Evidence based educational outreach visits: effects on prescriptions of non-steroidal anti-inflammatory drugs

E Bernal-Delgado, M Galeote-Mayor, F Pradas-Arnal, S Peiró-Moreno

Aims: To evaluate the effectiveness of an evidence based group educational outreach visit on prescription patterns of non-steroidal anti-inflammatory drugs (NSAIDs) in primary care.

Design: Randomised controlled simple blind trial, with randomisation into three groups: experimental (evidence based educational outreach visit), placebo (conventional education session), and control (without intervention).

Setting: The 24 primary care centres of the National Institute of Healthcare Network in a rural province of Aragon, Spain.

Participants: The 24 primary health care teams of the network, with 158 general practitioners (GPs). The teams were randomised into the groups, experimental (8 teams, 48 GPs), placebo (8 teams, 54 GPs), and control (8 teams, 56 GPs).

Intervention: Experimental group: one group educational outreach visit, conveying data based on a systematic review of the literature that was reinforced with printed material; placebo group: one non-structured educational session; control group: no intervention. Both educational sessions emphasised that there are no differences in the effectiveness of the NSAIDs reviewed (diclofenac, piroxicam, and tenoxicam); a recommendation was made to prescribe diclofenac over tenoxicam because of price differences.

Main outcome measures: Changes in the number of prescriptions prescribed for each of the drugs and changes in the cost per package of NSAIDs prescribed during the six months before, and after the intervention.

Results: There were no differences in the basal characteristics of the three groups, except for the number of prescriptions during the six months before the intervention. Prescriptions for NSAIDs decreased homogeneously in the three groups. For tenoxicam, the experimental group reduced prescriptions by 22.5% (95%CI: 34.42 to -10.76), compared with a reduction of 9.78% (95%CI: -17.70 to -1.86) in the placebo group and an increase of 14.44% (95%CI: 5.22 to 23.66) in the control group. The average cost per prescription decreased by 1.91% (95%CI: -0.33% to -3.49%) in the experimental group, 0.16% (95%CI: -0.27% to -2.93%) in the placebo group, and rose by 1.76% (95%CI: 0.35% to 3.17%) in the control group.

Conclusions: Evidence based educational outreach visits are more effective than no intervention at all. Results suggest that evidence based educational outreach visits are incrementally more effective than conventional educational sessions, which in turn are more effective than no intervention at all.

During the past few years different strategies to modify or improve medical practices have been implemented and evaluated. These have included distributing educational materials and guides of practice, holding conferences, informational sessions, educational outreach visits, interventions by local opinion leaders, meetings to hear patients’ opinions, conducting audits with or without feedback, sending reminders, and other activities borrowed from the field of marketing, fostering processes to reach consensus on the local level and various combinations of all of these interventions. Interpreting the results of these interventions is a complex task, making it difficult to draw conclusions concerning their effectiveness, the scope of their expected effect or their comparative advantages.

An educational outreach visit (academic detailing) consists of a visit to a healthcare provider in his or her own setting by a trained educator. A recent review has shown that these visits are effective in modifying the behaviour of healthcare professionals, and are particularly useful in changing prescribing practices. However, little has been studied of the effectiveness of this technique applied to groups of professionals, or when recommendations are made based on the systematic review of scientific evidence. This last aspect is relevant, because the additional effort of conducting a systematic review of the evidence to prepare the visit should translate into an increase in the effectiveness of these visits compared with conventional educational sessions.

The object of this study is to evaluate the effectiveness of group educational outreach visits that convey information based on the systematic review of evidence in modifying prescription patterns for non-steroidal anti-inflammatory drugs (NSAIDs) in primary care.

METHODS

Design

Pragmatic, simple blind trial, with random assignment to one of three groups: experimental group (participating in an evidence based educational outreach visit), placebo group (participating in a conventional educational session), and control group (with no intervention).

Setting

The study was conducted in the primary health care network of the Instituto National de la Salud (National Institute of Healthcare) in Teruel, a rural province in Aragon, Spain. The network is comprised of 26 primary care teams (24 at the start of the study), which usually include six or seven general practitioners (GPs), one to three paediatricians, and 6 to 10 nurses.

www.jech.com
Among the relevant features of this health care system administered by the Spanish government are the universal nature of coverage, the cost free status of care (except 40% copayment of pharmaceutical products for non-retired users), funding of health centres by means of global budgets, and the semi-civil servant status of the doctors, who are paid by salary.

Population
The 24 primary care teams that existed in Teruel at the start of the project participated in the study. These included a total of 158 GPs. By using a table with randomised numbers, the teams were assigned to one of three groups: the experimental group (8 teams, 48 GPs), the placebo group (8 teams, 54 GPs), and the control group (8 teams, 56 GPs). All of the groups were blind to the randomisation.

Systematic review of the effectiveness of NSAIDs
Three active agents from the group of NSAIDs were selected for review: piroxicam, tenoxicam, and diclofenac. These drugs were selected for the following reasons: (a) they are frequently used in primary care; (b) they are drugs that are prescribed and managed by GPs, with very little input from specialists; (c) they are practically identical in terms of effectiveness, though vary considerably in price; (d) the environment studied (fig 1) showed wide variations in their use, with differences in rates of prescription per inhabitant and month, adjusted for the age and sex of the population, of up to 7 times for piroxicam, 18 for tenoxicam, and 22 for diclofenac.

The effectiveness and safety of piroxicam, tenoxicam, and diclofenac used to treat arthrosis was analysed in light of the following measures of results: decrease in pain, residual pain, pain with movement, pain the day after onset of treatment, appearance of adverse effects, interruption of treatment because of adverse effects, suspension of treatment because of subjective ineffectiveness, subjective effectiveness for the doctor, and subjective effectiveness for the patient. The search of the literature was based on the review of pre-existing systematic reviews of the effectiveness of the NSAIDs in the Cochrane Library, using two complementary search strategies (appendix 1). Only reviews meeting the following criteria were accepted: (1) they included only randomised, double blind clinical trials; (2) they included the active drugs that were the subject of this analysis; (3) they analysed at least one of the measures of results cited; (4) they achieved a minimum of 15 points out of a possible 28 on the scale used for quality analysis (appendix 2). Of the three reviews found, one was selected for the study and the other two were excluded. Appendix 3 provides the references of the reviews and, when applicable, the criteria used for exclusion.

The systematic review led to the main conclusions that there are no differences between the three drugs studied, except for a lower rate of interruptions of treatment with tenoxicam compared with piroxicam. In light of these conclusions, during the outreach visit, diclofenac was recommended over the other drugs because it is cheaper, and when an oxicam was required, piroxicam was recommended, unless there was a risk of suspending treatment. The comparative information was summarised in a printed brochure, similar to those given out by pharmaceutical companies (available from the authors upon request), and was distributed during the visit. The visit emphasised the evidence based methodology followed, the conclusions reached after the review, and the recommendations formulated above (table 1).

Intervention
The experimental intervention consisted of holding a group educational outreach session with the GPs of each of the primary care teams. During the session, oral and visual information was presented about the effectiveness and safety of the NSAIDs reviewed. The printed brochure was distributed. Diclofenac was recommended as the anti-inflammatory of choice to treat arthrosis and, when an oxicam was required, piroxicam was recommended as the best alternative. The visits were held in each primary care centre and were led by the primary care pharmacist for the area of Teruel, who had been previously trained in the principles of evidence based medicine (EBM) and the basic tools for the dissemination of this information. The GPs comprising the placebo group participated in a conventional educational session about NSAIDs, and did not receive printed support material. The primary care

Table 1 Conclusions and recommendations resulting from the systematic review conveyed to the experimental group

<table>
<thead>
<tr>
<th>Conclusions</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Although there is some evidence that tenoxicam is more effective than piroxicam, it is not conclusive. The choice of tenoxicam may be based on the fact that there is a lower interruption rate.</td>
<td>Diclofenac is recommended as the anti-inflammatory drug of choice to treat arthrosis. If an oxicam is indicated, the recommended drug is piroxicam, except when there is a threat of interrupting treatment.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
area pharmacist also conducted these sessions, although before she received training in EBM, and without her having any knowledge of the results of the systematic review. The control group received no information during the period of the study. The intervention period lasted less than four weeks.

**Measurements of results**
The main outcome was defined as the detection of changes in prescription patterns for the active drugs, and other related drugs, measured as the proportional change in the number of packages prescribed during the six months before intervention and the six months afterwards. We expected to observe a decrease in the prescription of the drugs that had not been recommended, and an increase in the prescription of those that had. A secondary outcome was the average cost per package during the six months after the intervention, minus the average number of packages prescribed per doctor and the cost per prescription referred to the average of the six months before the intervention.

**Other variables**
Other variables that were analysed were the percentages of retirees and women, the average number of medical visits per inhabitant per year, the average number of patients seen by each doctor per day and each doctor’s previous prescription patterns, divided by the latter. Significant differences were established in function of the confidence intervals of 95% (95%CI) of this reduction. The calculation of the 95%CI values was based on eight people (the primary care teams) in each one of the groups.

**Results**
Because prescriptions of NSAIDs show seasonal behaviour, changes in costs per prescription (as a secondary variable) were also evaluated with a deterministic time series analysis. The object of this analysis was to find a function to serve as the analytical representation of a set of data that evolves over time with reference to an imaginary line. Of the functions available, the S or Gompertz curve was selected. Its analytical expression, similar to linear regression, is: \( y = \exp(a + (b/t)) \); where \( a \) is the constant for the model, \( b \) is the slope \( y \) to time. This model best adjusts to the empirical data for the 12 months before intervention, calculated with the \( r^2 \) coefficient of determination and Snedecor’s \( F \) test. A time series adjusted curve of the cost per prescription for each group and its confidence intervals was thus plotted, and the post-intervention changes registered in the groups were projected on this adjusted curve. The reduction in the cost in prescription (cost of average package during the six months after the intervention, minus the cost of average package during the six previous months, divided by the latter) between the periods prior and post-intervention was also calculated for each of the groups, with their respective confidence intervals of 95%.

**Results**
Table 2 shows the basal characteristics of the experimental, placebo, and control groups. There were no significant differences between the groups in terms of the percentage of retirees or women, average number of patients seen by each doctor each day, the average number of visits per inhabitant per year, or the average cost per prescription. However, there were significant differences in the number of packages prescribed per month in the six months before intervention, which was higher in the experimental group compared with the control group. Participation in the group sessions was also evaluated with a deterministic time series analysis. The object of this analysis was to find a function to serve as the analytical representation of a set of data that evolves over time with reference to an imaginary line. Of the functions available, the S or Gompertz curve was selected. Its analytical expression, similar to linear regression, is: \( y = \exp(a + (b/t)) \); where \( a \) is the constant for the model, \( b \) is the slope \( y \) to time. This model best adjusts to the empirical data for the 12 months before intervention, calculated with the \( r^2 \) coefficient of determination and Snedecor’s \( F \) test. A time series adjusted curve of the cost per prescription for each group and its confidence intervals was thus plotted, and the post-intervention changes registered in the groups were projected on this adjusted curve. The reduction in the cost in prescription (cost of average package during the six months after the intervention, minus the cost of average package during the six previous months, divided by the latter) between the periods prior and post-intervention was also calculated for each of the groups, with their respective confidence intervals of 95%.

**Results**
Table 2 shows the basal characteristics of the experimental, placebo, and control groups. There were no significant differences between the groups in terms of the percentage of retirees or women, average number of patients seen by each doctor each day, the average number of visits per inhabitant per year, or the average cost per prescription. However, there were significant differences in the number of packages prescribed per month in the six months before intervention, which was higher in the experimental group compared with the control group. Participation in the group sessions was also evaluated with a deterministic time series analysis. The object of this analysis was to find a function to serve as the analytical representation of a set of data that evolves over time with reference to an imaginary line. Of the functions available, the S or Gompertz curve was selected. Its analytical expression, similar to linear regression, is: \( y = \exp(a + (b/t)) \); where \( a \) is the constant for the model, \( b \) is the slope \( y \) to time. This model best adjusts to the empirical data for the 12 months before intervention, calculated with the \( r^2 \) coefficient of determination and Snedecor’s \( F \) test. A time series adjusted curve of the cost per prescription for each group and its confidence intervals was thus plotted, and the post-intervention changes registered in the groups were projected on this adjusted curve. The reduction in the cost in prescription (cost of average package during the six months after the intervention, minus the cost of average package during the six previous months, divided by the latter) between the periods prior and post-intervention was also calculated for each of the groups, with their respective confidence intervals of 95%.

**Results**
Table 2 shows the basal characteristics of the experimental, placebo, and control groups. There were no significant differences between the groups in terms of the percentage of retirees or women, average number of patients seen by each doctor each day, the average number of visits per inhabitant per year, or the average cost per prescription. However, there were significant differences in the number of packages prescribed per month in the six months before intervention, which was higher in the experimental group compared with the control group. Participation in the group sessions was also evaluated with a deterministic time series analysis. The object of this analysis was to find a function to serve as the analytical representation of a set of data that evolves over time with reference to an imaginary line. Of the functions available, the S or Gompertz curve was selected. Its analytical expression, similar to linear regression, is: \( y = \exp(a + (b/t)) \); where \( a \) is the constant for the model, \( b \) is the slope \( y \) to time. This model best adjusts to the empirical data for the 12 months before intervention, calculated with the \( r^2 \) coefficient of determination and Snedecor’s \( F \) test. A time series adjusted curve of the cost per prescription for each group and its confidence intervals was thus plotted, and the post-intervention changes registered in the groups were projected on this adjusted curve. The reduction in the cost in prescription (cost of average package during the six months after the intervention, minus the cost of average package during the six previous months, divided by the latter) between the periods prior and post-intervention was also calculated for each of the groups, with their respective confidence intervals of 95%.
of 1.76% (95%CI: 0.35% to 3.17%) was observed. In the placebo group. In the control group, a relative increase of 0.16% (95%CI: −0.27% to 2.93%), and 14.1% for the experimental, control and placebo groups respectively. Additionally, the selective change in prescription patterns observed, where there was a decrease in the prescriptions of the drugs that had not been recommended, while volumes for the others either remained stable or increased slightly (though not significantly), cannot be accounted for by evoking the differences detected in the overall basal prescription volumes between the groups, the only variable that showed any differences when basal values were compared.

Other possible threats to the internal validity of a study of this nature. The first is the potential failure of the randomisation process, or other selection biases. Although this cannot be entirely discarded when the number of items to be randomised is small (as in this case, 24 centres into three groups), it is unlikely to have occurred, as no differences were registered—except for one—in the basal variables that affect the volume of prescriptions (table 1), nor were differences registered in the overall relative increase of packages prescribed (~15.8% to ~14.1% and ~14.1% for the experimental, control and placebo groups respectively). Additionally, the selective change in prescription patterns observed, where there was a decrease in the prescriptions of the drugs that had not been recommended, while volumes for the others either remained stable or increased slightly (though not significantly), cannot be accounted for by evoking the differences detected in the overall basal prescription volumes between the groups, the only variable that showed any differences when basal values were compared.

Other possible threats to the internal validity of a study of this nature. The first is the potential failure of the randomisation process, or other selection biases. Although this cannot be entirely discarded when the number of items to be randomised is small (as in this case, 24 centres into three groups), it is unlikely to have occurred, as no differences were registered—except for one—in the basal variables that affect the volume of prescriptions (table 1), nor were differences registered in the overall relative increase of packages prescribed (~15.8% to ~14.1% and ~14.1% for the experimental, control and placebo groups respectively). Additionally, the selective change in prescription patterns observed, where there was a decrease in the prescriptions of the drugs that had not been recommended, while volumes for the others either remained stable or increased slightly (though not significantly), cannot be accounted for by evoking the differences detected in the overall basal prescription volumes between the groups, the only variable that showed any differences when basal values were compared.

Other possible threats to the internal validity of a study of this nature. The first is the potential failure of the randomisation process, or other selection biases. Although this cannot be entirely discarded when the number of items to be randomised is small (as in this case, 24 centres into three groups), it is unlikely to have occurred, as no differences were registered—except for one—in the basal variables that affect the volume of prescriptions (table 1), nor were differences registered in the overall relative increase of packages prescribed (~15.8% to ~14.1% and ~14.1% for the experimental, control and placebo groups respectively). Additionally, the selective change in prescription patterns observed, where there was a decrease in the prescriptions of the drugs that had not been recommended, while volumes for the others either remained stable or increased slightly (though not significantly), cannot be accounted for by evoking the differences detected in the overall basal prescription volumes between the groups, the only variable that showed any differences when basal values were compared.
the three groups, and which did not conduct any other differential strategies during the period in question. Regression to the mean pre-intervention score would not explain changes in the type of active drugs prescribed and, furthermore, would be minimal in a time series adjusted as a moving average. Possible misspecification of the results of the intervention was minimal, as the information system loses less than 4.5% of the prescriptions for the whole healthcare area.

The conditions under which the intervention was conducted made it highly unlikely that the participants might have suspected which of the three groups they were in, and virtually rule out any other form of contamination. The healthcare centres are spread far apart, all the sessions were conducted in a short period of time and the entire healthcare area had already received all manner of materials from the pharmacy unit of the primary care directorate: all the doctors had been sent individualised information about their pharmacological consumption and expenditures, and had participated in other pharmacotherapeutical sessions or had received pharmacology bulletins, etc. This made it improbable that the experimental or placebo sessions could have been perceived as exceptional situations, which would lead the participants to suspect that they were immersed in an intervention trial. For this reason as well, “novelty” or “social desirability” effects would not constitute alternative explanations for the results obtained. Finally, although the person who conducted the sessions was not blind to the randomisation of the groups, she had not participated in the systematic review of the evidence and was only given sufficient information and training to allow her to conduct the evidence based sessions once she had completed the intervention in the placebo group.

Limitations affecting the analysis of the data begin with the study’s low statistical power (n=24 primary care teams distributed into three groups) to detect differences in the reduction in the percentage of prescriptions of the non-recommended drugs. Had the samples been larger, the differences detected between the experimental (−25%) and the placebo (−13%) groups, and the placebo and the control group (−1%) might have been significant. When analysing the cost per prescription, because of the seasonal nature of the use of the NSAIDs, a time series was chosen. More complex models, such as a random effects regression, may have been more suitable. However, as this variable was a secondary outcome, and the differences detected in the simple analysis were irrelevant, the more sophisticated analysis tools were considered unnecessary.

It is true that when evaluating the results of educational interventions it is difficult to isolate the effect of the intervention from other variables in the environment and the context of the study. Factors affecting outcome range from the choice and framing of the issue, to the way in which the visit is conducted, the critical analysis of the evidence, the preparation of materials for dissemination, empathy towards the session leader, and the overall climate within the organisation itself. None the less, the findings of this study indicate that the evidence based outreach session was more effective than no intervention at all, and provide support for the inclusion of evidence based interventions as part of the primary care doctors’ routine, even in environments, such as the Spanish National Health System, where there are very few incentives, economic or otherwise, for changing behaviour.

ACKNOWLEDGEMENTS
Thanks to Soledad Márquez and Ricard Meneu for their comments and suggestions.

Contributors
EBD, MGM, and FPA were responsible for the conception and design of the study. EBD was responsible for the systematic revision, and MGM conducted the educational visits. EBD and SP were responsible for the analysis and interpretation of the data. All the authors contributed to writing the article and revising it for important intellectual content. All approved the final version.

APPENDIX 1 SEARCH STRATEGY

Search 1 Effectiveness
1 OSTEOARTHRITIS*:ME
2 OSTEOARTHRITIS
3 (#1 or #2)
4 DRUG-THERAPY*:ME
5 THERAPY
6 (DRUG and THERAPY)
7 (OUTCOMES-RESEARCH and (SEE and OUTCOME-ASSESSMENT–HEALTH-CARE--*:ME))
8 OUTCOMES
9 (#3 and #4)
10 (#3 and #8)
11. (#9 or #10)

Search 2 Adverse effects
1 ANTI-INFLAMMATORY-AGENTS-NON-STEROIDAL*:ME
2 ADVERSE-DRUG-REACTION-REPORTING-SYSTEMS*:ME
3 (ADVERSE and EFFECT)
4 (SIDE and EFFECT)
5 SAFETY
6 PEPTIC-ULCER*:ME
7 (PEPTIC and ULCER)
8 (GASTROINTESTINAL and HEMORRHAGE)
9 HEMATEMESIS
10 MELENA
11 (DUODENAL and ULCER)
12 (STOMACH and ULCER)
13 (ESOPHAGITIS and PEPTIC)
14 (GASTROINTESTINAL and INJURY)
15 (GASTROINTESTINAL and BLEEDING)
16 (#1 or #2)
17 (#2 or #3) or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12) or #13) or #14) or #15)
18 (#1 and #17)

APPENDIX 2 CRITERIA TO EVALUATE THE ARTICLES RECOVERED

1 Existence of a research protocol. From 0 to 2 points. Categories: “No protocol”, “Essential parts are missing”, “Detailed protocol”.
2 Existence of search strategy or background. From 0 to 2 points. Categories: “Not specified”, “Insufficient”, “Specified”.
3 List of studies analysed and reasons for inclusion. From 0 to 2 points. Categories: “Not specified”, “List of studies”, “Reasons for inclusion”.
4 List of studies excluded and reasons for exclusion. From 0 to 2 points. Categories: “Not specified”, “List of studies”, “Reasons for exclusion”.
5 Analysis of quality of studies. From 0 to 4 points. Categories: “Qualitative”, “Quantitative”.
6 Comparability of studies. From 0 to 8 points. Categories: “Qualitative analysis of sources of heterogeneity”, “Statistical evaluation of heterogeneity”, “Suitability of the type of analysis”, “Analysis of sensitivity”.
7 Control of biases. From 0 to 8 points. Categories: “Analysis of blind studies”, “Analysis between two reviews and degree of concordance”, “Existence of different search strategies”, “Search for non-published materials”.

www.jech.com
APPENDIX 3 STUDIES INCLUDED AND EXCLUDED AND REASONS FOR EXCLUSION

Studies used for review

Studies excluded from review

Reason for exclusion
The study compares drugs that are infrequently used and compares analgesics to NSAIDs.


Reason for exclusion
The study evaluates a drug that is not available in our environment (etodolac).

.........

Authors’ affiliations
E Bernal-Delgado, F Pradas-Arnal, S Peiró-Moreno, Fundación Instituto de Investigación en Servicios de Salud, Spain
E Bernal-Delgado, M Galeote-Mayor, F Pradas-Arnal, Dirección de Atención Primaria, INSALUD, Teruel, Spain
S Peiró-Moreno, Escuela Valenciana de Estudios para la Salud, Spain

Funding: this study was financed in part with FIS grant 98/0730 from the Fondo de Investigación Sanitaria and complementary financing from the Instituto de Investigación en Servicios de Salud.

Conflicts of interest: none.

REFERENCES