Analysis of individual patient data from clinical trials: epidural morphine for postoperative pain

R. J. Ni Mhuircheartaigh†, R. A. Moore*† and H. J. McQuay

Pain Research and Nuffield Department of Anaesthetics, University of Oxford, John Radcliffe Hospital, Level 6 West Wing, Oxford OX3 9DU, UK

*Corresponding author. E-mail: andrew.moore@pru.ox.ac.uk

Background. Individual patient information from clinical trials is infrequently available, but can provide insights for clinical trials and practice.

Methods. We analysed individual patient information from five randomized trials (913 patients) of i.v. patient-controlled analgesia (IVPCA) plus epidural placebo, morphine sulphate (MS) 5 mg, or extended-release epidural morphine (EREM; DepoDur™) at doses of 5–30 mg, to explore effects of a range of epidural morphine doses. Pain and opioid requirement on first and second postoperative days, dose–response, clinically relevant comparisons of IVPCA without epidural morphine, 5 mg MS, and 10 mg EREM, and relationship between patient rating and other measures were described.

Results. There were three strong findings. Epidural morphine resulted in greater patient satisfaction, despite higher rates of adverse events. Those describing their analgesic medication as ‘very good’ or ‘excellent’ used IVPCA opioid less and had pain scores significantly below the global mean, whereas those describing their medication as ‘poor’ or ‘fair’ had pain scores significantly above the mean. Epidural morphine meant less need for postoperative IVPCA opioid than epidural placebo. The therapeutic gain with EREM was lower pain scores with less IVPCA opioid. Moderate or severe pruritus was more common with IVPCA plus epidural morphine, whatever the formulation, compared with IVPCA plus placebo.

Conclusions. Analysis of individual patient data from high-quality clinical trials provides important insights into characteristics of new agents not immediately apparent from original trials, and also informing clinical practice. Prophylactic epidural morphine provides a better patient experience than IVPCA alone.

Br J Anaesth 2009; 103: 874–81

Keywords: anaesthetic techniques, epidural; analgesia, patient-controlled; analgesia, postoperative; analgesics opioid, morphine; pain, postoperative

Accepted for publication: September 10, 2009

Industry-sponsored, large-scale, clinical trials are designed to answer specific questions, often constrained by regulatory requirements to demonstrate safety and efficacy. Clinical trials may answer questions for regulatory purposes, but miss the point as far as clinical practice is concerned. These studies, however, constitute the greater part of evidence available to guide clinical practice during the early post-approval period, before the emergence of further insights and practical clinical considerations as experience and the literature grow.

Clinical trials published in medical journals suffer from tight word limits. Company clinical trial reports offer a considerably greater opportunity to inspect results of clinical trials than is the case with individual published papers. They are unconstrained by space, often running to many thousands, or even tens of thousands, of pages, and

†Declaration of interest. R.A.M. has acted as a consultant for Flynn Pharma, and R.N.M. has spoken on a voluntary basis at sponsored symposia.
Report information in much greater detail for pain outcomes, adverse events, and discontinuations.\textsuperscript{1–5} Company clinical trial reports usually contain data at the level of the individual patient, not just the population average often found in published reports of the same trial. This is important because in many painful conditions, the distribution of response is highly skewed, including postoperative pain, arthritis, or neuropathic pain; the mode is far from the mean.\textsuperscript{5} Although painful conditions demand analysis at the individual patient level,\textsuperscript{6} it is also true for other medical conditions.\textsuperscript{7}

We were provided with clinical trial data from the development programme of an extended-release epidural morphine sulphate formulation (DepoDur\textsuperscript{TM}). We set out to explore the effects of a wide range of epidural morphine doses not often encountered in clinical practice and the relative merits of pain prevention and rescue analgesia, by using individual patient data from five large, randomized, controlled, predominantly double-blind clinical trials of DepoDur\textsuperscript{TM}.

Epidural administration of morphine is an effective route for an effective drug.\textsuperscript{8–9} In spite of its efficacy, single-injection use is limited by its short duration of action relative to postoperative pain and adverse effects associated with higher doses. DepoDur\textsuperscript{TM} is a lipid foam-based formulation of morphine for epidural administration with an ‘extended release’ pharmacokinetic profile (extended-release epidural morphine (EREM)) designed to extend analgesia from a single injection into the second postoperative day, reducing the need to use an epidural catheter to deliver repeated doses. Pharmaceutical details have been described in detail.\textsuperscript{10} Safety concerns about indwelling epidural catheters\textsuperscript{11–13} may make longer duration epidural morphine from a single injection, an attractive option in some countries. The ASA Task Force on Acute Pain Management has advised caution when using continuous opioid infusion, because drug accumulation may contribute to adverse events.\textsuperscript{14}

**Methods**

Ten randomized trials involving EREM are known, nine performed as part of the development programme and one independently.\textsuperscript{15} Flynn Pharma Ltd provided unrestricted access to individual patient data in the form of PDF files (more than 1 GB of data and more than 10,000 pages) to the Pain Research Group of the University of Oxford. Individual patient data from five large, high-quality, multicentre, randomized, double-blind clinical trials investigated the safety, efficacy, and pharmacokinetics of DepoDur\textsuperscript{TM}. These five trials were conducted between 1998 and 2003 and results from the trials themselves have been published in peer-reviewed journals.\textsuperscript{16–20} We excluded data from four trials. A trial of knee replacement in 168 patients using 20 and 30 mg EREM\textsuperscript{21} used a rescue analgesic regimen [i.v. bolus hydromorphone and i.v. patient-controlled analgesia (IVPCA) morphine] very different from the included five, making pooling unreliable. Another excluded study involved healthy volunteers not in pain, and two others tested effects of lidocaine at various intervals before EREM administration to measure pharmacokinetic changes from chemical interactions between lidocaine and the foam vehicle. Reporting quality of included trials was examined using the Oxford scale.\textsuperscript{22}

Briefly, the five studies included in this analysis were:

(i) a Phase 2, open-label dose finding study in which either EREM or morphine sulphate (MS) 5 mg was administered to 39 patients undergoing total hip arthroplasty;\textsuperscript{19}

(ii) a Phase 2, randomized, double blind, placebo-controlled, dose-finding study of EREM in 120 patients undergoing total hip arthroplasty;\textsuperscript{20}

(iii) a Phase 3, randomized, double-blind, placebo-controlled study of EREM in 194 patients undergoing hip arthroplasty;\textsuperscript{18}

(iv) a Phase 3 randomized, double-blind study of EREM or MS 5 mg in 546 patients undergoing lower abdominal surgery;\textsuperscript{17}

(v) a randomized, double-blind study of EREM or MS 5 mg in 79 patients undergoing elective lower segment Caesarean section (LSCS).\textsuperscript{16}

All patients received IVPCA and also an epidural injection. The three main treatment groups were therefore IVPCA plus placebo (vehicle or saline), IVPCA plus MS 5 mg, and IVPCA plus EREM at 5, 10, 15, 20, 25, or 30 mg doses.

Each of the studies had the same basic design and conduct. Visual analogue scores (VAS) and categorical ratings of pain intensity were recorded at rest and with a standardized procedure-appropriate activity. These were recorded at the time of the patient’s first postoperative request for analgesia and at 2, 4, 8, 12, 18, 24, 30, 36 and 48 h after surgery. Visual analogue pain intensity scores at rest and with activity were incorporated into an individual’s ‘mean pain intensity score’ over the first 24 h, the second 24 h, and the total 48 h of the postoperative period.

All patients were provided with IVPCA containing fentanyl, except in the study of patients undergoing LSCS in which IVPCA morphine was provided. IVPCA was being used as a measure of analgesic efficacy.\textsuperscript{23} Records were kept of subjects who were converted to oral analgesia within the first 48 h or who received other opioid analgesia for any reason. Non-fentanyl opioids were incorporated into the total consumption calculations for analysis as ‘fentanyl equivalents’;\textsuperscript{24} totals at 24 and 48 h were available for each patient.

Patients were asked to rate their pain medication at 48 h as ‘poor’, ‘fair’, ‘good’, ‘very good’, and, in some of the studies, ‘excellent’. Ratings of ‘very good’ and ‘excellent’ were combined. Five specific adverse effects associated with epidural opioids (pruritus, nausea, vomiting, urinary retention, and respiratory depression) were reported, and
their severity rated during each study as none, mild, moderate, or severe.

Ventilatory frequency and end-tidal CO₂ at intervals after operation and haemoglobin oxygen saturation were measured continuously for the first 24 postoperative hours and at intervals thereafter. There was a consistent definition of respiratory depression of sustained saturation <90% on 2 litre min⁻¹ oxygen via a nasal cannula, or end-tidal CO₂ >50 mm Hg for two consecutive measurements, but no reported standardized criteria for severity. In addition, interventions for respiratory depression, like increased oxygen administration or use of opioid antagonists on these observations, were not standardized across institutions, and there was widespread use of opioid antagonists for the treatment of pruritus.

We extracted various data into a database for analysis, including patient characteristics (age, sex, and BMI), type of surgery, type of anaesthesia (general or subarachnoid), treatment assignment, occurrence and severity of common adverse events, pain intensity scores at 10 postoperative time points, patient ratings of their pain medication 48 h after operation, and consumption of rescue analgesia during those first 48 h. Adverse event severity was taken as most severe at any time over the 48 h period.

Analysis

Our analyses included the following.

(i) Beneficial and harmful effects on the first and second postoperative days, separately and together; specifically pain intensity scores, postoperative opioid use, and adverse events, with regard to:
   (a) dose of EREM
   (b) clinically relevant comparisons: although it is interesting to see effects of higher doses, it is more important to compare clinically relevant treatments, in this case IVPCA plus placebo epidural, IVPCA plus epidural MS 5 mg, and IVPCA plus EREM 10 mg.
   (ii) Patient rating of satisfaction with pain medication.
   (iii) The relationship between patient rating and analgesic consumption, pain scores, and adverse events.

The effect of dose on pain intensity scores was assessed with an analysis of variance, as were normally distributed patient characteristic variables. Tukey’s method for multiple pairwise comparisons (with a family-wise error rate=0.05) was used to test for significant differences between doses. Rescue analgesia use was asymmetrically distributed and was analysed using a Kruskal–Wallis test for equality of medians with post hoc comparisons of clinically relevant doses carried out using Mann–Whitney U-tests with a Bonferroni multiple comparisons correction. The incidence of adverse effects in each treatment group was tested using χ² tests. Post hoc analyses specifically investigated differences in clinically relevant treatment groups.

Patient ratings of pain medication were used to explore differences in the experiences (pain intensity scores, rescue analgesia use, and adverse effect) of patients and the possible reasons underlying those differences. Ratings were converted into binary outcomes with ‘poor’ and ‘fair’ regarded as unsuccessful pain management, and ‘good’, ‘very good’, or ‘excellent’ regarded as successful pain management. A binary logistic regression analysis was carried out with factors for the treatment group, surgery and anaesthesia type as factors, and pain intensity scores, rescue analgesia use, and adverse effect severity as covariates. Statistical analysis was carried out using Minitab® Statistical Software (version 15.1.1, Minitab Inc., State College, PA, USA, 2007).

Results

All five trials were randomized, with information about withdrawals; four were double blind. Quality scores were 3 out of a maximum of 5 points in four trials, indicating little risk of bias; the single open study (quality score 2/5) contained only 4% of the patients.

Results were available for 970 patients. Fifty-seven of these patients were excluded from our analysis for a number of reasons, including failure to administer the assigned study drug (dural puncture, failed epidural placement, or a change in anaesthesia or surgical plan unrelated to the study), patient withdrawal of consent before study drug administration, or failure to record analgesia consumption. Overall, 913 (94%) of randomized patients were included in the per-protocol analysis.

All 913 patients had IVPCA available and 92% had received epidural morphine at some dose and formulation. IVPCA plus epidural was given to 75 patients, 95 had IVPCA plus MS 5 mg, and 742 had IVPCA plus EREM in doses ranging from 5 to 30 mg [5 mg (n=120), 10 mg (n=127), 15 mg (n=152), 20 mg (n=168), 25 mg (n=129), and 30 mg (n=47)]. Among these patients, 474 had abdomino-pelvic surgery (colectomy, total abdominal hysterectomy, or radical prostatectomy), 366 had hip surgery (total hip arthroplasty, primary, or revision), and 73 underwent LSCS.

The mean (range) age was 52 (18–99) yr; 53% were women. There were significant differences in the treatment group ages, arising from the populations of the component studies. Hip surgery studies compared IVPCA plus EREM with IVPCA plus placebo and tended to involve older subjects than abdominal surgery and LSCS which used MS 5 mg as a control rather than placebo. There was no significant difference in ages in the clinically relevant comparisons.

Pain intensity scores

Patients who received IVPCA plus any dose of epidural morphine had significantly lower mean 48 h VAS pain
intensity scores at rest than those with IVPCA plus epidural placebo (Fig. 1). There was no significant difference in VAS with activity between the placebo group and those who received either MS 5 mg or EREM 5 mg, but those who received at least EREM 10 mg had VAS scores with activity that were significantly lower than those in the placebo group.

Significant differences were limited to the first postoperative day; there was no significant difference in pain intensity scores on the second postoperative day, although average scores at rest on the second postoperative day were low, at ~20 mm (range 0–100 mm) (Fig. 1c). IVPCA plus higher doses of EREM tended to produce lower pain intensity scores. IVPCA plus MS 5 mg and IVPCA plus EREM 10 mg were significantly better than placebo at rest and with activity on the first but not on the second postoperative day. IVPCA plus EREM 10 mg was significantly better than IVPCA plus MS 5 mg for VAS with activity on the first postoperative day and overall.

Postoperative opioid use

Patients in each of the epidural morphine treatment groups used significantly less rescue analgesia than those in the placebo group (Fig. 2) on both the first and second postoperative days and overall. IVPCA plus higher doses of EREM tended to result in less postoperative IVPCA opioid consumption. Postoperative IVPCA opioid use was significantly lower on the second but not on the first postoperative day in the IVPCA plus EREM 10 mg group than in the IVPCA plus MS 5 mg group (Fig. 2).

Adverse events

Patients reported their subjective experience of pruritus, nausea, and vomiting at each rating time point. The assessment of urinary retention and respiratory depression was limited by inconsistencies across studies in definition and management. Detection of urinary retention was confounded by varied policies on postoperative urinary catheter removal. In some centres, opioid antagonists were used as a first-line treatment of pruritus, potentially masking respiratory depression.

Table 1 shows the incidence of five common adverse events (pruritus, nausea, vomiting, respiratory depression, and urinary retention) over 48 h by severity across all patients with recordings. Adverse events were predominantly absent or mild, and severe in fewer than 2% of
patients. Only for nausea and pruritus were adverse events moderate in >10% of patients.

Figure 3 shows the incidence by treatment group and severity. There was no significant difference in the incidence of moderate and severe nausea and vomiting between the placebo group and those receiving IVPCA plus MS 5 mg or IVPCA plus EREM below 25 mg. There was a tendency for more moderate or severe cases of respiratory depression or urinary retention with higher IVPCA plus EREM doses, but differences between studies precluded any formal dose–response analysis.

Table 1 Severity of different adverse events across all treatment groups

<table>
<thead>
<tr>
<th>Adverse event</th>
<th>Number of patients</th>
<th>Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>None</td>
<td>Mild</td>
</tr>
<tr>
<td>Pruritus</td>
<td>906</td>
<td>49</td>
</tr>
<tr>
<td>Nausea</td>
<td>905</td>
<td>37</td>
</tr>
<tr>
<td>Vomiting</td>
<td>905</td>
<td>70</td>
</tr>
<tr>
<td>Urinary retention</td>
<td>905</td>
<td>88</td>
</tr>
<tr>
<td>Respiratory depression</td>
<td>902</td>
<td>85</td>
</tr>
</tbody>
</table>

Fig 2 (A) Each individual’s rescue analgesia use over 48 h analysed by the treatment group. (b) Comparison for clinically relevant groups on first and second postoperative days. Boxplots represent the IQR with whiskers extending 1.5 IQR from the box. Medians are shown as a horizontal line and means as a diamond. *P<0.05.
Patient rating of pain medication

Most patients rated their pain medication as at least ‘good’. We arbitrarily divided patients into those describing their postoperative medication as ‘poor’ or ‘fair’ (the ‘dissatisfied’ group) and those who described it as ‘good’, ‘very good’, or ‘excellent’ (the ‘satisfied group’). At the end of the first 48 postoperative hours, the proportion of patients satisfied with their pain medication was 85% overall, with the lowest rating with placebo (77%) and the highest with IVPCA plus EREM 25 mg (93%); 80%+ satisfaction rates were achieved with all but IVPCA plus placebo and IVPCA plus EREM 5 mg. The percentage of dissatisfied patients decreased as epidural morphine dose increased (Pearson’s $\chi^2=17.4$, $P<0.001$).

Logistic regression analysis identified only VAS at rest and activity and total postoperative IVPCA consumption as significant predictors of patient satisfaction (80% concordant pairs). More severe pain and greater postoperative opioid use were associated with less satisfaction with pain medication (Fig. 4). Most of those rating pain medication very good or excellent had no or mild pain at rest, and moderate pain or less on activity. Even with a ‘good’ rating, the pain experience included moderate pain at rest and moderate or severe pain on activity. Patients rating their pain medication as very good or excellent consumed half to one-third of postoperative opioid compared with those with lesser ratings (Fig. 4).

Dissatisfied patients used significantly more IVPCA (median fentanyl equivalent 1305 µg) than the satisfied patients (median fentanyl equivalent 550 µg) ($P<0.0001$, Mann–Whitney $U$-test). Dissatisfied patients also reported significantly higher pain intensity scores at rest (mean 35 vs 21; $P<0.001$, unpaired $t$-test) and with activity (mean 52 vs 33; $P<0.001$). There was no significant difference in the prevalence of moderate or severe pruritus, nausea, vomiting, respiratory depression, or urinary retention between satisfied and dissatisfied patients.

Discussion

The five trials on which this analysis is based are unusual, in that all 913 patients received both IVPCA and an epidural injection, with MS at various doses and in various formulations in 92% of them, and epidural placebo in the remainder. IVPCA was used as a measure of the analgesic efficacy of epidural morphine, with lower postoperative IVPCA opioid requirement indicating greater analgesic effect from the epidural morphine. The joint use of these two analgesic strategies would be unlikely in clinical practice, but what we have are data that potentially illuminate aspects of postoperative pain care.

Pain (at rest and on activity) and adverse events were monitored 10 times over 48 postoperative hours, with
additional information about postoperative analgesic requirement and other outcomes like patient satisfaction, providing tens of thousands of observations at the level of the individual patient. This allowed correlation between metrics rather than population level analysis which allows only for descriptive associations. To have this quantity and quality of information is unusual in trials in the postoperative period, although analyses from clinical trial reports and at the level of the individual patient have been used before in painful conditions.\textsuperscript{1, 5, 6}

There were three strong findings.

Epidural morphine resulted in greater patient satisfaction, despite higher rates of adverse events. Those describing their analgesic medication as ‘very good’ or ‘excellent’ had IVPCA opioid use and pain intensity scores significantly below the global mean, whereas those describing their medication as ‘poor’ or ‘fair’ were significantly above.

Epidural morphine meant less need for postoperative IVPCA opioid than epidural placebo, with a dose–response. This was predictable, but the relationship between analgesic requirement and pain intensity is complicated. For example, IVPCA plus epidural morphine treatment was better than IVPCA plus placebo in the first postoperative day, with both lower pain intensity scores and lower opioid requirements. On the other hand, pain on activity was lower with IVPCA plus EREM 10 mg than IVPCA plus MS 5 mg on the first postoperative day, and on the second postoperative day, it was better in terms of lower opioid requirement, although with no difference in pain intensity score. The therapeutic gain with EREM was lower pain intensity scores with less IVPCA opioid.

Perioperative epidural morphine can produce more itch. Moderate or severe pruritus was over twice as common with IVPCA plus epidural morphine, whatever the formulation, compared with IVPCA plus placebo, but there was no significant difference in the frequency of moderate or severe nausea or vomiting between any of the treatment groups.

This was unanticipated. IVPCA would be expected to lead to the achievement of an adequate level of analgesia, with patients titrating to the same level of comfort unless there were features that prevented it, like small boluses, or long lock out times. In these trials, boluses were sensible and lock out times short: those with higher postoperative IVPCA opioid use on day 1 had higher pain intensity scores and never got on top of their pain. We cannot, however, exclude the possibility that IVPCA as an analgesic measure failed because it lacks sensitivity, unable to differentiate treatments of different efficacy. As has been pointed out before,\textsuperscript{25} there is limited evidence about how robust analgesic consumption designs are, especially those using PCA devices. Robustness of the system is assumed, but to our knowledge, it has not been rigorously tested in the prophylactic or perioperative analgesic setting. If the analgesic consumption design is less robust, then misleading conclusions may be made about the true efficacy of interventions such as prophylactic analgesia. Adverse events, from epidural morphine, such as sedation or nausea, could also result in less IVPCA drug consumption, leading to apparently better analgesic effect.

In these trials, each patient was questioned 10 times about adverse events. This frequent questioning may make it important to distinguish between adverse events of any severity (mild, moderate, or severe) and those involving intervention or patient distress (moderate or severe). Inclusion of reports of ‘mild’ adverse effects suggested that these adverse events occurred with high frequency. Symptoms of greater severity were found to occur less frequently, but to different degrees in individual adverse events.

We defined three treatments as relevant to clinical practice, IVPCA plus epidural placebo, IVPCA plus epidural MS 5 mg, and IVPCA plus EREM 10 mg. Both doses of epidural morphine were associated with lower pain scores compared with IVPCA/placebo, and EREM 10 mg had some advantages over MS 5 mg. One-third of patients receiving EREM had doses of 5 mg or 10 mg, whereas two-thirds had doses $>10$ mg, dosing levels unlikely to be used in clinical practice.

In conclusion, prophylactic analgesia leads to a more satisfactory patient experience than IVPCA. The logic of IVPCA use means that patients must by definition experience pain to access pain relief. Adequate pain prevention predicts patient satisfaction while the occurrence of unpleasant yet non-life threatening adverse events does not. Epidural morphine plus IVPCA produced superior analgesia than IVPCA alone, providing superior analgesia extending into the second postoperative day. Analysis of individual patient data from high-quality clinical trials provides important insights into the expected characteristics of new agents not immediately apparent from the original trials, and also informing thinking about improved clinical practice.

Acknowledgement

We are grateful to Flynn Pharma for making the individual patient data available.

Funding

No specific funding was received for this work. R.A.M. is funded by NIHR Biomedical Research Centre Programme.

References


7 Christakis NA. Does this drug work for you? Br Med J 2008; 337: a2281


13 Ruppen W, Derry S, McQuay HJ, Moore RA. Infection rates associated with epidural indwelling catheters for seven days or longer: systematic review and meta-analysis. BMC Palliat Care 2007; 6: 3


