

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/20986>

Please be advised that this information was generated on 2017-12-05 and may be subject to change.

- 40 Reuler JB. Hypothermia: pathophysiology, clinical settings, and management. *Ann Intern Med* 1978; 89: 519-27.
- 41 Jolly BT, Ghezzi KT. Accidental hypothermia. *Emerg Med Clin North Am* 1992; 10: 311-27.
- 42 Lloyd ELI, Conliffe NA, Orgel H, Walker PN. Accidental hypothermia: an apparatus for central re-warming as a first aid measure. *Scot Med J* 1972; 17: 83-91.
- 43 Shields CP, Sixsmith DM. Treatment of moderate-to-severe hypothermia in an urban setting. *Ann Emerg Med* 1990; 19: 1093-97.
- 44 Lash RF, Burdette JA, Ozdil T. Accidental profound hypothermia and barbiturate intoxication. *JAMA* 1967; 201: 269-70.
- 45 Grossheim RL. Hypothermia and frostbite treated with peritoneal dialysis. *Alaska Med* 1973; 15: 53-55.
- 46 Linton AL, Ledingham IMcA. Severe hypothermia with barbiturate intoxication. *Lancet* 1966; i: 24-26.
- 47 Lee HA, Ames AC. Haemodialysis in severe barbiturate poisoning. *BMJ* 1965; 1: 1217-19.
- 48 Gregory JS, Bergstein JM, Aprahamian C, Wittmann DH, Quebbeman EJ. Comparison of three methods of rewarming from hypothermia: advantages of extracorporeal blood warming. *J Trauma* 1991; 31: 1247-52.
- 49 Gentilello LM, Cobean RA, Offner PJ, Soderberg RW, Jurkovich GJ. Continuous arteriovenous rewarming: rapid reversal of hypothermia in critically ill patients. *J Trauma* 1992; 32: 316-27.
- 50 Kugelberg J, Schüller H, Berg B, Kallum B. Treatment of accidental hypothermia. *Scand J Thor Cardiovasc Surg* 1967; 1: 142-46.
- 51 Davies DM, Millar EJ, Miller IA. Accidental hypothermia treated by extracorporeal blood-warming. *Lancet* 1967; i: 1036-37.
- 52 Fell RH, Gunning AJ, Bardhan KD, Triger DR. Severe hypothermia as a result of barbiturate overdose complicated by cardiac arrest. *Lancet* 1968; i: 392-94.
- 53 Splittgerber FH, Talbert JG, Sweezer WP, Wilson RF. Partial cardiopulmonary bypass for core rewarming in profound accidental hypothermia. *Am Surg* 1986; 52: 407-12.
- 54 Letsou GV, Kopf GS, Eleftheriades JA, Carter JE, Baldwin JC, Hammond GL. Is cardiopulmonary bypass effective for treatment of hypothermic arrest due to drowning or exposure? *Arch Surg* 1992; 127: 525-28.
- 55 Hill JG, Bruhn PS, Cohen SE, et al. Emergent applications of cardiopulmonary support: a multiinstitutional experience. *Ann Thorac Surg* 1992; 54: 699-704.
- 56 von Segesser LK, Garcia E, Turina M. Perfusion without systemic heparinization for rewarming in accidental hypothermia. *Ann Thorac Surg* 1991; 52: 560-61.
- 57 Chinard FP. Hypothermia treatment needs controlled studies. *Ann Intern Med* 1979; 90: 990-91.
- 58 Danzl DF, Hedges JR, Pozos RS. The hypothermia study group. Hypothermia outcome score: development and implications. *Crit Care Med* 1989; 17: 227-31.

Clinical practice

Randomised controlled trial of routine individual feedback to improve rationality and reduce numbers of test requests

R A G Winkens, P Pop, A M A Bugter-Maessen, R P T M Grol, A D M Kester, G H M I Beusmans, J A Knottnerus

Summary

Feedback can be described as a way to provide information on doctors' performance to enable changes in future behaviour. Feedback is used with the aim of changing test-ordering behaviour. It can lead to reductions in test usage and cost savings. It is not sufficiently clear, however, whether feedback leads to more appropriate test use. Since 1985, the Diagnostic Coordinating Center Maastricht has been giving feedback on diagnostic tests as a routine health care activity to all family doctors in its region. Both quantity and quality of requests are discussed. In a randomised, controlled trial over 2.5 years, discussion of tests not included previously was added to the existing routine feedback. One group of family doctors (n=39) received feedback on test-group A (electrocardiography, endoscopy, cervical smears, and allergy tests), the other (n=40) on test-group B (radiographic and ultrasonographic tests). Thus, each group of doctors acted as a control group for the other. Changes in volume and rationality of requests were analysed. The number of requests decreased during the trial (p=0.036). Request numbers decreased

particularly for test-group A (p=0.04). The proportion of requests that were non-rational decreased more in the intervention than in the control groups (p=0.009). Rationality improved predominantly for test-group B (p=0.043). Thus, routine feedback can change the quantity and quality of requests.

Lancet 1995; 345: 498-502

Introduction

One method of influencing doctors' test-ordering behaviour is feedback. The effects of feedback have been extensively studied, with contradictory findings.¹⁻¹⁰ Most studies were of limited duration. One investigation followed the effects after feedback was stopped and found that they diminished soon afterwards.⁵ In many studies the effects were expressed in terms of a reduction in the number of tests ordered or a saving in expenses for diagnostic testing.⁴⁻¹⁰

Apart from such quantitative results there is also an effect on quality—the rationality or appropriateness of requests. Until now, effects of feedback on the quality of test use have been studied only incidentally. Schectman et al¹¹ focused on the compliance of test-ordering behaviour with guidelines for the use of thyroid function testing. Given the limited evidence so far, more research on the effects of feedback on the rationality of test-ordering behaviour is needed. Also the coherence between changes in volume and rationality of diagnostic testing should be considered. The goal of an intervention should be to improve quality of care and patient outcome; changes in the rationality and volume of requests can be a step in this direction.

Diagnostic Coordinating Center Maastricht (R A G Winkens PhD, Prof P Pop PhD, A M A Bugter-Maessen MD), **Center for Research on Quality in Family Practice, Universities of Nijmegen and Maastricht** (Prof R P T M Grol PhD); and **Departments of Methodology and Statistics** (A D M Kester PhD) and **Family Practice**, (G H M I Beusmans PhD, Prof J A Knottnerus PhD), **University of Limburg, Maastricht, Netherlands**

Correspondence to: Dr R A G Winkens, Diagnostic Coordinating Centre Maastricht, P O Box 5800, 6202 AZ Maastricht, Netherlands

	Group 1	Group 2
Number of doctors	39	40
Mean (SE) experience (years)	15 (7)	15 (7)
Male/female	35/4	36/4
Practice setting		
Single handed	21	23
Group practice	11	10
Health centre	7	7
Urbanisation		
City	28	25
Rural area	11	15

Table 1: Characteristics of practices and doctors at randomisation

At the Diagnostic Coordinating Center Maastricht (DCC), feedback on test ordering is provided routinely.^{12,11} The main aim is to improve the quality of test-ordering behaviour. Also, feedback is a way to reduce the number of unnecessary requests.¹³ To that end, at set times the rationality and volume of large numbers of requests from individual family doctors are discussed. Surveys showed that the introduction of the feedback was followed by a pronounced reduction in the number of requests in subsequent years. After 2 years, the total number of requests was reduced by 24%. A reduction, for some tests up to 95%, was particularly seen for tests discussed in the feedback and was specific for the Maastricht region.¹² A causal relation between our feedback and these changes has not yet been proven, however.

The effects of such thorough and personal feedback, provided routinely for many years (and not as part of a specifically organised, probably temporary, study setting), have not been studied in a randomised controlled trial. In such a trial we investigated whether routine individual feedback causes a reduction in the number of tests ordered and an increase in the rationality of test-ordering behaviour.

Methods

Since 1985, feedback on diagnostic actions has been given twice a year to the 85 family doctors affiliated to the DCC. These doctors provide care for about 187 000 patients. Every individual

family doctor receives a feedback report with critical comments on his or her requests. The feedback is provided by a respected expert (internal medicine specialist) and is based on an analysis of request forms completed by the family doctor. As a result, the feedback concerns real cases from daily practice and covers a wide variety of diagnostic tests. To help the doctor's recall, patients are mentioned by name and date of birth. The feedback reports discuss both the volume and the rationality of requests submitted in the previous month. Since request forms contain clinical data on the patient (history, physical signs and findings, suspected diagnosis) it is possible to assess the rationality of the tests ordered. Rationality is determined on the basis of accepted regional guidelines and standards of the Dutch College of General Practitioners (NHG).¹⁴ Comments about inappropriate requests are made, and recommendations and alternatives are offered for more rational diagnostic testing. The details and contents of the feedback have been described elsewhere.¹³

A randomised controlled trial was carried out from October, 1989, until May, 1992, among all family doctors affiliated to the DCC. During this period, routine feedback was provided five times (October, 1989; April and November, 1990; May and December, 1991); at each time requests from the preceding month were discussed. Before the trial, feedback concentrated on various blood, urine, and faeces tests. Other tests, such as radiography and electrocardiography, were discussed in less than 5% of all feedback comments and histology not at all. Feedback on these previously undiscussed tests was added to the routine feedback for the study.

The selected tests were divided into two test-groups. Test-group A consisted of cervical smear, electrocardiography, endoscopy, and allergy tests (Phadiatop, IgE, radioallergosorbent test). Test-group B consisted of radiographic and ultrasonographic tests (radiography of chest, cervical spine, thoracic spine, lumbar spine, pelvis, knees, ankles, and sinuses; ultrasonography of kidneys and liver/biliary tract). The selected tests were not split up randomly since they were not always independent from each other. Feedback on tests in one group should not affect requests for tests in the other group.

Because of retirement, 6 family doctors were not included in the trial. The remaining 79 were randomly assigned to receive feedback on one of the two test groups while serving as a control for the other test group. The first group of doctors received feedback on test-group A; the others received feedback on test-group B. After randomisation, the doctor groups were similar in terms of practice setting and degree of urbanisation, and experience and sex of the doctors (table 1). To ensure that the family doctors were unaware of the trial, and of their assigned

	Quantity data*			Quality data†		
	Before study	Year 1	Year 2	Before study	Year 1	Year 2
Group A						
Electrocardiography	1764	1924	1935	305	383	321
Endoscopy	1012	1447	1449	150	167	175
Cervical smears	5444	4804	4263	897	861	719
Allergy tests	764	990	944	61	127	110
Total	8984	9165	8591	1413	1538	1325
Group B						
Radlography						
Chest	2730	2787	2526	498	549	455
Cervical spine	560	619	567	109	108	93
Thoracic spine	298	347	348	59	54	50
Lumbar spine	1128	1212	1127	215	223	187
Pelvis/hip	664	753	844	144	152	139
Knee joint	798	824	715	127	132	99
Ankle joint	340	394	354	67	80	53
Sinuses	662	695	559	110	167	93
Ultrasound						
Liver/biliary tract	584	663	633	93	116	65
Kidneys	128	215	252	67	65	49
Total	7892	8509	7925	1489	1646	1283

*All requests per year. †Requests from 2 months per year.

Table 2: Request forms analysed

	Mean (and quartiles) of number of requests per doctor			
	Test-group A		Test-group B	
	Intervention group	Control group	Intervention group	Control group
Before trial	110 (51, 112, 160)	125 (70, 111, 173)	106 (42, 98, 157)	100 (40, 101, 128)
Year 1	109 (48, 106, 154)	143 (107, 146, 197)	113 (64, 91, 156)	103 (58, 95, 136)
Year 2	102 (42, 100, 159)	141 (79, 135, 190)	101 (54, 89, 137)	100 (62, 91, 134)

Table 3: Request numbers before trial and in years 1 and 2 of trial.

group, they were not informed about the trial. This course was possible because the routine feedback given since 1985 was continued and the new feedback on the tests discussed as part of the trial was integrated within it according to the assignment. It was not practicable for the expert carrying out the assessment of rationality to be unaware of the assignment.

Rationality of requests involved in the trial was assessed by comparison of request data and the accompanying clinical data with regional guidelines. The DCC has prepared guidelines on various tests for primary care, compiled by expert family doctors in collaboration with specialists.¹⁵ Guidelines for tests involved in the trial were used by the expert reviewer to assess the rationality of requests during the trial but they were not distributed to family doctors in the region until the trial ended. Therefore, the participants could not obtain any of these guidelines.

A request was rational when it was in accordance with the guidelines or standards concerned. If not in accordance, it was deemed to be non-rational (unless the request was at the insistence of the patient). Since 1979, DCC request forms have offered the possibility of providing clinical data on the patient and the reason for the request. With time, forms have been filled in satisfactorily. Therefore, in most cases, rationality can be assessed reliably.

To calculate the volume of requests, data from every request for the tests involved in the trial between July, 1988, and December, 1991, were entered on a computer database. Data from 1 year before the trial (July, 1988, to June, 1989) were used as baseline measurement and were compared with data from the first (1990) and second (1991) years of the trial. For each group of tests, the intervention group was compared with its control group. Differences between the baseline measurement and the last year of the trial were compared by the Mann-Whitney test. As an overall test for the two test-groups together, we adapted a test for crossover trials to our design.¹⁶

Rationality data were obtained from all requests in 2 selected months per year, since assessment of rationality for the whole study period was too laborious. Data from March and September, 1989, counted as baseline measurement. They were compared with data from March and October, 1990 (first year of the trial), April and November, 1991 (second year of the trial), and May, 1992 (final measurement). Each request was compared with the guidelines and classified as rational or non-rational. In feedback, emphasis is put on advice to refrain from non-rational requests in future. Hence, we expected a lower number of requests classified as non-rational. In the rationality data it was possible that, owing to holidays or illness, there were no requests from an individual doctor in a month that was to be discussed in a feedback report. These were taken as missing data.

Rationality data were analysed by means of repeated measures ANOVA capable of correcting for missing data.¹⁷ For each family doctor we calculated the number of rational and non-rational requests per test group per month. To see whether the relative rate of non-rational requests was affected by feedback, we expressed the numbers of non-rational requests as a percentage of the total number of requests in each month. By a Wald test in the ANOVA we compared the numbers and percentages of the last three monthly analyses of both groups of family doctors, with pre-trial values as covariates. As an overall test for both test groups together, we adapted the crossover test to our model.¹⁶

Month	Mean number of non-rational requests/total per doctor (%)			
	Test-group A		Test-group B	
	Intervention group	Control group	Intervention group	Control group
March, 1989*	42.2	41.5	60.3	68.6
September, 1989*	41.9	39.4	62.5	62.5
March, 1990	47.6	38.6	50.0	61.6
October, 1990	36.8	37.2	53.7	68.8
April, 1991	29.8	32.9	43.9	61.9
November, 1991	37.1	32.9	55.5	60.2
May, 1992	29.7	37.6	52.2	60.8

*Before trial.

Table 4: Effect of feedback on rationality of test requests

Month	Mean number of non-rational requests per doctor					
	Cervical smear		Chest radiograph		Lumbar spine radiograph	
	I	C	I	C	I	C
March, 1989*	2.73	2.91	2.69	2.21	1.92	1.38
September, 1989*	2.80	2.30	2.58	1.82	1.38	1.54
March, 1990	3.48	2.77	2.22	2.29	0.96	1.39
October, 1990	2.14	2.64	2.29	2.88	1.44	1.22
April, 1991	1.39	2.25	1.64	1.28	0.91	1.65
November, 1991	1.67	2.13	2.17	2.17	1.04	1.24
May, 1992	1.29	2.39	1.32	1.94	1.21	1.67

*Before trial. I=intervention, C=control group.

Table 5: Effect of intervention on rationality of requests for individual tests

Rationality data were also analysed by test. Before the trial, many requests for cervical smears and radiographs of chest and lumbar spine seemed to be non-rational. We expected the number of non-rational requests to decrease for these 3 tests. For individual tests, when the numbers of requests were too small for a repeated measures analysis, rationality was assessed with the Mann-Whitney test. Here the number of non-rational requests was calculated per family doctor per test per monthly analysis. For analyses at test-group level a two-sided significance level of 0.05 was used. For the analyses on rationality at test level, because of multiple testing, a significance level of 0.01 was used.

Results

Numbers of request forms analysed are given in table 2. In the intervention group the number of tests from test-group A decreased by 1% after 1 year and by 7% after 2 years, whereas in the control group the number increased by 14% after 1 year and by 13% after 2 years (table 3). The difference between the intervention and control groups was significant ($p=0.04$). The difference was due mainly to a substantial reduction in requests for cervical smears in the intervention group (37% in 2 years compared with 10% in control group).

Absolute numbers of tests from test-group B decreased in the intervention group by 4% but they also decreased slightly in the control group; the difference between the two groups of doctors was not significant ($p=0.11$).

For both test-groups together there was a significantly greater decrease in numbers of requests in the intervention groups than in the appropriate control groups ($p=0.036$).

The trend for non-rational requests showed a decline in test-group A, but the strongest decline was found in test-group B (table 4). For test-group A, the difference between the intervention and control groups in the decrease in the percentage of non-rational requests per physician did not reach significance ($p=0.11$), whereas for test-group B the difference between the groups was significant ($p=0.04$). Overall for both test groups there was a significant difference between the intervention and

control groups in the decline in percentage of non-rational test requests ($p=0.009$).

Analyses of non-rational requests for individual tests (table 5) showed that the number of non-rational requests for radiographs of the lumbar spine decreased in the intervention group but not in the control group ($p=0.004$). For chest radiography and cervical smears, no significant differences were found between the intervention and control groups in the changes in numbers of non-rational tests ($p=0.29$ and 0.65 , respectively).

Discussion

In general, the feedback provided in the trial affected both the quantity and quality of the ordering behaviour for tests discussed in the feedback. However, before any conclusion can be drawn, we must take into account the following features. In our study, a routine health care activity that had been applied for more than 4 years previously was evaluated in a randomised clinical trial. Several adjustments were therefore necessary. An important restriction was that for the trial we had to choose tests to be discussed in the feedback for the first time. Ideally, we would have discussed in the experimental feedback the tests for which ordering behaviour was least rational. These tests had, however, already been included in the feedback before the experiment. A trial on routine health care also has advantages. Automatically, all family doctors in the region were involved in the trial, so selection bias was ruled out. An important prerequisite for success is the masking of the intervention. The addition of the experimental feedback to the existing routine feedback meant that we could perform the trial without seeking informed consent. During and after the trial, there was no sign that any family doctor suspected the existence of a trial on the feedback provided by our centre. Besides, as we have found since 1985, the family doctors do not customarily discuss feedback reports with colleagues. Therefore, we were confident that randomisation at the level of the individual doctor was appropriate.

Some might find it unethical to withhold information and not to obtain informed consent. This study would not have been possible if we had obtained informed consent. In addition, our intervention was not harmful and normally available information was not withheld. Consequently, the care provided was of at least the usual standard. It might be judged more unethical not to do a trial—effectiveness would remain unknown and an ineffective intervention might undeservedly continue.

The effects achieved in the trial are smaller than those the feedback had in the first years (1985–88).¹² This difference is not surprising, since in that period the feedback was followed not only by a reduction in the number of requests for the tests discussed but also (though to a lesser extent) by a reduction in requests for tests not discussed, probably because of a general learning effect. It is more difficult to achieve a reduction in test requests when numbers have already been reduced before the experiment.

The results we obtained are not uniform. While the number of tests decreased especially for tests from test-group A, rationality predominantly improved for tests from test-group B. The nature of the tests may account for these differences. From separate data it seems that the radiographic and ultrasonographic investigations in test-

group B are often requested at the insistence of patients. Thus, it is more difficult for a family doctor to follow the feedback guidelines and refrain from a request. The same is true for electrocardiography (test-group A). Furthermore, the proportions of rational requests for several tests from test-group A were already high before the trial, endoscopy particularly. For these tests, the feedback could lead to only a minor improvement in rationality. The overall analyses, however, show significant reductions in both the number of tests and the proportion of non-rational requests.

Since rationality was analysed by requests per month, patterns may be affected by seasonal or even monthly changes in the numbers of requests. In March and November, for instance, there are always many requests. These patterns are evident in the intervention group as well as the control group and do not affect the analyses of differences.

The improvement in rationality for test-group B as a whole is not reflected in the three individual tests analysed, probably because the numbers of requests for these individual tests during the 2-month samples were small. Assessment of rationality was too laborious for requests from all months to be possible.

Feedback is reported to be more effective when it is personal and provided by a respected colleague.^{1,2} Probably feedback has to be repeated to maintain an effect.¹⁸ Our feedback bears these features. This trial shows that the routine feedback is indeed effective in reducing the volume and improving the rationality of test requests. The extent of the effect can be assessed from the results of an observational study, carried out before this trial. In that study, reductions of up to 95% were noticed and the feedback resulted in a substantial reduction in the total number of requests, associated with major cost savings.¹² If the feedback described here is to be applied elsewhere, we recommend discussion of many more tests than those involved in our trial. The two major prerequisites are a respected expert to give feedback and a request form that offers the possibility of giving clinical data on the patient and the reason behind the request.

References

- 1 Mugford M, Banfield P, O'Hanlon M. Effects of feedback of information on clinical practice: a review. *BMJ* 1991; **303**: 398–402.
- 2 Buntinx F, Winkens RAG, Grol RPTM, Knottnerus JA. Influencing diagnostic and preventive performance in ambulatory care by feedback and reminders: a review. *Family Pract* 1993; **10**: 219–28.
- 3 Axt-Adam P, van der Wouden JC, van der Does E. Influencing behaviour of physicians ordering laboratory tests: a literature study. *Med Care* 1993; **31**: 784–94.
- 4 Fowkes FGR, Evans KT, Hartley G, et al. Multicentre trial of four strategies to reduce use of a radiological test. *Lancet* 1986; **i**: 367–70.
- 5 Tierney WM, Miller ME, McDonald CJ. The effect on test ordering of informing physicians of the charges for the outpatient diagnostic tests. *N Engl J Med* 1990; **322**: 1499–504.
- 6 Martin AR, Wolf MA, Thibodeau LA, Dzau V, Braunwald E. A trial of two strategies to modify test-ordering behaviour of medical residents. *N Engl J Med* 1980; **303**: 1330–36.
- 7 Berwick DM, Coitin KL. Feedback reduces test use in a Health Maintenance Organization. *JAMA* 1986; **255**: 1450–54.
- 8 Marton KI, Tul V, Sox HC. Modifying test-ordering behaviour in the outpatient medical clinic; a controlled trial of two educational interventions. *Arch Intern Med* 1985; **145**: 816–21.
- 9 Tierney WM, McDonald CJ, Martin DK, Hui S, Rogers MP. Computerized display of past test results: effect on outpatient testing. *Ann Intern Med* 1987; **107**: 569–74.
- 10 Dowling PT, Alfonsi G, Brown MI, Culpepper L. An education program to reduce unnecessary laboratory tests by residents. *Acad Med* 1989; **64**: 410–12.

- 11 Schectman JM, Elinsky EG, Pawlson LG. Effect of education and feedback on thyroid function testing strategies of primary care clinicians. *Arch Intern Med* 1991; 151: 2163-66.
- 12 Winkens RAG, Pop P, Grof RPTM, Kester ADM, Knottnerus JA. Effect of feedback on test ordering behavior of general practitioners. *BMJ* 1992; 304: 1093-96.
- 13 Pop P, Winkens RAG. A diagnostic centre for general practitioners: results of individual feedback on diagnostic actions. *J R Coll Gen Pract* 1989; 39: 507-08.
- 14 Rutten GEHM, Thomas S. NHG-Standaarden voor de huisarts. Utrecht: Bunge, 1993.
- 15 Beusmans GHMI, Does RJMM. Werkafspraken in het Diagnostisch Centrum Maastricht. *Huisarts Wet* 1988; 31: 154-58.
- 16 Armitage P, Berry G. Statistical methods in medical research, 2nd ed. Oxford: Blackwell, 1987.
- 17 Jennrich RI, Schluchter MD. Unbalanced repeated-measures models with structured covariance matrices. *Biometrics* 1986; 42: 805-20.
- 18 Eisenberg JM. Physician utilization: the state of research about physicians' practice patterns. *Med Care* 1985; 23: 461-83.

Viewpoint

Risk assessment and factor VIII concentrates

N C Hughes-Jones

Litigation in France and in Switzerland, the existence of a Commission of Inquiry in Canada, and an investigation by the US National Academy of Sciences Institute of Medicine have kept the issue of HIV infection from factor VIII therapy in the public domain. There is much public confusion, particularly as to why so many haemophiliacs became infected with HIV when the clinicians who dealt with these patients were caring and concerned. Much discussion has been about uncertainty over the nature of the aetiological agent in factor VIII concentrates during the period 1983-85. What has been ignored is that their continued use during these years was determined by risk analysis; an understanding of events in these terms removes the confusion.

The transmission of infectious agents by transfusion is well known and the resultant morbidity and mortality are well documented, but no attempt has been made to discuss these hazards in terms of risk analysis¹—ie, in terms of benefit and harm. The aspect of risk analysis that concerns us most is risk tolerance, a term first introduced by Layfield² in his report on the benefits and risks of nuclear power in the UK. He commented that "tolerability does not mean acceptability. It refers to the willingness to live with a risk to secure certain benefits and in the confidence that it is being properly controlled. To tolerate a risk means that we do not regard it as negligible or something that we might ignore, but rather as something we need to keep under review and to reduce still further if and as we can."

The weighing of benefits and harm is subjective and so we need some kind of yardstick. Examination of the risk of hepatitis from blood and its products is illuminating. We have lived for 50 years with the knowledge that hepatitis can follow transfusion; before the introduction of screening tests, reports of the incidence of post-transfusion hepatitis varied between 1% and 10% with mortality varying between 2% and 25% of those with clinically overt disease.³ There was the implicit but unformulated assumption that these values were a tolerable risk but were not acceptable, as shown by the volume of relevant research at the time.

In January, 1983, two articles that appeared in the

New England Journal of Medicine on the finding of AIDS in haemophiliacs were accompanied by an editorial by Dr Jane Desforges stating that the evidence that "large pool" factor VIII concentrates contained an infective agent could not be ignored. She suggested that we should seriously consider returning to use of cryoprecipitate—"preventing the complications of the present treatment may have to take precedence over preventing the complications of haemophilia itself".⁴ During the ensuing two years, evidence continued to accumulate that there was indeed some form of agent in factor VIII concentrates responsible for the signs and symptoms of the immunological anomalies associated with AIDS, but the nature of that agent was unknown.⁵ However, by the end of 1984 it had become fairly certain that a virus was present in factor VIII concentrates and this coincided with evidence that the causative agent could probably be inactivated by heat treatment.⁶

With the wisdom of hindsight, we know that the decision of most haemophilia centres to continue the use of factor VIII concentrates during 1983-84 was wrong even though most of the haemophiliacs had probably already been infected. The reason for the continued use of concentrates was simple—namely, that there was a general consensus that the benefits obtained from factor VIII concentrate therapy greatly exceeded any harm that might ensue. The incidence of AIDS in Europe at the beginning of 1984 was about 1 in 1000 treated haemophiliacs, although it was about twice that in the USA.⁷ Several instances of "pre-AIDS" were described but it was believed that only 10-15% would progress to the full-blown disease, and Montagnier had said that the signs could regress.⁵ In the widely circulated UK publication, *General Practitioner*, there was a report in October, 1984, entitled "Need for factor VIII outweighs AIDS risk", quoting a UK Haemophilia Centre Director saying that "although (AIDS) is an important cause of morbidity and mortality it is not a major worry". The report added that it was important to remember that haemophiliacs had other problems: liver disease—even cirrhosis and death—neurological problems, and analgesic abuse were all quite common. This view was supported by haemophiliacs; at a meeting of the World Federation of Haemophiliacs in September, 1984, there was a statement that "urged haemophiliacs not to withhold treatment for fear of AIDS".⁸