comparisons between

Table I. Frequency distribution of the first significant digit[94]

First significant digit D	Expected frequency $P(D)$
1	0·301
2	0·176
3	0·125
4	0·097
5	0·079
6	0·067
7	0·058
8	0·051
9	0·046

centres are particularly informative if there are more than a few observations per centre (in which case fraud in any one centre may have a sizeable impact on the overall result). Such comparisons are useful against different types of fraud; for instance, the presence of outliers or the consistency in the effect of treatment may reveal fraud aimed at exaggerating the effect, while the presence of 'inliers' or underdispersion in the data may reveal invented cases.

5.3. Univariate methods

Beyond range checks and missing data checks, which are performed as part of routine data management, one can use other univariate statistical techniques to inspect the data (Table II).

Statistical checks may reveal unusual data patterns that are often the mark of fraud (Table III). Invented or manipulated data tend to have too little variance, no outliers or an abnormally flat distribution.[84,85] Their distribution may be too close to a simple but implausible model, such as a normal distribution with round numbers for the mean and standard deviation.[98]

Since fraud usually occurs in a single centre (except in the unlikely situation of a co-ordinated fraud across several centres), statistical checks must be performed within each centre as well as overall. A comparison of the results reported by different centres may reveal too little variability in one or more centres as compared to the overall variability.[86] Such a comparison may also reveal 'slippage' of one or more centres, the null hypothesis being that the means of the variable of interest are equal, but for random fluctuations, to the overall mean.[99] These tests are not informative if there are many centres and few patients per centre; on the other hand, grouping small centres could mask a problem in any one of them and is therefore not generally advisable.

Table IV shows an example of suspect data in a trial in which the patients were asked to perform a self-evaluation and report it in a diary every day for three weeks.[71] Whereas in most centres the number of self-evaluation days varied from 18 to 25 (the maximum allowed by the diary), in one suspect centre the vast majority of the patients had filled out exactly 21 days. Such perfect compliance with the protocol requirements is simply too good to be true.

5.4. Multivariate methods

Data management usually includes logical checks to ensure the consistency between the values of two or more variables. Multivariate statistical techniques offer more checking possibilities, but

Table II. Some statistical techniques that may be used to uncover fraud

One variable at a time	Descriptive statistics Box and whisker plot Frequency histogram Stem and leaf plots Tests for slippage
Several variables at a time	Cross-tabulation/scatter plot Correlation/regression Cook's distance Mahalanobis' distance Cluster analysis Discriminant analysis Chernoff faces Star (needle, spike) plots Hotelling's T^2 Tests for treatment contrasts
Repeated measurements	Autocorrelations Profiles Polynomial contrasts Runs tests
Calendar time	Residual plots CUSUM Control charts

Table III. Some patterns that may reveal fraud in clinical trial data

One variable at a time	Digit preference Round number preference Too few or too many outliers Too little or too much variance Strange peaks Data too skewed
Several variables at a time	Multivariate inliers Multivariate outliers Leverage Too weak or too strong correlation
Repeated measurements	Interpolation Duplicates Invented patterns
Calendar time	Breach of randomization Days of week (Sundays or holidays) Implausible accrual Time trends

Table IV. Frequency distribution of the number of self-evaluation days in suspect centre and in other centres.[71] The patients were supposed to perform a self-evaluation and report it in a diary every day for three weeks (21 days)

Number of self-evaluation days	Suspect centre	Other centres
18 or less	—	0·13
19	—	0·04
20	—	0·10
21	0·98	0·10
22	0·01	0·13
23	0·01	0·17
24	—	0·10
25	—	0·23

they are seldom used in clinical trials, if at all. Multivariate statistical methods include correlations between several patient-related variables as well as comparisons between the randomized groups (Table II). Simple two-way cross-tabulations or scatter plots for various pairs of variables can be compared across centres, and unusual patterns investigated further. Outlying observations, or outlying groups of observations coming from the same centre, can be detected more effectively in multi-dimensional space than in a single dimension. Moreover, in multi-dimensional space, 'inliers' can be detected through the use of the Mahalanobis' distance just as well as outliers; inliers have an abnormally low Mahalanobis' distance (they fall too close to the multivariate mean), while outliers have an abnormally high Mahalanobis' distance (they fall too far from the multivariate mean (Table III). The Mahalanobis' distance is computed by standardizing the variables of interest (subtracting the mean and dividing by the standard deviation), and summing the squares of these standardized variables. The sum approximately follows a χ^2 distribution with N degrees of freedom, if N variables are considered. The detection of inliers may be more useful to detect fraud than the detection of outliers, because fabricated data will tend not to contain outliers which are at higher risk of being detected than are values close to the (multivariate) mean.[85] This method can also be used to see if the N variables of interest are too close to each other for some pair(s) of individuals, in which case one of the individuals in the pair may be a (slightly modified) copy of the other. Robust methods such as using ranks in place of the observations are advisable for the detection of outliers because these can create severe departures from multivariate normality.

5.5. Repeated measures

When, as is often the case, some variables are measured repeatedly over the course of the trial on the same patient, these measures lend themselves well to a variety of checks (Table II). Here again, an insufficient variability over time may reveal propagation of previous values rather than genuine observations (Table III). Sometimes the fraud involves a mechanism or computer algorithm for making up data. Scherrer reports the case of an invented series of patients.[71] When polynomial contrasts of increasing degree (linear, quadratic, cubic, etc.) were fitted to the response

data over time, all were highly significant up to the seventh degree, which revealed unusually sophisticated data fabrication. Although the data looked plausible and reasonably random, tests for time trends identified highly unlikely deterministic patterns that suggested these data had been fabricated using a computer program.

5.6. Calendar time

In any trial with prolonged patient entry and follow-up, one can use calendar time to perform additional checks on the data (Table II). Simple checks can be performed on the day of the week, as certain events or examinations are unlikely to have taken place on a Sunday. Time intervals between successive visits and the number of visits per unit time provide further opportunities for checking the plausibility of a sequence of events (Table III). A comparison of treatment groups by week or month of randomization can reveal suspect periods during which all treatments were not allocated with equal probability.[40] More advanced checks may sometimes be performed, such as the variance of observations over time. An excellent example is provided by an animal study, in which the variability of the heart rates of dogs treated consecutively showed far too little variance initially, leading to a strong suspicion of data fabrication in the early stages of the study.[100]

6. TREATMENT OF FRAUD

6.1. Impact of fraud on the trial results

The highly publicized case of fraud in the National Surgical Adjuvant Breast and Bowel Project (NSABP) provides a framework to examine the impact of such fraud on the results of clinical trials. Briefly, one of the investigators in breast cancer trials, Dr. Roger Poisson of St Luc's Hospital in Montreal, systematically altered some baseline patient data so that these patients became eligible for entry into the trials.[1] The data subject to falsification were the dates of surgery and biopsy or oestrogen receptor values. For example, in one study, the delay between the surgery and randomization had been set to a maximum of 30 days by the trial protocol, and dates were falsified for a few patients in whom this limit had been exceeded. The fraud was serious as it involved several members of Dr. Poisson's staff and concerned about 7 per cent of all patients entered by St Luc's Hospital in NSABP trials.[20] However, it was clearly not aimed at distorting the results of the trials one way or another (it could only have done so had the treatment effect been substantially different among the wrongly entered patients than among the others). As a matter of fact, a careful reanalysis of NSABP trials without the data from St Luc's Hospital confirmed that the trial outcomes had not been materially affected by the fraud.[11,13] In another large trial published recently, all data from one centre suspected of fraud were excluded from the analysis, again with negligible impact on the study results. Yet this centre had contributed 452 (6·4 per cent) of the 7054 patients randomized in the study, and statistical analysis of the variability of their data supported the belief that no real patients had actually been studied in this centre![101]

Fraud is unlikely to affect the results of a trial if any of the following conditions hold:

(i) the fraud is limited to a few investigators (perhaps one centre in a multicentric setting) and/or to a few data items;
(ii) the fraud bears on secondary variables that have little or no effect on the primary endpoint of the trial;

(iii) the fraud affects all treatment groups equally, and hence does not bias the results of the trial. Fraud committed without regard to the treatment assignments (for example, prior to randomization or in double blind trials) generates noise but no bias.

At least one of these conditions frequently holds, and therefore fraud should not be expected to have a major impact on the results of multi-centre clinical trials. As a matter of fact, a search of Medline from 1966 to 1997 revealed that 235 articles had been retracted, 86 of which were deemed to be due to misconduct.[102] These numbers do not bear specifically on clinical trial reports, but they are quite small compared to the total number of articles published during the same period.

6.2. Actions in cases of fraud

When fraud is suspected in a centre, all analyses can be repeated with and without that centre, in order to assess the sensitivity of the trial results to the fraud.[101] Although fraudulent data would in general have to be excluded from the main analysis of the trial, other validated data from the same centre might well be kept in the analysis. If the trial is overlooked by a Data Monitoring Committee, it seems appropriate to leave such decisions to their discretion.[103]

Biostatistical methods can only point at problems; further investigations and hard evidence are needed to confirm fraud.[71,104] When fraud is confirmed, the only sensible course of action is to acknowledge it. Incidentally, this had been the case in the NSABP affair, which indicates that whistleblowing is not without dangers to those who uncover the fraud.[105] The NSABP Chairman and Chief Statistician were removed from their positions, even though their personal integrity had at no time been questioned. These removals were decided under considerable media and public pressure, in contravention to elementary requirements of a properly conducted investigation of alleged misconduct:

> 'Staff who are the subject of allegations are entitled to expect that their work will be regarded as honest unless proved to be otherwise, and that they will be protected against ill-founded, frivolous, mischievous, or malicious compliants. […] The demands of justice also require that arbitration and appeal arrangements are available. Responsibility for establishing such a facility might be undertaken by national funding agencies and/or professional bodies.'[106]

Further discussion of appropriate courses of action in case of fraud falls for beyond the scope of the present paper, as does the complete lack of formal instruction in ethical conduct in science training programmes.[107] Useful guidance is available in several papers,[108–111] in the booklet *Integrity and Misconduct in Research* edited by the U.S. Department of Health and Human Services[38] and in the book *Fraud and Misconduct in Medical Research* edited by Lock and Wells.[27] We believe that scientific societies such as the International Society for Clinical Biostatistics could play a role in providing independent advice on actual cases of suspected fraud.

7. DISCUSSION

Randomized clinical trials constitute, by design, the most reliable type of medical experiment. Their data can be verified using statistical techniques that take advantage of their highly structured nature. Their results are robust to occasional cases of data falsification and fabrication at some participating centres. As George put it, 'the methodology of clinical trials *de facto* provides a measure of protection against deliberate deception that is generally unappreciated by

those not familiar with the methodology'.[39] These observations do not condone the perpetration of fraud, but they put actual cases of fraud in clinical trials into their proper perspective. The NSABP case is exemplary because the public was either misinformed or seriously misguided on the substantive issues involved.[13] The statistical profession was largely ignored in the public NSABP debates. Perhaps the most disturbing fact in the NSABP case is that the premise upon which it was based was fundamentally flawed. Once the word 'fraud' was out, public outrage was inevitable, because patients with a serious illness were involved and the evaluation of common medical practices was being questioned. It should have been emphasized that the falsifications were not aimed at distorting the outcome of the trial, and that there was no way they could have done so.

Claims to the contrary notwithstanding, we did not find quantitative evidence that fraud is common in clinical trials. However, fraud is a cause for concern regardless of its prevalence or consequences because the 'habit of truth' is the cardinal value in scientific endeavours.[112,113] Fraud must be fought, but attempts to impose more bureaucracy and heavier monitoring on clinical trials is the wrong answer to an over-rated problem.[114] As the Presidents of the U.S. National Academy of Sciences and of the Institute of Medicine wrote:

> 'If we do not police ourselves, others may step in to do so. The result could be a scientific enterprise that is increasingly constrained by legal strictures, financial oversight, and bureaucratic provisions. […] If scientific research is beset with paperwork and regulation, much of the joy and creativity in doing science could disappear. Such a cultural change would not only impede scientific progress, it would also make our field much less attractive to the dedicated and talented young researchers who represent the future'.[115]

Our view is that fraud can largely be prevented through design of the trial protocol and case report form, and detected by statistical procedures and computerized checks that make use of the unique structure of clinical trial data.

ACKNOWLEDGEMENTS

The views expressed in this paper are those of the authors and of the International Society for Clinical Biostatistics, not necessarily of other organizations with which the authors are affiliated.

REFERENCES

1. Anonymous, 'Final Report of the Division of Research Investigations on St. Luc Hospital NSABP Project', Office of Research Integrity, 1993, No. 91-08, pp. 1–94.
2. Dingell, J. D. 'Shattuck lecture: Misconduct in medical research', *New England Journal of Medicine*, **328**, 1610–1615 (1993).
3. Anderson, C. 'Breast cancer. How not to publicize a misconduct finding', *Science*, **263**, 1679 (news) (1994).
4. Anonymous. 'NCI issues information on falsified data in NSABP trials (news)', *Journal of the National Cancer Institute*, **86**, 487–489 (1994).
5. Angell, M. and Kassirer, J. P. 'Setting the record straight in the breast-cancer trials', *New England Journal of Medicine*, **330**, 1448–1450 (1994).

6. Bivens, L. V. and Macfarlane, D. K. 'Fraud in breast cancer trials', *New England Journal of Medicine*, **330**, 1461 (letter) (1994).
7. Broder, S. 'Fraud in breast cancer trials', *New England Journal of Medicine*, **330**, 1460–1461 (letter) (1994).
8. Fisher, B. and Redmond, C. K. 'Fraud in breast cancer trials', *New England Journal of Medicine*, **330**, 1458–1460. (letter) (1994).
9. Poisson, R. 'Fraud in breast-cancer trials', *New England Journal of Medicine*, **330**, 1460 (letter) (1994).
10. Christian, M. C., McCabe, M. S., Korn, E. L., Abrams, J. S., Kaplan, R. S. and Friedman, M. A. 'The National Cancer Institute Audit of the National Surgical Adjuvant Breast and Bowel Project Protocol B-60', *New England Journal of Medicine*, **333**, 1469–1474 (1995).
11. Fisher, B., Anderson, S., Redmond, C. K., Wolmark, N., Wickerham, D. L. and Cronin, W. M. 'Reanalysis and results after 12 years of follow-up in a randomized clinical trial comparing total mastectomy with lumpectomy with or without irradiation in the treatment of breast cancer', *New England Journal of Medicine*, **333**, 1456–1461 (1995).
12. Smigel, K. 'Top cancer-related news stories focus on fraud, breast cancer and the hope of early detection', *Journal of the National Cancer Institute*, **87**, 12–14 (1995).
13. Peto, R., Collins, R., Sackett, D., Darbyhsire, J., Babiker, A, Buyse, M., Stewart, H., Baum, M., Goldhirsch, A., Bonadonna, G., Valagussa, P., Rutqvist, L., Elbourne, D., Dabies, C., Dalesio, O., Parmar, M., Hill, C., Clarke, M., Gray, R. and Doll, R. The trials of Dr. Bernard Fisher: a European perspective on an American episode', *Controlled Clinical Trials*, **18**, 1–13 (1997).
14. Broad, W. and Wade, N. *Betrayers of the Truth*, Simon and Schuster, New York, 1982.
15. Kohn, A. *False Prophets*, Basil Blackwell, Oxford, 1986.
16. Bell, R. *Impure Science*: *Fraud*, *Compromise*, *and Political Influence in Scientific Research*, Wiley, New York, 1992.
17. Miller, D. J. and Hersen, M. *Research Fraud in the Behavioral and Biomedical Sciences*, Wiley, New York, 1992.
18. Babbage, C. *Reflections on the Decline of Science in England and on Some of its Causes*, B. Fellowes, London, 1830.
19. Rao, C. R. *Statistics and Truth*, International Co-operative Publishing House, Burtonsville, 1989.
20. Piantadosi, S. *Clinical Trials. A Methodologic Perspective*, Wiley, New York, 1997.
21. Fisher, R. A. 'Has Mendel's work been rediscovered?', *Annals of Science*, **1**, 115–137, (1936).
22. Weiling, F. 'What about R. A. Fisher's statement of the "too good" data of J. G. Mendel' s pisum paper?', *Journal of Heredity*, **77**, 281–283 (1986).
23. Fletcher, R. *Science*, *Ideology*, *and the Media*: *The Cyril Burt Scandal*, Transaction Publications, New Brunswick, 1991.
24. Dorfman, D. D. 'The Cyril Burt question: new findings', *Science*, **201**, 1177–1186 (1978).
25. Gould, S. J. *The Mismeasure of Man*, 2nd edn, W. W. Norton and Company, New York, 1981.
26. Knox, R. 'The Harvard fraud case: where does the problem lie?', *Journal of the American Medical Association*, **249**, 1797–1807 (1983).
27. Lock, S. and Wells, F. (eds). *Fraud and Misconduct in Medical Research*, BMJ Publishing Group, London, 1993.
28. Husemeyer R. P. 'Handling scientific fraud. Pearce's editors were not to blame', *British Medical Journal*, **311**, 261–262 (1995).
29. Zuckerman, H. 'Deviant behaviour and social cultural in science', *in* Sagarin, E. (*ed.*), *Social Change*, Sage Publications, Beverly Hill, CA, 1977, pp. 87–138.
30. Medawar, P. *The Strange Case of the Spotted Mice*, Oxford University Press Paperback, Oxford, 1986.
31. George, S. L. 'Perspectives on scientific misconduct and fraud in clinical trials', *Chance*, **10**, 3–5 (1997).
32. Paulos, J. A. *Innumeracy. Mathematical Illiteracy and its Consequences*, Collins Publishers, Toronto, 1988.
33. Andersen, B. *Methodological Errors in Medical Research*, Blackwell scientific Publications, Oxford, 1990.
34. National Academy of Sciences. *Responsible Science*: *Ensuring the Integrity of the Research Process*, National Academy Press, Washington, DC, 1992.
35. Saint James Roberts, I. 'Cheating in science', *New Scientist*, **72**, 466–469 (1976).

36. Lock, S. 'Misconduct in medical research: does it exist in Britain?', *British Medical Journal*, **297**, 1531–1535.
37. Verma, B. L. and Shukla, G. D. 'Scientific misconduct in medical research', *Journal of the Indian Medical Association*, **90**, 222–225 (1993).
38. Anonymous. 'Integrity and misconduct in research', Report of the Commission on Research Integrity, U.S. Department of Health and Human Services, U.S. Government Printing Office, 1996-746-425, 1996.
39. George, S. L. 'The scientific integrity of clinical trials: is there a problem?', *Proceedings of the Biometrics Section of the American Statistical Association*, 34–41 (1995).
40. Stewart, L. A. and Clarke, M. J. on behalf of the Cochrane group on meta-analysis using individual patient data. 'Practical methodology of meta-analyses (overviews) using updated individual patient data', *Statistics in Medicine*, **14**, 2057–2079 (1995).
41. Weiss, R. B. 'Systems of protocol review, quality assurance, and data audit', *Cancer Chemotherapy and Pharmacology*, **42** (Suppl) S88–S92 (1998).
42. Petersdorf, R. G. 'A matter of integrity', *Academic Medicine*, **64**, 119–123 (1989).
43. Rippere, V. 'Handling scientific fraud. Clinical fraud is common', *British Medical Journal*, **311**, 262 (1995).
44. Ward, P. 'Europe takes tentative steps to combat fraud', *Applied Clinical Trials*, **September**, 37–42 (1995).
45. Miers, M. L. 'Current NIH perspectives on misconduct in science', *American Psychology*, **40**, 831–835 (1985).
46. Weiss, R. B., Vogelzang, N. J., Peterson, B. A., Panasci, L. C., Carpenter, J. C., Gavigan, M., Sartell, K., Frei, E. and McIntyre, O. R. 'A successful system of scientific data audits for clinical trials', *Journal of the American Medical Association*, **270**, 459–464 (1995).
47. Sunderland, M., Kuebler, S., Weiss, G. and Coltman, C. 'Compliance with protocol: quality assurance data from the Southwest Oncology Group', *Proceedings of the American Society of Clinical Oncology*, **9**, abstract 229 (1990).
48. Schmidt, J., Gertzen, H., Aschenbrenner, K. M. and Ryholt-Jensen, S. 'Detecting fraud using auditing and biometrical methods', *Applied Clinical Trials*, **May**, 40–50 (1995).
49. Shapiro, M. F. and Charrow, R. P. 'The role of data audits in detecting scientific misconduct: results of the FDA program', *Journal of the American Medical Association*, **261**, 2505–2511 (1989).
50. Shapiro, M. F. 'Data audits in investigational drug trials and their implications for detection of misconduct in science', *In* Lock, S. and Wells, F. (eds), *Fraud and Misconduct in Medical Research*, BMJ Publishing Group, London, 1993.
51. Wells, F. 'Investigator fraud in clinical research', *Applied Clinical Trials*, **October**, 40–50 (1996).
52. Altman, D. G. 'Statistics and ethics in medical research. VIII. Improving the quality of statistics in medical journals', *British Medical Journal*, **282**, 44–47 (1981).
53. Bailar, J., 'Science, statistics, and deception', *Annals of Internal Medicine*, **104**, 259–260 (1986).
54. Huth, E. J. 'Irresponsible authorship and wasteful publication', *Annals of Internal Medicine*, **104**, 257–259 (1986).
55. Chalmers, I. 'Underreporting research is scientific misconduct', *Journal of the American Medical Association*', **263**, 1405–1408 (1990).
56. Nowak, R. 'Problems in clinical trials go far beyond misconduct', *Science*, **264**, 1538–1541 (1994).
57. Chalmers, I., Gray, M. and Sheldon, T. 'Handling scientific fraud. Prospective registration of health care research would help', *British Medical Journal*, **311**, 262 (1995).
58. Bulstrode, C. and Fulford, P. 'Fraudulent and redundant publication', *Journal of Bone and Joint Surgery* **77-B**, 845–846 (1995).
59. Marshall, E. 'Secretiveness found widespread in life sciences', *Science*, **276**, 525 (1997).
60. Vogel, G. 'Long-suppressed study finally sees light of day', *Science*, **276**, 525–526 (1997).
61. Dresser, R. 'Defining scientific misconduct. The relevance of mental state (see comments)', *Journal of the American Medical Association*, **269**, 895–897 (1993).
62. Buyse, M. 'The case for loose inclusion criteria in clinical trials', *Acta Chirurgica Belgica*, **90**, 129–131 (1990).
63. Yusuf, S., Held, P., Teo, K. K. and Toretsky, E. R. 'Selection of patients for randomized controlled trials: implications of wide or narrow eligibility criteria', *Statistics in Medicine*, **9**, 73–86 (1990).

64. Stenning, S. The "Uncertainty Principle": selection of patients for cancer clinical trials', in Williams, C. J. (ed.), *Introducing New Treatments for Cancer*, Wiley, New York, 1992.
65. George, S. L. 'Reducing patient eligibility criteria in cancer clinical trials', *Journal of Clinical Oncology*, **14**, 1364–1370 (1996).
66. Mackintosh, D. R. and Zepp, V. J. 'Detection of negligence, fraud, and other bad faith efforts during field auditing of clinical trial sites', *Drug Information Journal*, **30**, 645–653 (1996).
67. Brundage, M. D., Pater, J. L. and Zee, B. 'Assessing the reliability of two toxicity scales: implication for interpreting toxicity data', *Journal of the National Cancer Institute*, **85**, 1138 (1993).
68. Vantongelen, K., Rotmensz, N. and Van der Schueren, E. 'Quality control of validity of data collected in clinical trials', *European Journal of Cancer and Clinical Oncology*, **25**, 1241–1247 (1989).
69. Pullar, T., Kumar, S., Tindall, H. and Feely, M. 'Time to stop counting the tablets?', *Clinical Pharmacology and Therapeutics*, **46**, 163–168 (1989).
70. Anonymous. *La Fraude dans les Essais Cliniques*, Médicament et Santé STS Edition, Paris, 1991.
71. Scherrer, B. 'L'apport de la biométrie', *in La Fraude dans les Essais Cliniques*, Médicament et Santé, STS Edition, Paris, 1991, pp. 47–58.
72. Yusuf, S., Collins, R. and Peto, R. 'Why do we need some large, simple randomized trials?', *Statistics in Medicine*, **3**, 409–422 (1984).
73. Buyse, M. 'Potential and pitfalls of randomized clinical trials in cancer research', *Cancer Surveys*, **8**, 91–105 (1989).
74. Peto, R. 'Monitoring of cancer patients in clinical trials need not be precise', *in* Symington, T., Williams, A. E. and McVie, J. G. (eds), *Cancer*: *Assessment and Monitoring*, Churchill Livingstone, Edinburgh, 1980.
75. Buyse, M. 'Regulatory vs public health requirements in clinical trials', *Drug Information Journal*, **27**, 977–984 (1993).
76. Decoster, G. and Buyse, M. 'Clinical research after drug approval: what is needed, what is not', *Drug Information Journal*, **33**, 627–634 (1999).
77. Schwarz, R. P. 'Maintaining integrity and credibility in industry-sponsored clinical research', *Controlled Clinical Trials*, **12**, 753–760 (1991).
78. Seachrist, L. 'NIH tightens clinical trials monitoring', *Science*, **264**, 499(1994).
79. Seachrist, L. 'NIH trial monitoring: hit or miss?', *Science*, **264**, 1534–1537 (1994).
80. Neaton, J. D., Bartsch, G. E., Broste, S. K., Cohen, J. D. and Simon, N. M. 'A case of data alteration in the Multiple Risk Factor Intervention Trial (MRFIT)', *Controlled Clinical Trials*, **12**, 731–740 (1991).
81. Knatterud, G. L., Rockhold, F. W., George, S. L., Barton, F. B., Davis, C. E., Fairweather, W. R., Honohar, T., Mowery, R. and O'Neill, R. T. 'Guidelines for quality assurance procedures for multicenter trials: a position paper', *Controlled Clinical Trials*, **19**, 477–493 (1998).
82. Efird, J. T., Yavner, S. B. and Karthigesan, J. 'Detecting health care fraud and abuse: a computer-based statistical approach', *Health Policy Review*, **2**, 1–16 (1996).
83. Enas, G. G., Sanger, T. M. and Huster, W. J. 'Essential efficacy data analysis (with discussion)', *Biopharmaceutical Report*, **2**, 1–12 (1993).
84. Collins, M., Evans, S., Moynihan, J., Piper, D., Thomas, P. and Wells, F. 'Statistical techniques for the investigation of fraud in clinical research', Report of the ABPI Fraud Statistics Working Party, February 1993.
85. Evans, S. 'Statistical aspects of the detection of fraud', *in* Lock, S. and Wells, F. (eds), *Fraud and Misconduct in Medical Research*, BMJ Publishing Group, London, 1993, pp. 61–74.
86. Evans, S. 'Fraud and misconduct in medical science', *in* Armitage, P. and Colton, T. (eds), *Encyclopaedia of Biostatistics*, Wiley, Chichester, 1998, pp. 1583–1588.
87. Haldane, J. B. S. 'The faking of genetic results', *Eureka*, **6**, 21–28 (1948).
88. Mosimann, J. E., Wiseman, C. V. and Edelman, R. E. 'Data fabrication: can people generate random digits?', *Accountability in Research*, **4**, 31–55 (1995).
89. Preece, D. A. 'Distribution of final digits in data', *Statistician*, **30**, 31–60 (1981).
90. Benford, F. 'The law of anomalous numbers', *Proceedings of the American Philosophical Society*, **78**, 551–572 (1938).
91. Raimi, R. A. 'The first digit phenomenon', *American Mathematics Monthly*, **83**, 521–538 (1976).
92. Hill, T. P. 'The significant-digit phenomenon', *American Mathematics Monthly*, **102**, 322–327 (1995).

93. Hill, T. P. 'A statistical derivation of the significant-digit law', *Statistics in Science*, **10**, 354–363 (1996).
94. Hill, T. P. 'The first-digit phenomenon', *Scientific American*, **86**, 358–363 (1998).
95. Nigrini, M. 'A taxpayer compliance application of Benford's law', '*Journal of the American Tax Association*, **18**, 72–91 (1996).
96. Mosimann, J. E. and Ratnaparkhi, M. V. 'Uniform occurrence of digits for folded and mixture distributions on finite intervals', *Communications in Statistics – Simulations*, **25**, 481–506 (1996).
97. DeMets, D. L. and Meinert, C. L. 'Data integrity', *Controlled Clinical Trials*, **12**, 727–730 (1991).
98. Dawson, R. J. and Mac, G. 'How many light bulbs does it take to generate a data set?', *American Statistican*, **50**, 247–249 (1996).
99. Canner, P. L., Huang, Y. B. and Meinert, C. L. 'On the detection of outlier clinics in medical and surgical trials: I. Practical considerations', *Controlled Clinical Trials*, **2**, 231–240 (1981).
100. Bailey, K. R. 'Detecting fabrication of data in a multicenter collaborative animal study', *Controlled Clinical Trials*, **12**, 741–752 (1991).
101. The ESPS2 Group. 'European Stroke Prevention Study 2. Efficacy and safety data', *Journal of Neurological Sciences*, **151**, (Suppl) S1–S77 (1997).
102. Farthing, M. J. G. 'Coping with fraud', *Lancet*, **352** (Suppl IV) II (1998).
103. Fleming, T. R 'Data Monitoring Committees and capturing relevant information of high quality', *Statistics in Medicine*, **12**, 565–570 (1993).
104. Ryan, R. P. 'Handling scientific fraud. Serious allegations are hard to believe', *British Medical Journal*, **311**, 262 (1995).
105. Goldbeck-Wood, S. 'Scientists call for whistleblowers' charter', *British Medical Journal*, **315**, 1251–1254 (1997).
106. Evered, D. and Lazar, P. 'Misconduct in medical research', *Lancet*, **345**, 1161–1162 (1995).
107. Gunsalus, C. K. 'Ethics: sending out the message', *Science*, **276**, 335 (1997).
108. Mishkin, B. 'Responding to scientific misconduct. Due process and prevention', *Journal of the American Medical Association*, **260**, 1932–1936 (1988).
109. Friedman, P. J. 'Research ethics, due process, and common sense', *Journal of the American Medical Association*, **260**, 1937–1938 (1988).
110. Woolf, P. K. 'Science needs vigilance not vigilantes', *Journal of the American Medical Association*, **260**, 1939–1940 (1988).
111. Parrish, D. M. 'Improving the scientific misconduct hearing process', *Journal of the American Medical Association*, **277**, 1315–1319 (1997).
112. Bronowski, J. *Science and Human Values*, Harper and Row, New York, 1965.
113. Medawar, P. *The Limits of Science*, Oxford University Press, Oxford, 1984.
114. James, W. 'Fraud and hoaxes in science', *Nature*, **377**, 474 (1995).
115. Alberts, B. and Shine, K. 'Scientists and the integrity of research', *Science*, **266**, 1660–1166 (1994).

