The integrity of the scientific literature

Walter W. Stewart and Ned Feder

A case of admitted scientific fraud has shed new light on the system that ensures the integrity of the scientific literature. Certain lapses from generally accepted standards of research may be more frequent than is commonly believed.

WE wish to report the results of an unusual investigation. In May 1981, the colleagues of a respected young scientist were shocked to discover that he was a forger of data. Initially, Dr John Darsee confessed to a single forgery, but it was subsequently found by his colleagues and by three investigating committees that he had fabricated much of the data that formed the basis of his more than 100 publications[1-129] over a period of about three years.

Although many of these publications embody data now acknowledged to be fraudulent, we have scrutinized these publications for internal consistency, accuracy and completeness. Our sample consists of the 109 publications by Darsee with the 47 scientists who were his co-authors, all of whom worked at two leading US medical schools. We also studied the reports of the committees set up to investigate the Darsee affair. What is special about the publications in our sample is that, as a by-product of the extensive investigations into Darsee's work, the circumstances surrounding their publication were known in unusual detail.

Our objective was to discover whether the published scientific reports would throw light on the vigilance of referees, of editors of journals and of Darsee's co-authors in meeting the standards conventionally accepted as necessary in the scientific literature. We have found many of the reports to be flawed in ways that could have been recognized by those concerned with their publication.

We emphasize the distinction between the original forgery and the lapses from standard publication practices we have found in our "sample of convenience"[130]. Other studies will be needed to determine whether or not our results are representative of clinical scientists, of biomedical scientists or of scientists in general.

Background

The circumstances of the Darsee affair have been widely reported[131-141]. In 1978 to 1981 inclusive, he was author or co-author of 18 full-length research papers published in major biomedical journals and of about 100 abstracts, book chapters, letters, reviews and short papers in clinical and experimental cardiology[1-129]. In part because he was so prolific, he was highly regarded first at Emory University School of Medicine and then at Harvard Medical School.

Five months after Darsee's confession, in May 1981, to a single act of data fabrication, it became apparent that there had been more than one such episode. Ultimately, three committees were appointed to investigate. Harvard appointed a committee headed by R.S. Ross, which issued its report[131] in January 1982. The NHLBI (National Heart, Lung, and Blood Institute) appointed a panel headed by H.E. Morgan to investigate the circumstances at Harvard, where most of the research had been supported by the National Institutes of Health (NIH); this report[134] was completed in June 1982. Finally, at Emory, an investigating committee headed by N.C. Moran was formed, and issued its report[132,133] in March 1983.

Shortly after the appearance of the report of the Moran committee, and almost two years after the frauds themselves had come to light, we embarked on our study. We were especially careful to discriminate between the direct effects of Darsee's forgeries, which are not the subject of this report, and the defects in the papers of which he was a co-author. We recognize that some may disagree with some of our decisions as to what constitutes a defect, and with our judgements of the degree to which Darsee's co-authors may be held to have been responsible.

Darsee's publications were in experimental and clinical cardiology, a field of research quite different from our own. For this reason, we have no comments to make on the pointedness or importance of the research described in the publications or about the design of the experiments. No doubt we may also have overlooked some errors in the publications because of our lack of first-hand knowledge of the field. At the same time, we believe that our position as outsiders had the advantage that none of those who published with Darsee was known to us personally or professionally. It may also be an advantage that we have been able to carry out our analysis in the same way that other scientists outside the field of cardiology might approach such a task.

Of Darsee's 47 co-authors, 24 worked at Emory University School of Medicine and 23 at Harvard Medical School. Biographical data were found for 44 of the 47 co-authors. Thirty-nine were apparently MDs, one was a PhD and one held both qualifications; three apparently had no advanced degree. At the time of their involvement with Darsee, about half of the co-authors were 30–40 years old and had received their advanced degrees 5–15 years previously. Approximately equal numbers were senior and junior to this group, the latter including technicians, medical students, resident staff at teaching hospitals and junior faculty members. The senior co-authors included professors and department chairmen. (Darsee himself was 33 in 1981.)

Abundance of errors

Many of the published errors and discrepancies were presumably introduced by Darsee. We nonetheless considered all errors, major and minor, to be of interest because they might reflect on the care with which the co-authors had checked the publications that bore their names, and might reflect indirectly on the care with which they had carried out their share of the research.

All but a few of the 18 research papers[6,24-26,50-53,87-96] (Fig. 1) and many of the abstracts contain errors or discrepancies that can be recognized simply by examining them carefully. We emphasize again that many of the errors and discrepancies are minor, for example, a small discrepancy between a numerical value printed in a table and what is supposed to be the same value printed in the text. Some errors,

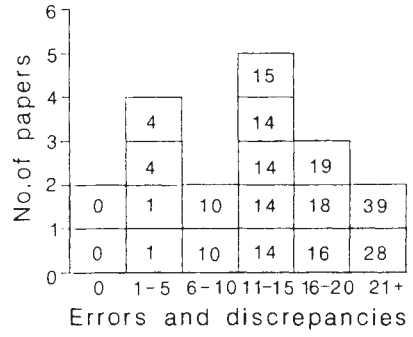

Fig. 1 Errors and discrepancies in published papers. Each box represents a single paper, and the number within the box is the sum of the errors and discrepancies for that paper.

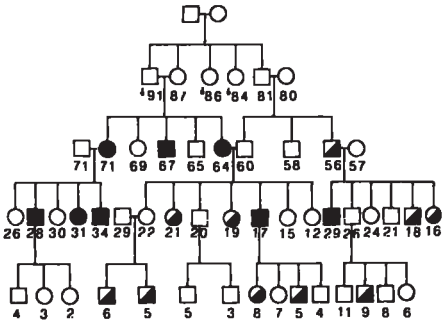

Fig. 2 Pedigree reproduced from ref. 87. The age of each family member at the start of the investigation is given below each symbol. Arrows (second row) denote those who had died before the start of the investigation. Note the 17-year-old male (row 4) with children ages 8,7,5 and 4; at the age of 8 or 9 (probably 8) he impregnated the mother of his oldest child.

however, are so glaring as to offend common sense.

We present a detailed analysis of three sample research papers from which the reader may judge the types of errors we counted. Two papers[87,88] from Emory and one from Harvard[96] are discussed below as examples of our procedure. Two of these[87,96] were eventually retracted wholly[127,142] or in part[143], but one[88] has not been retracted and is still cited as valid[132,133].

The most striking example, in ref. 87, is that in which the authors reported an arresting pedigree depicting a family with a high incidence of an unusual form of heart disease. Inspection of the pedigree (Fig. 2), reveals that a 17-year-old male had 4 children, ages 8,7,5 and 4. (It is unlikely that the father's age is a misprint: it appears in the figure and in two places in the text.) Thus the father was 8 or 9 (probably 8) when he impregnated the mother of his first child and 9 or 10 when he impregnated the mother of his second. This bizarre feature of the pedigree, which perhaps could have raised questions about the validity of the entire paper, was apparently not noticed by co-authors or referees, nor were the following unlikely groupings of ages: his sister, brother and first cousin had their first children at ages 16,15 and 15, and three women in the preceding generation had their last children at ages 41,45 and 52.

The summary of the paper gives the urinary taurine levels of the family members with congestive cardiomyopathy as ranging from 411 to 536 mg taurine per g creatinine, but the text and Table 1 of the same paper give for the same measurement 426 ± 45 mg taurine per g creatinine (mean ± standard deviation). Simple inspection suggests that these two sets of numbers cannot simultaneously be valid. Likewise, the summary of the paper states that the urinary taurine levels of the family members with mitral valve prolapse ranged from 215 to 265 mg taurine per g creatinine, whereas the text and Table 1 give 220 ± 31 mg taurine per g creatinine. These values are central to the thesis of the whole paper. The paper has three numerical discrepancies with ref. 40, and five with refs 19 and 28, abstracts which appeared before the paper and which deal with the same family. All three abstracts of this work indicated that, at the beginning of the study, all 46 members of the kindred were alive, which contradicts Fig. 1 of ref. 87, in which 3 of the 5 oldest are shown as having died before the beginning of the study.

A more subtle example of a major error is illustrated in Fig. 3, taken from a paper[88] also from Emory by six members of our sample. The figure depicts changes in haemodynamic values in five patients at various time intervals after a surgical procedure. Almost all the data in this paper[88] are contained in three figures which cannot be reconciled with each other. The mean and standard deviation for haemodynamic values in the five patients are given in our Fig. 3 (Fig. 1 in the original paper). It is easy to see that many error bars are the same size, and also that mean SVR (systemic vascular resistance) is almost constant at −160 units (the bottom of the graph) in four serial measurements. (Abbreviations are explained in the legend to Table 1.)

These results are especially implausible in the light of data given elsewhere in the paper, where it is shown that SVR and other haemodynamic values for two individual patients vary markedly in the course of time. Moreover, many standard deviations in the figure we reproduce appear on inspection too small to be reconciled with the widely scattered values shown elsewhere in the paper; this is true of SVR at 0.5 h, PVR at 1 h, SWI and SVR at 8 h, and CO at 12 h. Indeed, simple calculations show that it is virtually impossible to reconcile most of the data in the figure we reproduce with the data in the other two figures of the paper.

The introductory section in this paper[88] indicates that the study of five patients was carried out "in four uncomplicated cases of peritoneovenous shunt insertion, and in one subject who developed signs of cardiac failure and pulmonary edema", but the figure legends indicate that two of the five patients developed pulmonary oedema. Specifically, one of the four patients said to be without complications "developed clinical and radiographic signs of pulmonary edema 45 minutes after the shunt was opened" (Fig. 2 legend). The other patient, whose pulmonary oedema was more severe, is described variously as developing this complication "one hour after the shunt was opened" (Fig. 3 legend) and "within a few hours" after the shunt was opened (results section[88]).

Furthermore, the statements that the changes of the mean shown in Fig. 1 after the first half hour were "very small" and that there was "very little change" appear to be contradicted by the actual maximum changes of the mean — a factor of two or more in the change from the value at time zero (corresponding to roughly a 25–60% change in the mean itself) — for three of the five parameters in the figure.

There are obvious inconsistencies between this paper[88] and a preceding abstract[30] describing the same five patients. Eight of ten values for mean or standard deviation given in the two publications are inconsistent (Table 1). This includes most of the numerical data in the abstract. Another comparison (Table 1) shows discrepancies in two out of five standard deviations between one

Table 1 Discrepancies between a paper and two abstracts

	Ref. 88	Ref.30	Refs 20 and 88	Refs 20 and 30
Mean or s.d.	m.±s.d.	m.±s.d.	s.d.	s.d.
Parameter measured:				
CO	2.1±0.8	2.6*±0.8	1.0†	0.9
PCWP	10.0±3.7	12*±2*	4.4	1.8
PVR	−43±33	−64*±24*	30	43‡
SVR	−123±33	−192*±168*	199†	158
SWI	10.6±4.2	11.3±5.8*	4.8†	5.2‡

Ref. 88 is a full-length paper; refs 20 and 30 are abstracts. Discrepancies are indicated between columns 2 and 1 (*), columns 3 and 1 (†), and columns 4 and 2 (‡). Values in column 1 were obtained by measurements of the 0.5 h points in Fig. 1 of ref. 88. The values in column 2 were copied directly from ref. 30, except that for SVR and PVR, the R units[30] have been converted to dyn s cm^{-5} (factor of 80; ref. 144). Columns 3 and 4 contain standard deviations for 5 patients calculated from the means and standard deviations given in ref. 20 for 4 patients and the means given in refs 88 or 30 for 5 patients. The values were derived from calculations based on "population" standard deviations; calculations based on "sample" standard deviations gave similar values and conclusions. When comparing values, allowance was made for rounding of the final digit in refs 20 and 30 and for uncertainties of measurement in ref. 88. Abbreviations: CO, cardiac output; PCWP, pulmonary capillary wedge pressure; PVR, pulmonary vascular resistance; SVR, systemic vascular resistance; SWI, stroke work index. Units are given in the publications, except as noted above.

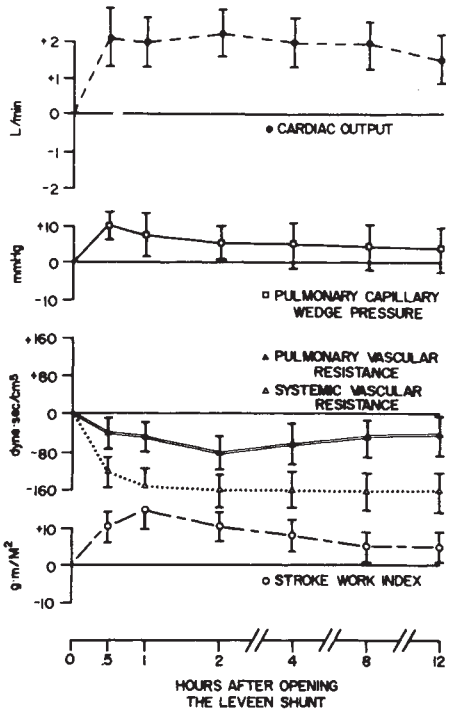

Fig. 3 Reproduction of Fig. 1 from ref. 88. The graph depicts the mean and standard deviation for changes in haemodynamic values of all five patients in the study. Note the remarkable similarity in size of many error bars. Note also the relative constancy of SVR (average value for five patients) at −160 for four serial measurements, despite the variability of SVR (individual values for two patients) in Figs 2 and 3 of ref. 88 (not reproduced here).

abstract[20] and the other[30], and in three of five standard deviations between the abstract[20] and the paper[88]. Thus the two abstracts are inconsistent with each other and with the full-length paper.

In both research papers described above, the major errors are fundamental and call into question the validity of the paper's main conclusions. More typically, however, the errors simply reflect on the care with which the paper had been prepared.

The final example was contributed to a major journal by one of the co-authors about two months after Darsee had admitted to him and another co-author that he had forged data in another study. At the time the paper[96] was contributed, these two co-authors were carrying out, privately, a "detailed review and analysis"[131] of the extent of Darsee's forgery in their laboratory. Thus these two scientists were simultaneously publishing a paper containing the results of their collaboration with Darsee and serving in private as investigators into the extent of his forgery of data. One author stated[138] that he had "total confidence" in this particular paper, which he also said[135] had "passed special scrutiny". The Ross Committee, asked to review Darsee's work at Harvard and prepare a statement of record, affirmed the co-authors' view that this publication was accurate.

The paper[96] describes the recovery of heart muscle injured reversibly by a brief experimental occlusion of a coronary artery. One measure of heart muscle function (systolic shortening) is said in the legend to Fig. 3 not to recover fully "until after 7 days of reperfusion". This description is at least unclear, and we regard it as erroneous, since the figure itself and the data in Table 1 show full recovery at or before 7 days.

We present this as an example of an error of very minor significance, but one that we think suggests that the co-authors were not sufficiently careful in examining their own paper before publication. There are small discrepancies between the summary (p. 7152) and the text (p. 7153) with regard to ATP concentration at both 90 min and 72 h of reperfusion. There is a discrepancy between the error bar for ischaemic subendocardium at 15 min of occlusion in Fig. 1A and the corresponding number in the summary and text.

The paper also contains discrepancies with another paper[95] and with an abstract[105], both published previously, and both describing similar or identical studies on similar groups of dogs. The values of maximum negative dp/dt (all between about − 4,000 and − 4,300 mm Hg s^{-1}) for the dog in Fig. 3 of ref. 96 cannot be reconciled with the values (means from − 2,248 to −2,562; standard deviations from 200 to 316) presumably calculated from this dog plus 6 others and listed in Table 2 of ref. 95. Also, the size of the experimental groups for measurement of ATP and creatinine phosphate at 15 min occlusion and 90 min reperfusion is different in the three publications: 14 or 21 dogs (the text is ambiguous) in ref. 95; 34 or 41 in ref. 96; and 36 in ref. 105.

Despite these differences in the numbers of dogs in the experimental groups, the means and standard errors, wherever given, are the same for all three publications. For example, in refs 95 and 96 (Figs 1 and 2, 15 min occlusion and 90 min reperfusion), there are 16 values and 16 error bars; these are identical in the two papers. Also, in the cases where numerical data are given in the text of the three publications, the numbers after rounding are the same. Such a coincidence of means and standard errors from experimental groups of different sizes is extremely implausible. The paper[96] contains additional errors and discrepancies not described here.

Among the 18 research papers, there were as many as 39 errors and discrepancies in a single research paper, with an average of about 12 per paper (Fig. 1). Of the 22 scientists who were co-authors of a research paper, 19 were co-authors of at least one research paper containing 10 or more errors or discrepancies.

The analysis of errors and discrepancies given above is based entirely on our sample of publications in scientific journals. In contrast, some of the analyses that follow depend on information given in the committee reports.

Retention of data

The reports of the investigating committees showed that in almost all cases most of the co-authors had not retained the experimental or clinical data on which the publications were based. This was true of 7 of the 8 Harvard papers, which involved experiments on dogs, and of 8 of the 10 papers from Emory, 8 of which involved human subjects. [These figures appear mostly attributable to the failure to retain data by Darsee, the chief source of experimental data — Editor, *Nature*.] In the latter case, there were 7 papers for which some of the co-authors had retained neither names nor other identifying information for the human subjects. For the purposes of the present report we have disregarded all papers not involving human subjects except for the cases discussed below.

Of 14 who were co-authors of papers (all from Emory) involving human subjects, at least nine failed to retain the list of subjects (Table 2). One paper[25] is particularly interesting: a report on the incidence of mitral valve prolapse in presumably healthy young men. The paper is considered valid[133] and is heavily cited[145]. The co-authors "maintain that the data in the paper are valid although none of them has a copy of the original research results. They state that they did the study and oversaw the writing of the paper"[133]. Although Darsee was first author of the paper, the Moran report provides no information on his role in the project and, in particular, on his opportunities for altering or forging the data; possibly this in-

Table 2 Frequency of lapses classified by publication and by co-author

Type of lapse	No. of publications	No. of co-authors
1	28	26
2	42	21
3	7	9
4	13	6
5	7	6
6	1	4
7	8	8
8	1	1
9	1	1
10	2	3

Column 1, type of lapse as defined in Table 5. Column 2, number of publications by co-authors in which there were one or more examples of the lapse in column 1. Column 3, number of co-author involved in each type of lapse. Some publications and some co-authors were associated with more than a single type of lapse.

formation is contained in the confidential appendices to the report. Thus the only available support for the validity of this paper is based on a statement attributed by the Moran Committee to the co-authors, given without supporting data or details.

Involvement in research

In considering involvement, we confined our attention to the 18 research papers. We defined "direct involvement" to mean actual participation in the acquisition and interpretation of data, or the making of an intellectual contribution which was essential to the research. We disregarded a person's rank in the laboratory hierarchy in judging whether an author was directly involved.

Our evidence for the extent of an author's involvement came from the reports of the investigating committees. For example, according to a committee report[132], one co-author's role was to encourage Darsee and provide some of the grant support for the research. In another committee report[134] a senior scientist is described as having other duties not allowing him sufficient time for direct participation in the research laboratory and also as being seldom in the laboratory and having no direct supervisory role. In both these cases we considered that the co-authors had not met our criteria for direct involvement, in the sense of direct participation in experimental work. Supervisory functions may be exercised with a greater or lesser degree of remoteness, and we acknowledge that the evidence available is limited and that our judgements on this point are necessarily subjective.

We consider those not directly involved in the research to be "honorary" authors. Honorary authorship was a common occurrence, according to our analysis. Of 18 papers, 13 had at least one author judged by us to be honorary. There was a total of 49 authorships (that is, occurrences of an author's name on one of the 18 research papers); in 16 authorships (33% of the total), involving 6 of the 47 co-authors, the author was judged to have little or no direct involvement (Table 2). Those who were honorary authors at least once also published significantly more papers than those who were never honorary authors (3.33 ± 1.63 versus 0.71 ± 1.42, mean ± standard deviation; $P < 0.01$ by Student's two-tail t-test). Their higher rate of publication compared with co-authors who were not honorary appeared to depend entirely on their honorary papers: the difference disappeared when an author's honorary papers were not counted (0.67 ± 0.82 versus 0.71 ± 1.42). The 13 papers with at least one honorary author had more errors and discrepancies than the 5 without an honorary author (15.1 ± 10.2 versus 5.0 ± 6.6; $P = 0.067$).

Table 3 No. of dogs for which results were shared

	Ref. 53	Ref. 90	Ref. 91	Ref. 92	Ref. 94
Ref. 53	**33**	8	12	0	9
Ref. 90	8	**22**	9	0	11
Ref. 91	12	9	**76**	6	7
Ref. 92	0	0	6	**48**	0
Ref. 94	9	11	7	0	**44**

The boldface figures (on diagonal) give the total numbers of dogs used in each study. The remaining entries give the numbers of dogs for which experimental data were shared between pairs of studies. Although not shown in this table, data for the same 2 dogs were used in 4 studies[53,90,91,94]. Reference 134 was the source of this table, which in turn was based on information provided by one of the senior co-authors at Harvard.

Incomplete statements

In many ways the most striking of our findings was the discovery that certain research papers (7 of 18) embody statements and data in a manner that impedes the reader in forming an accurate reconstruction of the way in which experiments were actually carried out. The circumstances were these.

Five papers[53,90-92,94] reported studies done with dogs on the effects of experimental coronary artery occlusion on the heart and on recovery from the effects of occlusion with or without drugs. In these papers the assignment of dogs to experimental and control groups is described as formally randomized, as with the statement[90], "Twenty-five minutes after occlusion of the LAD, the remaining dogs were randomized to treated or control group on the basis of a random odd or even number generated by a computer program."

During the subsequent investigation by the NHLBI panel, however, the co-authors acknowledged that some control data had been reused[134] — in other words, that the control dogs actually included some historical controls, a use which had not been mentioned in the published papers. (Historical controls are animals that have been used as controls in a previous study[130]. The same data recorded in the earlier study are then used again as control data in the later study. The number of dogs for which data were reused is shown in Table 3.)

One co-author of three of the five papers subsequently stated[146]: "In order to economize animal usage, as well as professional and technical time, control animals which had previously been randomized were sometimes used to provide data for more than a single experiment . . . we saw no reason not to allow him [Darsee] to use his randomized controls in more than a single experiment." Another scientist, the co-author of all five papers, acknowledged the reuse of data, but stated[147] that Darsee had misled him and the other co-author about the number of dogs for which data were reused and about the number of experiments in which the data for any one dog were reused.

Reuse of data was approved by one of these co-authors in at least one paper not involving Darsee (mentioned in ref. 134), and both co-authors have subsequently defended the practice. One co-author, commenting on a study not involving Darsee that was carried out in his laboratory, acknowledged[147] that control data were reused and that the reuse was not described on publication: "Although the use [of some control data from another study] was not specified in the text [of a study not involving Darsee], the use did not bias the results of the study in any way". The other co-author defended the sharing of control results said to be randomized by stating[146] that "the practice of not specifying in the Methods Section that a control dog had been used in more than one study, while certainly not ideal, is common," and also that "the practice of utilizing controls in more than a single experiment is not uncommon and, *although it should be*, it is not always spelled out in published papers" (emphasis added).

In a series of experiments in which the objective is to study the effects of various interventions on animals subjected to a certain procedure, in this case arterial occlusion, the ideal is that described in this group of five papers: all animals would be subjected to arterial occlusion, and afterwards assigned randomly either to the control group or to one or other of the experimental groups.

Although the use of historical controls may have several practical advantages, not least those of economizing in the use of animals, this practice is clearly less satisfactory than the ideal; for example, the experimenter's foreknowledge could systematically affect the outcome of experiments. The fact that Darsee may not have selected historical controls at random, as required by the laboratory's policy described to the NHLBI panel, is irrelevant to our criticism, which is that these papers leave the reader with the impression that control animals were selected by a randomization procedure that most readers, we judge, would consider superior to that actually used.

We emphasize that we do not criticize the use of historical controls, which may or may not have been appropriate, but merely the misleading published description. Thus the use of historical controls is reconcilable with the statement in the

paper quoted above that "after occlusion . . . the remaining dogs" were randomly assigned to different groups only if "remaining dogs" is understood to refer to a number of historical controls.

It would appear from the subsequent statements of the two co-authors that they knew of and authorized the use of historical controls in the laboratory, as well as the failure to describe that use on publication. The described randomization procedure in any study was supposed to apply to all dogs in that study, but in fact did not apply to the historical controls when their data were reused. We conclude that the papers contained an incomplete and misleading description of experimental procedures.

Furthermore, the experiments reported in the five papers included procedures, such as arterial catheterization and administration of certain drugs or dyes, that differ from one study to another (Table 4). As a result, most of the historical controls had been subjected to procedures in an earlier study that were different from those described for the later studies. Thus the published descriptions of these later procedures, which are described explicitly and in detail as applying to all dogs, do not apply in some respects to the historical controls.

One author, writing to the NIH about the sharing of results, stated[146] that "the experimental conditions were similar" for the studies in question. The other referred[147] to studies (not involving Darsee) in which "experimental conditions are the same"; he has also indicated in correspondence with us that the procedures described in the five papers were similar enough in some cases so that the use of historical controls was justified. Differences in experimental conditions are summarized in Table 4.

Some might consider that in most cases shown in Table 4, the data from dogs in one study would be unsuitable as reused data for another study, but the facts enabling readers to form a judgment on this point were not given in any of the five papers. We believe that the descriptions of experimental procedures did not apply to the reused data obtained in previous studies and that these descriptions were therefore misleading.

A final point may be related to reuse of data: one of the five papers states[91] that "all dogs [were subjected to an experimental procedure, which is described]. Ventricular fibrillation or standstill was then induced by a barbiturate overdose, and the hearts were excised rapidly and placed on ice." Yet the co-authors have said (cited in ref. 134) that some of the same dogs were subsequently used for the 4-hour and 6-hour occlusions in other experiments. As the NHLBI panel report politely observes[134], "This would appear to be impossible". We have assumed that the two co-authors in our sample were unaware of this anomaly at the time of publication.

The misleading statements noted above appeared in five papers co-authored by two members of our sample. In addition to these five papers, there were two others (refs 95 and 96; see next section) containing statements that were imprecise or misleading. In total, there were 18 instances of statements made by 6 co-authors whose effect would have been, in our opinion, to mislead the reader.

Unacknowledged republication

We were surprised to discover several instances in which the same data were republished without adequate reference to that circumstance. The most striking example was a paper[95] containing data later published almost in their entirety as a major part of a subsequent paper[96] which appeared in a different journal. Although the second paper referred to the first by means of an index number in the text, there is no explicit statement in the second paper that the new data are a continuation of an experiment already published.

The subject matter of both papers concerns the recovery of heart muscle injured by temporary surgical occlusion of a coronary artery in dogs, described by the authors as a model with "important clinical implications for patients" with certain kinds of heart disease. The first paper states that "Seven dogs were also subjected to 15 min of LAD occlusion followed by 3 days of reperfusion", which is not false but which does not reveal that the same 7 dogs were actually reperfused for 14 days. The references in the second paper to the first would not suggest to those who were not familiar with the first that the research about to be described was a continuation of an experiment already published. In the second paper's opening paragraph, which is reproduced overleaf, the first paper is cited twice, but neither citation of the first paper provides a hint that some of the same data are being republished and that some of the same animals were used.

On neither paper was Darsee the first or last author, and it seems that the co-authors could have known of the statements that were imprecisely misleading. For the journal to which ref. 96 was submitted (*Proceedings of the National Academy of Sciences*), a contributing author customarily by-passes the refereeing system, which might have discovered the unacknowledged republication of data; instead the contributing author assumes "responsibility for the propriety and scientific standards"[148] of the manuscript he submits.

Table 4 Differences between experimental procedures in five papers

	Ref.53	Ref.90	Ref. 92	Ref. 94	Ref. 91
(1) Coronary artery occlusion (time)	2 – 6 h	6 h	1 or 3 h	4 – 6 h	1 min – 2h
(2) Randomization before/after coronary artery occlusion	After	After	Before	After	Before
(3) Time from removal of coronary artery clamp to end	~1 min – 4h	Not removed	5 or 72 h	Not removed or 2h	6 h
(4) Antiarrhythmic agent:	Lido	Lido	Lido	Proc'de	None
(i) Initial bolus inj., mg per kg	1.5	1.5	1.5	3	None
(ii) Further intermittent inj.	Yes	No(?)	No(?)	No(?)	No
(iii) Continuous infusion	No	No	No	Yes	No
(5) Dye administered initially	Fln	MeB,GV	Fln,GV	Fln	Fln,GV
(6) Dye injected terminally	TS,MeB	MeB	MeB	MonB	MonB
(7) Drug or NaCl soln, ml per kg	0(?)	1	1	2	0(?)
(8) Catheter in aorta/left carotid	Carotid	Carotid	Carotid	Aorta	Carotid
(9) A set of multiple procedures, (a) – (e), listed below	No	No	Yes	No	Yes

There are substantial differences in experimental procedures for at least nine of the ten pairs of papers. The procedures in item 9 include: (a) insertion of micromanometer through stab wound in ventricular apex, (b) application of coronary cuff, (c) repeated brief reocclusion of coronary artery, (d) implantation of crystal transducers in myocardium, and (e) delay in examination of heart after death (for flowmeter calibration). Abbreviations: Lido, lidocaine; Proc'de, procainamide; Fln, fluorescein (injection); MeB, methylene blue (injection); GV, gentian violet (topical); TS, thioflavine S; MonB, monastral blue. Information in the table comes from the papers themselves[53,90–92,94]; see also ref. 134, appendix 5.

Opening paragraph of ref. 96:

PREVIOUS studies from this laboratory have shown that brief (15 min) coronary occlusions followed by reperfusion do not cause necrosis but do result in prolonged abnormalities in myocardial biochemistry, function, and ultrastructure (1 – 3). At 72 h after release of a 15-min coronary occlusion, the ATP concentration in reperfused previously ischemic myocardium was significantly reduced (by 22% of normal), the percentage of systolic shortening of reperfused left ventricular segments was reduced (by 42% of normal), and ultrastructural abnormalities were present, despite the absence of necrosis (1, 3). However, it was not known whether these changes ever were reversed to normal after longer periods of reperfusion and, if so, what the time course of the recovery is. Recently there has been renewed interest in the treatment of acute myocardial infarction by means of coronary reperfusion, either by injecting fibrinolytic agents directly into the obstructed coronary artery or by coronary revascularization (4, 5).

This is the opening paragraph, printed here verbatim, of a research paper [96], discussed in the text. The two occurrences of reference 3 in the paragraph above are the only references in this paper[96] to the authors' own previous study[95] (see text).

A common example of republication involves abstracts of research submitted to scientific societies. Of the 88 abstracts, only 47 were published only once; all the rest involve duplicate or triplicate publication. This practice is apparently followed in many fields of biomedical science, where an abstract may be prepared for one meeting of a society and then submitted, perhaps in an amended form, to a second meeting. We nevertheless suspect that the extent of this practice may not be widely appreciated. We discuss four cases below but have not included abstract data in our tables.

We note that there were four pairs of Darsee abstracts in which the texts were nearly identical, but whose titles had apparently been changed. For example, an abstract[77] entitled "Can tetrazolium stains reliably identify myocardial infarction prior to definite histologic necrosis?" was also presented[73] under the title "Early pathologic detection of acute myocardial infarction". Sometimes the alteration consists of the substitution of synonyms[80,72]: "Persistent myocardial abnormalities following brief periods of temporary coronary occlusion not associated with necrosis" also appears as "Prolonged metabolic, functional, and ultrastructural abnormalities following transient myocardial ischemia without infarction".

In each pair the authors are the same, and the text and numerical values are identical through most of the abstract. The publication dates were generally within about a month of one another, as were the scheduled presentation dates. Although it could be argued that it may be necessary to tailor the titles of certain abstracts to make them acceptable for oral presentation before the different societies, we consider that the change in title has the effect of making the pair of abstracts, if listed only by title, appear different to the reader, whereas in fact the contents are almost the same.

There were 5 instances of unacknowledged republication of data. They appeared in publications, two from Emory and seven from Harvard, having a total of 8 co-authors (Table 2).

Lapses from standards

We were surprised to find that there is apparently no generally accepted written code of conduct for scientific research[149-159]. We think that there is nonetheless an unwritten code of which scientists are aware and to which they profess adherence. We have based the analysis that follows on standards that we believe are generally accepted — accepted in the sense that they are generally understood, not necessarily in the sense that they are always complied with.

In Table 5 we attempt to place in ten categories the departures from generally accepted standards that we observed in our sample; other categories would probably be more appropriate for other studies. We found it useful to divide our categories into two groups, A and B, with those of group A explicable simply by carelessness or excessive haste, while those of group B appeared to be more serious in some sense. We do not mean to suggest that certain departures from accepted practice which appear in group B may not be due solely to carelessness and haste; undoubtedly this is so in some cases. Furthermore, we recognize that the distinction between type A and type B is necessarily subjective and that other observers might classify the lapses we have found differently.

We emphasize that the instances on which Fig. 4 is based were supported by evidence we considered compelling, but not necessarily conclusive in each case. We acknowledge that our categories must be subjective.

Frequency of lapses

Of the 47 co-authors in our sample, 31 (66%) were authors of publications containing lapses of type A — a result that is the same whether or not honorary authorships are included. Furthermore, 13 (28%) were involved in type B lapses. Only 12 of the 47 were found not to have been involved in lapses of either type. Most of the 12 were authors of abstracts only; just 2 of the 12 were co-authors of a research paper.

Some scientists apparently do not regard abstracts as publications, even though they are archived and cited and are sometimes used to establish priority[160]. If the 88 abstracts and 3 chapters are excluded from consideration, the frequency of type A and B lapses is about the same or higher. Of the 22 scientists who were co-authors of the 18 research papers, 21 (95%) were involved with papers containing lapses of type A, and 6 (27%) were involved in lapses of type B.

Discussion

It is obvious that the lapses from generally accepted standards of research described here have at least the potential for interfering with the accuracy of the scientific literature or for harming the scientific enterprise in some other way. We also think that the instances described above represent departures from the generally accepted standards of scientific research. We recognize that the decision as to what constitutes a departure from generally accepted standards is necessarily subjective. Presumably some scientists have views different from those expressed here. We regard it as a matter of the greatest importance that these differences be discussed publicly.

Table 5 Lapses from generally accepted standards

Type A

(1) Presence of errors (usually in numerical values).
(2) Inconsistency with a research group's previously published data.
(3) Failure to retain names or identifying numbers for human subjects included in a published paper.
(4) Honorary authorship (see text).

Type B

(5) Statements that are misleading.
(6) Misleading citations of a previous paper containing data being republished, so that the true relationship between the two papers is obscured.
(7) Publication of very similar abstracts under very different titles.
(8) Failure to utilise unique knowledge that makes possible the recognition of certain serious errors in published work — errors not discoverable by anyone else.
(9) Failure to acknowledge the source of a substantial amount of research data received from someone else.
(10) Failure to take appropriate action after receipt of a complaint, later shown to be well founded, that a colleague may be involved in questionable data collection.

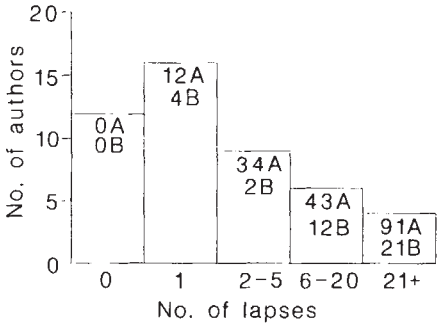

Fig. 4 Authors and their lapses. The histogram gives the number of authors in our sample having lapses within the specified range. The numbers give the total of type A and type B lapses for all authors in that range.

The abundance of errors that we observed in our sample obviously interferes with the accuracy of the scientific literature. Every scientist has, at a minimum, an obligation to ensure that what is published under his name is accurate.

There is a notable absence of explicit published standards for the retention of data. We believe that it would be generally accepted that data should remain accessible for as long as questions and criticisms are likely to be raised about what has been published. For most of the papers in our sample, data were not available. The data we give in Table 4 use only those cases where data involving human subjects had not been retained. We suggest that a reasonable minimum requirement is that each author of a paper involving human subjects should retain at least a list of the subjects' names or the hospital's identification numbers of patients. This information would permit access to hospital records and thus in many cases allow reconstruction of data in the paper.

The reader may ask: What harm is done by honorary authorship? Indeed some of our colleagues have argued that the custom of routinely placing the name of a senior scientist, usually the head of the laboratory, on a paper — regardless of his contribution — is widely followed and does no harm. We disagree, as have others[159,161-166]: honorary authorships falsify the assignment of responsibility for published research and increase the likelihood that inaccurate data will be published. The honorary author is in a poor position to judge the validity of the work, yet he often lends a prestige that may lull other co-authors, the reviewers or the readers into uncritical and inappropriate acceptance.

Similarly, the harm done by unacknowledged republication of data may not be immediately obvious. But it disguises the true relationship between published papers and yields no benefits other than for the authors.

In some ways the most disturbing of our findings is the discovery that certain of the publications (7 out of 18) embody statements and data in a manner that impedes the reader in forming an accurate reconstruction of the way in which experiments were carried out.

Representativeness

The evidence presented here indicates a surprisingly high frequency of lapses from accepted standards. One possible explanation is the association of the co-authors of the publications in our sample with Darsee: he may have induced them to behave in ways not typical of them. A second possibility is that the co-authors are not representative of their field. A third possibility is that lapses from accepted standards may be unusually common in the specialty in which these scientists worked, in clinical research generally or in research carried out at medical schools or by physicians. (The co-authors of the publications in our sample were mainly physicians, and the research was carried out exclusively in medical schools.) A fourth and disturbing possibility is that lapses from accepted standards may be more common among biomedical scientists than is currently appreciated.

Other studies will be needed to determine whether our results are representative of other groups of scientists.

Causes

Extraordinary emphasis is commonly placed on the sheer quantity of publications[158,165-179]. The director of one of the world's leading research institutions (but not an institution with which Darsee was ever affiliated) sent a signed, official memorandum to junior colleagues proposing incentives for publication: "There is no demand that these be literary masterpieces in first line journals; journeyman works for publication in second, third or fourth line archival publications will be quite satisfactory. Upon proper completion and submission of [two] manuscripts, [a technician]'s appointment will be extended to April 1, 1984. During that time it is expected that an additional manuscript on [subject] will be completed and submitted. If so, the period of employment will be extended an additional three months and again an additional manuscript on [subject] is an anticipated result of the extended employment." On another occasion a director of research said in all seriousness, "No research that results in a single paper is worthwhile."

Those promoting Darsee or recommending his promotions were well aware that his publications were unusually numerous[146]. But apparently none of these supporters, who were also his co-authors, made the effort that would have shown that many of the papers had serious flaws. How often does this occur?

Other causes of lapses from accepted standards may be the extreme competition for research grants, large research teams in which the leader is not directly involved in the research and the absence of an accepted code of conduct for scientific research.

In this paper, we have deliberately avoided a discussion of the cases of acknowledged fraud which have come to light in recent years, concentrating instead on defects of the scientific literature generally regarded in the scientific community as regrettable but less serious. Outright fraud, which might conveniently be called a lapse of type C in our nomenclature, is presumably rare. May not the scientific community be over-sanguine in attaching less importance to lapses such as we have found, which are generally unremarked but which debase the scientific literature?

Testing the system

The results of our study naturally raise the question: what fraction of papers in the biomedical sciences are not supported by primary data at the time of publication? The question cannot be answered with the papers in our sample, since most are known to be partial or complete forgeries and thus are atypical.

We suggest a study to answer the question: close examination of a random sample of published papers. Probability of selection could be uniform or could be weighted by a variable such as number of authors or cost of research. Papers would be examined by either of two methods: careful examination of the paper itself for errors and discrepancies by appropriate experts following an established procedure ("external audit") or examination of the primary data on which the paper is based, which would of course require access to laboratory records ("internal audit"). The advantages of the external audit are its non-invasive character and presumably lower cost; the advantage of the internal audit is its completeness.

What would this cost? This depends on the precision required and on the actual fraction of defective papers. We assume for the sake of illustration that an external audit could be done for about 3% of the cost of the research on which a paper is based, and an internal audit for about 10%. If the actual fraction of papers that could not be substantiated were as high as 15% and if the answer were required with a standard error of 7%, about 25 papers would have to be examined, which in the case of internal audits would altogether cost only about 2.5 times as much as one average paper. If, on the other hand, auditing 100 papers failed to disclose a single paper that could not be substantiated, this would show with $P = 0.05$ that the actual fraction of papers unsupported by primary data at the time of publication is less than 3%.

In addition to the financial costs of examining the practices of scientists, there are other costs which may be more serious

in the long run. Examination of scientific practices could cause unwarranted harm to individual scientists. Systematic examination of scientific practices might even weaken the fabric of trust that is essential to the functioning of science. Of all human endeavours, science is one of the most successful — prodigious in benefits, low in cost. But science, vulnerable to abuse from within by its practitioners, is perhaps even more vulnerable to harm by regulation, and at some point the cost of further regulation will outweigh the benefits.

Scientists have to an unusual degree been entrusted with the regulation of their own professional activities. Self-regulation is a privilege that must be exercised vigorously and wisely, or it may be lost. □

Walter Stewart and Ned Feder are in the National Institutes of Health, Bethesda, Maryland 20892, USA.

1. *Notre Dame Sci. Q.* **7**, no.2, 14-17 (1969).
2. *Notre Dame Sci. Q.* **8**, no.1, 9-11 (1969).
3. *Circulation* **56**, III–218 (1977).
4. *South. med. J.* **70**, 1249 (1977).
5. *Archs intern. Med.* **137**, 1639 (1977).
6. *Ann. intern. Med.* **89**, 867-870 (1978).
7. *Clin. Res.* **26**, 3A (1978).
8. *Clin. Res.* **26**, 226A (1978).
9. *Clin. Res.* **26**, 226A (1978).
10. *Chest* **74**, 334-335 (1978).
11. *Chest* **74**, 334 (1978).
12. *Circulation* **58**, II-237 (1978).
13. *Circulation* **58**, II-235 (1978).
14. *Circulation* **58**, II-229 (1978).
15. *Circulation* **58**, II-174 (1978).
16. *Circulation* **58**, II-42 (1978).
17. *Circulation* **58**, II-23 (1978).
18. *Clin. Res.* **26**, 773A (1978).
19. *Clin. Res.* **26**, 762A (1978).
20. *Clin. Res.* **26**, 759A (1978).
21. *Clin. Res.* **26**, 745A (1978).
22. *New Engl. J. Med.* **298**, 221-222 (1978).
23. *Am. J. Med.* **64**, 1089 (1978).
24. *Circulation* **59**, 492-497 (1979).
25. *Circulation* **59**, 619-622 (1979).
26. *New Engl. J. Med.* **300**, 877-882 (1979).
27. *Am. J. Cardiol.* **43**, 367 (1979).
28. *Clin. Res.* **27**, 274A (1979).
29. *Clin. Res.* **27**, 265A (1979).
30. *Clin. Res.* **27**, 264A (1979).
31. *Clin. Res.* **27**, 230A (1979).
32. *Clin. Res.* **27**, 161A (1979).
33. Abstract of report presented in annual Lloyd Hyde Medical House Staff and Fellows Research Program, Emory University School of Medicine (1979).
34. *Clin. Res.* **27**, 573A (1979).
35. *Clin. Res.* **27**, 561A (1979).
36. *Clin. Res.* **27**, 561A (1979).
37. *Chest* **76**, 371 (1979).
38. *Chest* **76**, 350 (1979).
39. *PACE* **2**, A43 (1979).
40. *Circulation* **60**, II-157 (1979).
41. *Circulation* **60**, II-149 (1979).
42. *Circulation* **60**, II-124 (1979)
43. *Circulation* **60**, II-90 (1979).
44. *Circulation* **60**, II-41 (1979).
45. *Circulation* **60**, II-38 (1979).
46. *Clin. Res.* **27**, 726A (1979).
47. *New Engl. J. Med.* **301**, 443 (1979).
48. *South. med. J.* **72**, 174-180 (1979).
49. *Circulation* **60**, II-252 (1979).
50. *Ann. intern. Med.* **92**, 735-741 (1980).
51. *Am. J. Cardiol.* **46**, 607-612 (1980).
52. *Am. J. Cardiol.* **46**, 613-618 (1980).
53. *Am. J. Cardiol.* **46**, 800-806 (1980).
54. *Am. J. Cardiol.* **45**, 492 (1980).
55. *Am. J. Cardiol.* **45**, 476 (1980).
56. *Am. J. Cardiol.* **45**, 395 (1980).
57. *Clin. Res.* **28**, 257A (1980).
58. *Clin. Res.* **28**, 163A (1980).
59. *Clin. Res.* **28**, 163A (1980).
60. *Clin. Res.* **28**, 653A (1980).
61. *Chest* **78**, 513 (1980).
62. *Chest* **78**, 513 (1980).
63. *Circulation* **62**, III-318 (1980).
64. *Circulation* **62**, III-317 (1980).
65. *Circulation* **62**, III-177 (1980).
66. *Am. J. Cardiol.* **45**, 394 (1980).
67. *Fedn Proc.* **39**, 1112 (1980).
68. *Fedn Proc.* **39**, 398 (1980).
69. *Clin. Res.* **28**, 538A (1980).
70. *Clin. Res.* **28**, 163A (1980).
71. *Clin. Res.* **28**, 174A (1980).
72. *Clin. Res.* **28**, 615A (1980).
73. *Clin. Res.* **28**, 614A (1980).
74. *Clin. Res.* **28**, 610A (1980).
75. *Clin. Res.* **28**, 609A (1980).
76. *Circulation* **62**, III-314 (1980).
77. *Circulation* **62**, III-246 (1980).
78. *Circulation* **62**, III-239 (1980).
79. *Circulation* **62**, III-196 (1980).
80. *Circulation* **62**, III-80 (1980).
81. *Circulation* **62**, III-76 (1980).
82. In *Clinical Methods: The History, Physical and Laboratory Examinations* (eds Walker, H. K., Hall, W. D. & Hurst, J. W.) 1122-1161 (Butterworth, Boston, 1980).
83. In *The Heart, Update III* (ed. Hurst, J. W.) 1-8 (McGraw-Hill, New York, 1980).
84. In *Heart Disease: A Textbook of Cardiovascular Medicine*, Vol. 2 (ed. Braunwald, E.) 1517-1582 (Saunders, Philadelphia, 1980).
85. *Primary Cardiol.* **6**, 101-109 (1980).
86. *Primary Cardiol.* **6**, 112-118, 122 (1980).
87. *New Engl. J. Med.* **304**, 129-135 (1981).
88. *Ann. Surg.* **194**, 189-192 (1981).
89. *Circulation Res.* **49**, 1017-1028 (1981).
90. *Circulation* **63**, 29-35 (1981).
91. *Am. J. Physiol.* **240**, H399-H407 (1981).
92. *J. clin. Invest.* **68**, 225-239 (1981).
93. *Archs Path. Lab. Med.* **105**, 403-406 (1981).
94. *Am. J. Cardiol.* **48**, 702-710 (1981).
95. *Am. J. Physiol.* **241**, H591-H599 (1981).
96. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7152-7156 (1981).
97. *Clin. Res.* **29**, 444A (1981).
98. *Clin. Res.* **29**, 428A (1981).
99. *Chest* **80**, 388 (1981).
100. *Circulation* **64**, IV-280 (1981).
101. *Circulation* **64**, IV-210 (1981).
102. *Circulation* **64**, IV-69 (1981).
103. *Circulation* **64**, IV-68 (1981).
104. *Fedn Proc.* **40**, 485 (1981).
105. *Clin. Res.* **29**, 562A (1981).
106. *Clin. Res.* **29**, 270A (1981).
107. *Clin. Res.* **29**, 270A (1981).
108. *Clin. Res.* **29**, 184A (1981).
109. *Clin. Res.* **29**, 184A (1981).
110. *Clin. Res.* **29**, 184A (1981).
111. *Z. Kardiologie* **70**, 317 (1981).
112. *J. nucl. Med.* **22**, P53 (1981).
113. *Chest* **80**, 385-386 (1981).
114. *Circulation* **64**, IV-280 (1981).
115. *Circulation* **64**, IV-236 (1981).
116. *Circulation* **64**, IV-222 (1981).
117. *Circulation* **64**, IV-178 (1981).
118. *Circulation* **64**, IV-122 (1981).
119. *Circulation* **64**, IV-99 (1981).
120. In *The Heart, Update IV* (ed. Hurst, J. W.) 219-286 (McGraw-Hill, New York, 1981).
121. *Am. Heart J.* **102**, 807-808 (1981).
122. *Am. Heart J.* **101**, 124-126 (1981).
123. *J. cardiovasc. Med.* **6**, 705-715 (1981).
124. *Acta med. scand., Suppl.* **651**, 123-132 (1981).
125. *J. cardiovasc. Med.* **6**, 1173-1177, 1181 (1981).
126. *Am. J. Cardiol.* **49**, 1023 (1982).
127. *New Engl. J. Med.* **308**, 1419 (1983) (retraction).
128. *Ann. intern. Med.* **99**, 275-276 (1983) (retraction).
129. *Circulation Res.* **53**, 837 (1983) (retraction).
130. Last, J. M.(ed.) *A Dictionary of Epidemiology*, 23, 94 (Oxford University Press, New York, 1983).
131. Ross, R. S. *et al.* Report of an ad hoc advisory committee to the dean of the Harvard Medical School on dishonesty in scientific research, *Harvard University Gazette* **77**, 1, 11-12 (29 January 1982).
132. Hall, E. C. *et al.*, Report of ad hoc committee to evaluate research of Dr. John R. Darsee at Emory University; [the "Moran report"]; 1 – 32 (Emory University, 28 March 1983); cf. *Minerva* **23**, 276 – 304 (1985).
133. Moran, N. C., Addendum to report on research of Dr John R. Darsee at Emory University School of Medicine, 1 – 3 (Emory University, 5 May 1983).
134. Morgan, H. E., Klocke, F. J., Shepherd, J. T., & Wildenthal, C. K., Report of the National Heart, Lung, and Blood Institute special panel to review alleged misconduct at Brigham and Women's Hospital/Harvard Medical School, Boston, Massachusetts, 1 – 36 and 7 Appendices (29 June 1982).
135. Knox, R. *J. Am. med. Ass.* **249**, 1797-1807 (1983).
136. Knox, R. A. *J. Am. med. Ass.* **249**, 2867-2876 (1983).
137. Knox, R. A. *Boston Globe*, 41,46 (23 May 1983).
138. Broad, W. J. *Science* **215**, 478-482 (1982).
139. Broad, W. J. *Science* **215**, 874-876 (1982).
140. Culliton, B. J. *Science* **220**, 31-35 (1983).
141. Culliton, B. J. *Science* **220**, 936 (1983).
142. Heymsfield, S. B. & Glenn, J. F. *New Engl. J. Med.* **308**, 1400 (1983).
143. [Braunwald, E.] *Proc. natn. Acad. Sci. U.S.A.* **79**, 6390 (1982).
144. Milnor, W.R. *Hemodynamics*, 17 (Williams and Wilkins, Baltimore, 1982).
145. *Science Citation Index, Source Index* (Institute for Scientific Information, Philadelphia, 1979-86).
146. Braunwald, E, "Comments concerning NHLBI panel's report and range of possible NIH actions", Memorandum to NIH staff members, 1-36 and 6 Appendices (21 September 1982).
147. Kloner, R.A., Memorandum to W. F. Raub, NIH, 1-21 (14 September 1982).
148. "Information To Contributors", *Proc. natn. Acad. Sci. U.S.A.* **78**, ix (January 1981).
149. Krevans, J. R. *et al. J. med. Educ.* **57**, 895-902 (1982).
150. Pigman, W. & Carmichael, E. B. *Science* **111**, 643-647 (1950).
151. Clapp, J. *Professional Ethics and Insignia*, 1-851 (Scarecrow Press, Metuchen, New Jersey, 1974).
152. "Yale Policy Statement on Collaborative Research", *Yale Weekly Bull. Calendar* **11**, 9 (13-20 September 1982).
153. Hastings Center Staff (eds) *The Hastings Center's Bibliography of Ethics, Biomedicine, and Professional Responsibility*, 10-11 (University Publications of America, Frederick, Maryland, 1984).
154. Danforth, W. H. *et al. Report of the Association of American Universities Committee on the Integrity of Research*, 1-6 (Association of American Universities, Washington, DC, [April 1983].
155. *The Chemist's Creed*, American Chemical Society (ACS, Washington, DC, 14 September 1965).
156. *Honor in Science*, 1-38 (Sigma Xi, New Haven, Connecticut, 1984).
157. Chalk, R., Frankel, M.S. & Chafer, S.B. *Professional Ethics in the Scientific and Engineering Societies*, 1- 224 (American Association for the Advancement of Science, Washington, DC, 1980).
158. Huth, E. J. *Ann. intern. Med.* **104**, 266-267 (1986).
159. Huth, E. J. *Ann. intern. Med.* **104**, 269-274 (1986).
160. Stossel, T.P. *New Engl. J. Med.* **313**, 123-126 (1985).
161. Cannon, W. B. *The Way of an Investigator: a Scientist's Experience in Medical Research*, 93 (Norton, New York, 1945).
162. [Huth, E. J.] *Ann. intern. Med.* **97**, 613-614 (1982).
163. Relman, A. S. *New Engl. J. Med.* **308**, 1415-1417 (1983).
164. Croll, R. P. *Perspect. Biol. Med.* **27**, 401-407 (1984).
165. Majerus, P. W. *J. clin. Invest.* **70**, 213-217 (1982).
166. Huth, E. J. *Ann. intern. Med.* **104**, 257-259 (1986).
167. Broad, W. & Wade, N. *Betrayers of the Truth* (Simon & Schuster, New York, 1982).
168. [Huth, E. J.] *Ann. intern. Med.* **99**, 266-267 (1983).
169. Burman, K. D. *Ann. intern. Med.* **97**, 602-605 (1982).
170. Gjerde, C. L. & Colombo, S. E. *J. med. Educ.* **57**, 157-162 (1982).
171. Angell, M. *Science* **219**, 1417-1418 (1983).
172. Mahoney, M. J. *Scientist as Subject: the Psychological Imperative* 79-82 (Ballinger, Cambridge, Massachusetts, 1976).
173. Jackson, C. I. & Prados, J. W. *Am. Scient.* **71**, 462-464 (1983).
174. Gratzer, W. *Nature* **302**, 774-775 (1983).
175. Stetten, D., Jr *Science* **226**, 1374, 1376 (1984).
176. Petersdorf, R. G. *Ann. intern. Med.* **104**, 252-254 (1986).
177. Woolf, P. K. *Ann. intern. Med.* **104**, 254-256 (1986).
178. Bailar, J. C., III. *Ann. intern. Med.* **104**, 259-260 (1986).
179. Angell, M. *Ann. intern. Med.* **104**, 261-262 (1986).

● Some editorial changes have been made in this manuscript without the consent of the authors. A reply from Braunwald follows on page 215.