Randomised controlled trial of reminders to enhance the impact of audit in general practice on management of patients who use benzodiazepines

Richard Baker, Azhar Farooqi, Clare Tait, Sara Walsh

Abstract

Objective—To determine whether reminder cards in medical records enhance the effectiveness of audit with feedback in improving the care of patients taking long term benzodiazepine drugs.

Design—Randomised trial, practices receiving feedback only in one group and practices receiving feedback plus reminder cards in the other group.

Setting—18 general practices in Leicestershire.

Subjects—Random samples of patients who had been taking a benzodiazepine anxiolytic or hypnotic drug for four weeks or longer.

Main outcome measures—Entries in medical records indicating compliance with five criteria of care: assessment of suitability for withdrawal; being told about dependency; withdrawal being recommended; withdrawal or continuing medication; and a consultation with the general practitioner in the past year. Data were collected before and after feedback or feedback plus reminders.

Results—Of a total population of 125 846 registered with the 18 practices, 2409 (1.9%) had been taking a benzodiazepine for four weeks or longer. Of the 742 in the first samples, 543 (73.2%) were women, the mean (SD) age was 68.7 (14.9) years, and they had been taking a benzodiazepine for 10.1 (6.7) years. The number of patients whose care complied with the criteria rose after the interventions to implement change. The increase was greater in practices receiving feedback plus reminders for only two of the five criteria: "told about dependency" increasing from 52 (11.1%) to 118 (25.8%) in the feedback only group, and from 27 (10.5%) to 184 (43.0%) in the feedback plus reminders group; odds ratio (OR) 1.46 (95% confidence interval (95% CI) 1.32 to 5.21); and "consulted in the past year" increasing from 434 (93.1%) to 411 (95.8%) in the feedback only group and 255 (96.6%) to 400 (99.8%) in the feedback plus reminders group, OR (95% CI) 13.5 (2.01 to 330.3).

Conclusions—Reminder cards had only a limited effect and cannot be recommended for routine use. There were improvements in the care of patients of both groups of practices and further studies are indicated to determine the impact of both systematically developed criteria and reminders embedded into restructured medical records.

Keywords: randomised controlled trial; reminders.

Introduction

Since the creation of medical audit advisory groups in 1991 clinical audit has become increasingly common in general practice. However, questions have been raised about the effectiveness of audit in implementing change. Because audit is now widely undertaken, it may offer an important method of improving performance, but strategies are required to ensure that it does regularly lead to appropriate change.

Several methods are available for implementing change in the performance of clinicians. In a systematic review of 91 studies of methods for implementing clinical guidelines, strategies were more effective if they were used within the consultation—for example, reminders or restructured medical records. In a systematic review of 99 trials of education strategies, those that were effective included reminders, patient mediated interventions, and outreach visits. In both reviews audit with feedback was found to have some effect, but to a lesser extent. However, no intervention could be relied on to be always effective.

Few implementation studies have been undertaken in United Kingdom general practice. In a study of guidelines for the initial investigation of infertile couples in the Grampian Region, study practices received structured infertility management sheets which acted as reminders at the time of consultations. There were significant improvements in compliance with the guidelines in the study practices in comparison with controls.

In a study of practices in east London, guidelines for asthma and diabetes were implemented through practice based education assisted by the use of a stamp to structure the medical records and act as a prompt at the time of consultations. There were improvements in the intervention practices, but in this study the use of prompts was not randomised and firm conclusions about their effectiveness were not possible. Therefore, further evaluation of the effectiveness of reminders in records is required.

The management of patients receiving long term benzodiazepines is a suitable topic for
audit as there is evidence that care could be improved. Concern about the consequences of long term prescribing has been expressed for more than 15 years. The number of prescriptions for these drugs has declined, although a recent audit in 15 general practices found that 3.7% of the population were taking benzodiazepines. This is despite evidence of the effectiveness of several methods for helping patients to stop their medication, including a letter from the general practitioner, the provision of information, and psychological techniques. Therefore, a study was undertaken to determine whether reminder cards in medical records plus audit with feedback was a more effective strategy than audit with feedback alone in improving the care of patients taking long term benzodiazepine drugs.

Method

AUDIT CRITERIA
A long term benzodiazepine user was defined for this study as a patient who had been taking either a benzodiazepine anxiolytic or hypnotic drug for a period of four weeks or longer, in accordance with the recommendation in the British National Formulary about the appropriate duration of treatment to minimise the risk of dependency. The audit criteria were developed in 1994 with a standard method. Guidelines produced by the Mental Health Foundation and guidance from the Royal College of Psychiatrists were used to identify key elements of care. Research evidence relating to each element of care was identified by a Medline literature search including English language papers published since 1989, supplemented by hand searches of general practice journals—British Journal of General Practice and Family Practice. Relevant references in papers encountered in these searches were also consulted. Criteria were formed from the research encountered, and given the priorities "must do" or "should do" and submitted to peer review by two general practitioners with particular knowledge of the topic and a psychologist providing a drug dependency service. The criteria are shown in the box.

Subjects and data collection
A letter inviting practices to participate in an audit of prescribing benzodiazepine was sent from Leicestershire medical audit advisory group to all 147 practices in the county in 1994. We were seeking between 20 and 30 participating practices and no additional attempts were made to recruit practices. A visit was made by one researcher to the practice managers of those practices that wished to take part to explain the audit and collect basic information about the practice. The practices were told that the study was an audit and to this extent were blinded to the trial of reminders.

In each practice a researcher compiled a list of long term patients who used benzodiazepines with the help of the registers of prescribing on computer, supplemented by checking computer details of patients and prescribing with manual records to ensure that each patient met the definition of long term user. In one practice without a computer, a register of users was developed by recording patients' names when a prescription for a benzodiazepine was issued. After the first data collection, practices were randomly allocated, with a table of random numbers and without stratification, to two groups, one to receive feedback only, and the other to receive feedback plus reminder cards. For the second data collection, the list of patients was supplemented by a further search of the practice prescribing system to identify new patients. No types of long term user were excluded from the study. In each practice, data about long term care were collected from a random sample of users. The sample sizes were in proportion to the total study size. To detect an increase from 35% to 45% of patients in criterion 3 (patients are offered withdrawal), with a power of 80% and significance level of 5% a sample of 295 subjects per experimental group was required. However, to allow for cluster sampling this was increased by a factor of about 1.5 to 450. Data were collected from the clinical records of long term users before and after the steps taken to implement change, the interval between data collections being about 12 months. As the first data collection involved the creation of a register of long term users no information was collected about criterion 1 (box). Furthermore, criteria 5, 6, 8, 9, and 10 related to patients undergoing withdrawal. As few patients would be undergoing withdrawal

Criteria for the management of long term benzodiazepine users

MUST DO
1 All patients on long term benzodiazepines must be recorded on a specific register or a computerised repeat prescribing system.
2 The records must show that long term users of benzodiazepines have been assessed on their suitability for withdrawal.
3 Patients assessed as suitable for withdrawal (a) must be told about dependency, and (b) withdrawal must be recommended and this recorded in their notes.
4 Patients who are unsuitable for withdrawal or are not ready to withdraw must continue to receive their medication.
5 Withdrawal from benzodiazepines must be a phased reduction over a period no less than six weeks.
6 The records must show that patients undertaking withdrawal have received information about symptoms to expect and coping strategies.
7 The records must show that patients failing to withdraw or not ready to withdraw are reviewed at least annually.

SHOULD DO
8 During withdrawal, patients should be seen at regular intervals for monitoring and support.
9 The records should show that patients attempting withdrawal but experiencing difficulty have been offered additional support or specialist treatment.
10 The records should show that patients receiving a benzodiazepine with a short half life have been transferred to a drug with a longer half life before withdrawal.
data about these criteria were not included in the trial. Data about five criteria were collected from the general practice records of the study patients: (a) a record of being assessed on suitability for withdrawal (criterion 2); (b) a record of those suitable for withdrawal being told about dependency (criterion 3 (a)); (c) a record of those suitable for withdrawal being offered withdrawal (criterion 3 (b)); (d) whether patients were continuing medication, undergoing withdrawal, or have withdrawn from their medication (criterion 4); (e) those continuing medication have been seen by a general practitioner in the past year (criterion 7). Patients were classified as unsuitable for withdrawal if they had epilepsy, were terminally ill, had organic brain disease, had a history of psychosis, or were thought to be likely to resort to illicit psychotropic drugs.

Standard data extraction rules were developed and data extracted by a researcher, one for the first data collection and another for the second. We were unable to blind the second data collector to study groups. To test accuracy of data extraction, random samples of 10 records from each of the two data collections were drawn and data were independently extracted by a second researcher. Levels of agreement and \( \kappa \) were calculated for each criterion.

**INTERVENTIONS TO IMPLEMENT CHANGE**

After the first data collection, all practices received a copy of the audit criteria, including a summary, detailed explanation, and references, which justified the priorities “must do” or “should do”. They also received feedback which enabled them to compare their performance with the criteria and the other participating practices. The feedback indicated the percentage of patients whose care was in compliance with each criterion. After the practices had received the criteria and feedback, a semi-structured interview was carried out on one member of each practice to assess their views about the criteria and identify obstacles to change.\(^{10}\)

The group of practices randomised to receive reminders as well as feedback were given cards to place in the records of long term users. Printed on the cards were the criteria (box) and on the reverse were boxes for the practitioner to record when the patient was assessed on suitability for withdrawal, discussion of withdrawal, and whether withdrawal was planned. Practices were not offered assistance with placing the cards in the records and could choose not to make use of them. This approach was adopted as it would be more likely to represent the way such cards would be used in routine audits.

**ANALYSIS**

In comparing changes between the two data collections a Z statistic was calculated from the changes in each practice in the proportion of patients whose care complied with the criteria.\(^{10}\) The analysis between the two groups of practices sought to determine the effect of reminders in enhancing the impact of audit with feedback. Logistic regression was undertaken with SABRE software, with a log normal mixture model with a practice specific random effect in the linear predictor.\(^{10}\) The analysis involved modelling the log OR — that is, the ratio of the odds of compliance with a criterion compared with non-compliance in the feedback plus reminders group of practices with the odds of compliance compared with non-compliance in the feedback only group. The model accounted for the degree of compliance in the first data collection. Practices given reminder cards were included in the feedback plus reminder group whether or not they chose to place the reminder cards in records. The analysis was based on intention to treat.

**Results**

**SUBJECTS**

Twenty practices agreed to take part but two withdrew after the first data collection. Both had been randomised to the feedback plus reminders group and therefore the number of patients in this group in the first data collection was reduced. Two of the eight practices offered reminder cards chose not to place them in medical records. The 18 practices had a total patient population of 125 846, of whom 2409 (1.9%) met the study definition of long term benzodiazepine user.

Table 1 shows the characteristics of the practices. There were no significant differences between the groups for numbers of patients, underprivileged scores (\( z \) tests), or numbers of partners (Mann-Whitney \( U \) test). There were 1618 (2.2% of total patient population) long term users in the feedback only practices and 791 (1.6%) in the feedback plus reminders practices. However, the patients in each group

---

**Table 1** Characteristics of the practices in the feedback only and feedback plus reminders groups

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Feedback practices</th>
<th>Feedback + reminders practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of practices</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Number computerised</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Number training</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Number of partners (median range)</td>
<td>4 (1 to 7)</td>
<td>3.8 (1 to 7)</td>
</tr>
<tr>
<td>Number of patients (mean, SD)</td>
<td>7508 (3343, 2120 to 12310)</td>
<td>6345 (3422, 2290 to 11991)</td>
</tr>
<tr>
<td>Underprivileged score (mean, SD)</td>
<td>4.2 (14.7, -13.9 to 26.6)</td>
<td>5.5 (16.7, -9.6 to 26.4)</td>
</tr>
</tbody>
</table>

**Table 2** Age and sex of patients in each study sample, and those taking different types of benzodiazepine medication (some patients were taking more than one benzodiazepine)

<table>
<thead>
<tr>
<th>First data collection</th>
<th>Second data collection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Feedback</strong></td>
<td><strong>Feedback + reminders</strong></td>
</tr>
<tr>
<td><strong>Number</strong></td>
<td>475 (15.2)</td>
</tr>
<tr>
<td><strong>Age (y, mean (SD))</strong></td>
<td>68.2 (15.2)</td>
</tr>
<tr>
<td><strong>Sex (mean (SD))</strong></td>
<td><strong>F</strong></td>
</tr>
<tr>
<td><strong>M</strong></td>
<td>136 (28.6)</td>
</tr>
<tr>
<td><strong>Drugs (n (%))</strong></td>
<td><strong>Total</strong></td>
</tr>
<tr>
<td><strong>Temazepam</strong></td>
<td>104 (32.5)</td>
</tr>
<tr>
<td><strong>Nirazepam</strong></td>
<td>115 (22.6)</td>
</tr>
<tr>
<td><strong>Diazepam</strong></td>
<td>130 (25.8)</td>
</tr>
<tr>
<td><strong>Chlordiazepoxide</strong></td>
<td>40 (7.9)</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td>55 (10.9)</td>
</tr>
</tbody>
</table>
of practices had been taking their benzodiazepine medication for a similar time (mean (SD) 9.9 (6.7) years for the feedback only group and 10.4 (6.7) years in the feedback plus reminders group). There were no significant differences in the mean ages or sex distribution of patients in the different study groups (table 2).

**DATA ANALYSIS**

The assessment of data accuracy had agreement of 94% in the first data collection and 98% in the second. In the first data collection agreement was complete for two criteria: whether medication was being continued or withdrawn and whether the patient had consulted the general practitioner in the past year. For offering withdrawal \( \kappa \) was 0.9, for assessment of suitability 0.8, and for informing the patient about dependency 0.6. For the second data collection agreement was complete for four criteria: assessment of suitability, offer of withdrawal, whether withdrawing or continuing medication, and consulting the general practitioner in the past year. For being informed about dependency \( \kappa \) was 0.8.

For both groups of practices there were substantial changes between the two data collections in the levels of compliance with the criteria (table 3). There were changes in both the recording of the process of care and also outcome — withdrawal from medication. Although only 11 (1.5%) patients were undergoing withdrawal before the strategies to implement change, afterwards 81 (8.2%) had stopped their medication and 13 (1.3%) were still undergoing withdrawal. The differences between the two groups of practices were less substantial (table 4). Compliance was greater in the group receiving feedback plus reminders for only two of the five criteria, being told about dependency (criterion 3 (a)) and reviewed in the past year (criterion 7).

Interviews were held in 16 practices, eight in each group. All interviewees were general practitioners, except for one who was a practice manager. Thirteen practices had found the criteria helpful and all found the feedback helpful. Twelve practices reported or anticipated difficulties in changing their performance, five indicating shortage of time, and another shortage of staff. Six practices thought that patients taking benzodiazepines long term were often resistant to advice about withdrawal.

**Discussion**

This study has investigated whether the provision of reminder cards to place in patients’ records will enhance the effectiveness of audit with feedback in changing performance in general practice. Reminder cards were associated with improved compliance for only two of five criteria, and in particular there was no effect on outcome — withdrawal from medication. Other studies have suggested that reminders in general practice might be an effective strategy for changing performance and the reasons for the findings in this study must be considered.

One possible explanation is that the format of reminders may influence their effectiveness. If they are presented as a simple list, as in this study, they may be less effective than when embedded in the medical record to prompt care directly. A second possibility is that there may be a ceiling to the amount of change that a practice can introduce in a limited period, and because substantial changes occurred in both groups of practices the potential impact of reminders may have been reduced. The criteria used in this study were given priorities on the basis of research evidence and may have had a greater impact on performance than those that are less systematically developed. Further research to investigate this possibility is required.

Thirdly, the apparent effect of the reminders may have been reduced because two practices chose not to use them. Steps might have been taken to enable these practices to use the reminders—in fact, the provision of help in placing the cards in records and additional meetings with the practice teams. However,

### Table 3 Changes in compliance with guidelines for care of patients between the first and second data collections

<table>
<thead>
<tr>
<th>Criteria</th>
<th>First data collection</th>
<th>Second data collection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>n (%)</td>
</tr>
<tr>
<td>2 Assessed on suitability</td>
<td>741</td>
<td>51 (6.9)</td>
</tr>
<tr>
<td>3 (a) Told about dependency</td>
<td>728</td>
<td>79 (10.9)</td>
</tr>
<tr>
<td>3 (b) Withdrawal offered</td>
<td>726</td>
<td>201 (27.7)</td>
</tr>
<tr>
<td>4 Withdrawn or withdrawing</td>
<td>740</td>
<td>11 (1.5)</td>
</tr>
<tr>
<td>7 Reviewed in past year</td>
<td>730</td>
<td>689 (94.4)</td>
</tr>
</tbody>
</table>

### Table 4 Numbers and percentages of patients in first and second data collections in the two groups of practices. C = compiled with criterion, N = not compiled

<table>
<thead>
<tr>
<th>Criterion</th>
<th>First data collection</th>
<th>Second data collection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C</td>
<td>N</td>
</tr>
<tr>
<td>2 Assessed on suitability for withdrawal</td>
<td>25</td>
<td>450</td>
</tr>
<tr>
<td>3 (a) Told about dependency</td>
<td>52 (5.3)</td>
<td>419 (88.9)</td>
</tr>
<tr>
<td>3 (b) Withdrawal offered</td>
<td>102 (21.7)</td>
<td>368 (78.3)</td>
</tr>
<tr>
<td>4 Withdrawal or withdrawn</td>
<td>46 (9)</td>
<td>464 (91)</td>
</tr>
<tr>
<td>7 Reviewed in past year</td>
<td>434 (93.1)</td>
<td>32 (6.9)</td>
</tr>
</tbody>
</table>
these steps would have constituted facilitation. Thus several strategies would have been used to encourage behavioural change and it would have been difficult to find which was responsible for any change that took place. Hence, the findings of our pragmatic intention to treat analysis are likely to reliably indicate the impact of the routine use of reminders in records.

The impact of reminders may have even been overestimated by this study. It is possible that the differences between the groups of practices may have been restricted to recording, the effect of reminders being to prompt clinicians to record their actions in the notes. The fact that the reminders failed to influence outcome may support this view.

The number of patients from the feedback plus reminders group of practices in the first data collection was reduced because of the withdrawal of two practices from the study. However, it is unlikely that this would have reduced the ability of the trial to detect a practically useful effect due to the reminder cards. Although the proportion of long term users in the feedback plus reminders group was lower, the patients were similar in terms of age, sex distribution, and duration of drug use. Therefore, it is unlikely that the patients in the reminders group were more resistant to advice about withdrawal.

Conclusions

The provision of reminder cards for placing in medical records only marginally enhanced the effectiveness of audit with feedback in improving the management of long term benzodiazepine users. They cannot be recommended for routine use. Reminders embedded in structured medical records may be more effective but further studies are required to answer this question and also to find the effectiveness of prioritised, evidence linked criteria.

We thank all the participating practices, Leicestershire Medical Audit Advisory Group for supporting the study, Professor Byron Jones and Michelle Lawson for help with the statistical analysis, Noelle Robertson for help with developing the audit criteria, and Dr M Ashworth, Dr A Tyte, and Ralph Barker for peer reviewing them. We are grateful to Dr Genie Feder for comments on an earlier draft of this paper. The study was funded by the Department of Health and Trent Regional Health Authority Directorate of Research and Development.