Baseline morphine consumption may explain between-study heterogeneity in meta-analyses of adjuvant analgesics and improve precision and accuracy of effect estimates

Brett Doleman MBBS,¹ Alex J Sutton PhD,² Matthew Sherwin BMBS,¹ Jonathan N Lund MD,¹ John P Williams PhD.¹

¹Department of Surgery and Anaesthesia
Division of Medical Sciences and Graduate Entry Medicine
University of Nottingham
Derby, UK
DE22 3NE

²Department of Health Sciences
College of Medicine
University of Leicester
Leicester, UK
LE1 7HA

Correspondence to: Dr Brett Doleman
University of Nottingham
Royal Derby Hospital
Uttoxeter Road
Derby
DE22 3DT
Telephone: 01332 724641
Fax: 01332 724697
Email: dr.doleman@gmail.com

Financial Disclosures: None

Conflicts of Interests: None (all authors)

Review registration: CRD42016039109 (PROSPERO)

Word count: 376 (Abstract), 531 (Introduction) and 1977 (Discussion)

Abbreviated title: Meta-regression of analgesics

Author Contributions:
Brett Doleman: conceived the review and developed methods, data analysis, writing manuscript and approving final version.
Alex Sutton: data analysis, editing manuscript and approving final version.
Matthew Sherwin: data collection, editing manuscript and approving final version.
Jonathan Lund: data analysis, editing manuscript and approving final version.
John Williams: data analysis, editing manuscript and approving final version.
Abstract

Introduction: Statistical heterogeneity can increase the uncertainty of results and reduce the quality of evidence derived from systematic reviews. At present, it is uncertain what are the major factors that account for heterogeneity in meta-analyses of analgesic adjuncts. Therefore, the aim of this review was to identify whether various covariates could explain statistical heterogeneity and use this to improve accuracy when reporting the efficacy of analgesics.

Methods: We searched for reviews using MEDLINE, EMBASE, CINAHL, AMED and the Cochrane Database of Systematic Reviews. Firstly, we identified the existence of considerable statistical heterogeneity ($I^2>75\%$). Secondly, we conducted meta-regression analysis for the outcome of 24-hour morphine consumption using baseline risk (control group morphine consumption) and other clinical and methodological covariates. Finally, we constructed a league table of adjuvant analgesics using a novel method of reporting effect estimates assuming a fixed consumption of 50mg of postoperative morphine.

Results: We included 344 randomized controlled trials with 28,130 participants. 91% of analyses showed considerable statistical heterogeneity. Baseline risk was a significant cause of between-study heterogeneity for acetaminophen, NSAIDS and COX-2 inhibitors, tramadol, ketamine, alpha-2 agonists, gabapentin, pregabalin, lidocaine, magnesium and dexamethasone ($R^2 \ 21-100\%$; $p<0.05$). There was some evidence that methodological limitations of the trials explained some of the residual
heterogeneity. Type of surgery was not independently associated with analgesic efficacy. Assuming a fixed baseline risk of 50mg (in order of efficacy), gabapentin, acetaminophen, alpha-2 agonists, NSAIDS and COX-2 inhibitors, pregabalin, tramadol, magnesium and lidocaine demonstrated moderate clinically significant reductions (>10mg). We could not exclude a moderate clinically significant effect with ketamine. Dexamethasone demonstrated a small clinical benefit (>5mg).

**Discussion:** We empirically identified baseline morphine consumption as the major source of heterogeneity in meta-analyses of adjuvant analgesics across all surgical interventions. Controlling for baseline morphine consumption, clinicians can use audit data to estimate the morphine reducing effect of adding any adjuvant for their local population, regardless which surgery they undergo. Moreover, we have utilized these findings to present a novel method of reporting and an amended method of graphically displaying effect estimates, which both reduces confounding from variable baseline risk in included trials and is able to adjust for other clinical and methodological confounding variables. We recommend use of these methods in clinical practice and future reviews of analgesics for postoperative pain.
**Introduction**

Meta-analyses have emerged as a useful method to summarize research findings and increase the statistical power of primary research studies. However, one of the major limitations of this form of analysis is the aggregation of trials conducted in both different populations and in different clinical circumstances. This is termed clinical heterogeneity. Such clinical heterogeneity, along with other methodological limitations, may give rise to statistical heterogeneity. Statistical heterogeneity occurs when estimates from individual trials vary by more than would be expected by chance which can be quantified using measures such as the $I^2$ statistic (0-100%).

Unexplained statistical heterogeneity can increase the uncertainty surrounding effect estimates derived from meta-analyses and reduce the quality of evidence used to inform healthcare decisions. In addition, in the presence of statistical heterogeneity, effect estimates may be inaccurate and lead to erroneous conclusions on the clinical significance of a particular intervention. Therefore, investigating causes for heterogeneity is essential using techniques such as meta-regression analysis. Baseline risk is a particular covariate that can help predict between-study heterogeneity in meta-analyses. For example, participants at higher risk of the condition under study (those with more pain) may derive larger absolute benefits compared to those at lower risk. However, conventional meta-regression analyses of baseline risk may be biased due to measurement error in baseline risk (x variable) and effect estimates (y variable) and the functional relationship between the two. This causes regression to the mean, where relationships may become apparent even if none
exist.\textsuperscript{4,5} Therefore, alternative analyses, which allow for the measurement error, such as previously developed Bayesian meta-regression models are recommended.\textsuperscript{6}

Heterogeneity is a particular problem in meta-analyses of analgesics used to prevent postoperative pain.\textsuperscript{7} Indeed, a previous review has suggested that type of surgery should be explored in these reviews.\textsuperscript{7} However, even within the same type of surgical procedure, pain levels and morphine consumption can be heterogeneous. In addition, differing analgesic protocols can further confound the association between type of surgery and the efficacy of the analgesic. Previous primary research has shown that the pain level experienced by a participant determines analgesic efficacy, with higher pain levels resulting in higher absolute pain score reductions following analgesic administration.\textsuperscript{8,9} We have previously demonstrated that using control group morphine consumption (baseline risk), we were able to explain a large degree of between-study heterogeneity,\textsuperscript{10,11} which may improve the quality of evidence as per Grades of Recommendation, Assessment, Development and Evaluation Group (GRADE).\textsuperscript{2} Moreover, if baseline risk is found to explain statistical heterogeneity, any one effect estimate from a meta-analysis may be confounded by the variable baseline risk in the included trials and therefore quoting meta-regression estimates from a fixed value may help reduce confounding.

Therefore, our \textit{a priori} hypothesis was that baseline risk and other methodological or clinical covariates could explain statistical heterogeneity in meta-analyses of analgesics. We then aimed to utilize these findings to report a novel method of
presenting effect estimates using meta-regression parameter estimates, which reduce confounding from variable baseline risk in the included studies. In addition, we used this novel method to produce a league table of adjuvant analgesics based on efficacy and illustrate how our novel method can better target clinical use of analgesics using local departmental audit data.
Methods

We prospectively registered this review on the PROSPERO\textsuperscript{12} website using the registration number CRD42016039109. Due to the numerous previous systematic reviews published on the subject, the aim of this study was to search for previous reviews of postoperative analgesics. We then identified randomized controlled trials from all included reviews and performed a secondary analysis of these (Figure 1). We searched all databases from inception to May 2016: MEDLINE, EMBASE, CINAHL, AMED and the Cochrane Database of Systematic Reviews. We used the following search terms: ‘postoperative AND pain’, ‘surgery’, ‘analgesi*', ‘morphine AND consumption’, ‘opioid AND consumption’ and we exploded the MeSH term ‘ACUTE PAIN’. We combined these terms with the specific generic term for the analgesic agent. We then limited our search to reviews and meta-analyses.

We extracted the data onto an electronic database. For reviews we extracted the following data: study author, year of publication, type of agent and methods for investigating heterogeneity. For randomized controlled trials included in the reviews, we extracted the following data: study author, year of publication, population, intervention, comparator, opioid used, risk of bias information and data used to calculate effect estimates. For randomized controlled trials, in order to reduce selective reporting bias, if standard deviations were not reported, we estimated these from other studies in the analysis.\textsuperscript{13} We did not attempt to estimate means and standard deviations from medians or inter-quartile ranges due to the high likelihood of
non-normal data. If results were not reported in the text, these were estimated from published graphs.

We had no language restrictions for inclusion in our review and we translated non-English language papers. We included reviews that included the following analgesic agents versus placebo for postoperative pain: acetaminophen, non-steroidal anti-inflammatory drugs (NSAIDS) and cyclooxygenase (COX) 2 inhibitors, tramadol, intravenous ketamine, alpha-2 agonists (clonidine and dexmedetomidine), gabapentin, pregabalin, nefopam, lidocaine, magnesium and dexamethasone. We aimed to identify reviews of prophylactic administration (defined as first dose given before the onset of pain or agents added to postoperative analgesic regimens, such as patient-controlled analgesia). We did not include reviews evaluating single dose analgesics for established postoperative pain or reviews in dental surgery, as these are unlikely to report 24-hour morphine consumption.

The outcome of interest was 24-hour opioid consumption. We chose opioid consumption as this serves as a surrogate measure for both how painful the procedure was and any concurrent analgesia used. In addition, as participants within these trials can use variable amounts of morphine to achieve a desired level of comfort, it may be more appropriate than pain score data, which may be confounded by variable morphine use between the groups. Moreover, one of the main goals of multimodal analgesia is to reduce opioid consumption. We only included primary studies where we could extract morphine consumption data. If studies reported dosage per kilogram,
we converted this to a 70-kilogram weight. We also used data from the day of surgery or postoperative day one and analysed this as 24-hour data. If alternative opioids were reported, these were converted to morphine equivalents using the following conversion factors: oral to intravenous morphine (3:1), pethidine/meperidine (10:1), ketobemidone (1:1), tramadol (20:1), fentanyl (1:100), remifentanil (1:100), piritramide (1:0.75), hydromorphone (1:3), oral hydrocodone (2:1), intravenous oxycodone (1:1.5), oral oxycodone (2.5:1), papaveretum (1.5:1), meptazinol (5:1), nalbuphine (1:1), propoxyphene (10:1), sublingual buprenorphine (1:25) and trimeperidine (2:1).

We undertook assessment of randomized controlled trials from included reviews using the Cochrane risk of bias tool. For blinding to receive low risk, studies had to describe in enough detail study drugs and placebos that were identical or similar in appearance rather than simply describe the study as ‘double-blind’. Outcome assessment also needed to be blinded. Attrition bias would receive high risk if patients were excluded from the analysis for reasons that may influence opioid consumption, such as those with uncontrolled pain or potential opioid adverse effects. If this information was unclear, we attempted to contact authors for further information. Studies only received low risk for selective outcome reporting if outcomes were pre-stated in a published protocol or trial registration referenced in the included study. Other bias included baseline characteristic imbalances which have been associated with influencing pain (for example gender and pre-operative pain) or industry sponsorship.
Statistical Analysis

To quantify the degree of statistical heterogeneity we used the $I^2$ statistic, with values exceeding 75% as evidence of considerable heterogeneity and those exceeding 50% as evidence of moderate statistical heterogeneity.\(^1\) For the available data, we calculated the mean difference (MD) in morphine consumption (mg) with 95% confidence intervals (CI) using a random-effects model. In order to identify whether control group morphine consumption could explain the between-study heterogeneity we undertook meta-regression analysis.\(^3\) This analysis is similar to conventional regression analysis, although it involves using study-level covariates, such as the dose of the analgesic used in the trial as the predictor variable and the effect estimate (MD) as the outcome variable, with each study weighted for the precision of the results (lower standard errors having more weight) (Table 1).

We performed conventional (frequentist) meta-regression initially using control group morphine consumption (baseline risk) as a covariate based on previous findings.\(^10\) Due to the problems with analyzing baseline risk using frequentist meta-regression, we additionally undertook Bayesian meta-regression using Markov Chain Monte Carlo (MCMC) with Gibbs sampling following recently developed methodology that incorporates the uncertainty of the covariate estimates, which avoids the problems of regression to the mean (Supplementary Text Document 1).\(^6\) We used vague prior distributions and burnt in the MCMC chains for 10,000 iterations and then used a sample of 50,000 iterations on which to base inferences. We checked convergence visually by looking at history plots of the sampled values. We present the results of
We also used the following covariates, which we added to the frequentist model with baseline risk and present as the R² change from the univariate analysis using baseline risk only (Table 2). Clinical covariates included dose or route of drug administration, type of agent (NSAIDS versus COX-2 inhibitor for example), type of surgery and type of anesthesia. For type of surgery, where possible, we aimed to include procedure-specific evidence, if this was not possible we grouped procedures by specialty or anatomical location. In addition, we assessed whether measures of internal validity were responsible for statistical heterogeneity including: randomization, allocation concealment, blinding and attrition bias. Except for attrition bias, these covariates were only included in models if they exaggerated effect estimates. Control group morphine consumption was initially added to the model, we then added other covariates to a multivariate model to adjust regression estimates for these confounding variables if they significantly improved the model, in a stepwise approach (p<0.1 for retention in the model). Studies with missing data were excluded from the analysis.

For conventional meta-regression, we used a restricted maximum likelihood, random-effects model. We also used the Knapp-Hartung modification as the variance estimator for effect estimates. We assessed linearity and heteroscedasticity from predicted versus residual plots and we assessed residuals for normality using
histograms. We assessed outliers from studentized residual values and leverage using Cook’s distance (with values greater than one regarded as a cause for concern). We present results as the proportion of between-study heterogeneity explained by the model ($R^2$ analogue) with a corresponding p value. We undertook sensitivity analysis removing studies that had significant leverage on the model and excluding studies that measured opioid consumption at other time points. We regarded p values for final models <0.005 as statistically significant following Sidak adjustment for multiple comparisons.

If we identified baseline risk as a significant cause of between-study heterogeneity, we produced a league table of analgesic adjuncts based on a fixed control group consumption of 50mg using Bayesian parameter estimates, which aims to reduce confounding from variable baseline risk in the included trials. In addition, we constructed amended scatter plots as an alternative graphical method to Forest plots in presenting effect estimates (Figure 2). These plots demonstrate the relationship between baseline risk and efficacy (red regression line). These plots also show the effect estimate from a fixed consumption of 50mg of morphine (where solid horizontal red line meets the Y-axis) with the associated 95% confidence interval (dashed red lines on the Y-axis) (STATA code used to calculate effect estimates and produce amended plots in Supplementary Text Document 2).

For results from a fixed consumption of 50mg of morphine, we regarded a difference of >20mg as a large clinically significant difference, >10mg a moderate clinically
significant difference and >5mg of small clinical significance. We were unaware of any literature regarding clinically significant reductions in morphine consumption. Therefore, we selected these values based on two (20mg) and one (10mg) standard dose(s) of morphine. This analysis allows comparison of analgesic adjuncts when adjusted for the variable control group morphine consumption from the included randomized controlled trials in order to reduce confounding. However, we ranked agents based on the point estimate and did not incorporate the uncertainty around these into these ranks and therefore these should be interpreted with caution. Where dose or route of administration was found to be a significant predictor, we included results from the most effective clinical situation and specified this where appropriate (for adjusted conventional estimates). We present both Bayesian parameter estimates (median) and adjusted conventional estimates with 95% CIs/CrIs.

The meta-regression parameter estimates presented in Table 1 can be used to estimate the likely reduction in 24-hour morphine consumption based on average consumption data from local departmental audits using the following formula:

\[
\text{Mean reduction in 24-hour morphine (mg) = intercept + beta coefficient x mean 24-hour consumption from departmental audit}
\]

However, it should be noted that at extremes of baseline risk, confidence intervals are wider and therefore the uncertainty of the estimate increases. In addition, extrapolating morphine reductions outside the limits of our models may be invalid.
We did not undertake assessment for publication bias due to recent findings from our group that traditional methods may be inaccurate. We conducted all analyses using Comprehensive Meta-analysis Version 3,\textsuperscript{31} STATA Version 14\textsuperscript{32} and WinBUGS Version 1.4.\textsuperscript{33}
Results

We included 344 randomized controlled trials with 28,130 participants (Table 1 and Supplementary Table 1). We identified these studies from 8 narrative reviews,34-41 25 systematic reviews42-66 and 72 meta-analyses10-11, 67-136 (Figure 1). Of the included reviews that conducted a meta-analysis, 78% investigated heterogeneity. In 75%, investigation of heterogeneity was conducted using subgroup or sensitivity analysis and only 18% conducted meta-regression. In 32% of meta-analyses, investigation of heterogeneity was based on type of surgery, 35% used dose and 11% used type of anesthesia. In 31% of meta-analyses, heterogeneity was investigated using methodological covariates. On risk of bias assessment of the individual randomized controlled trials, adequate randomization was described in 58% of studies, adequate allocation concealment in 30%, adequate blinding in 51% and lack of attrition bias in 73%.

From the included randomized controlled trials, there was evidence of considerable statistical heterogeneity ($I^2 > 75\%$) in most analyses (91%). On meta-regression analysis (Table 1), control group morphine consumption (baseline risk) explained between-study heterogeneity for acetaminophen, NSAIDS and COX-2 inhibitors, tramadol, ketamine, alpha-2 agonists, gabapentin, pregabalin, lidocaine, magnesium and dexamethasone (Figure 2). We could not analyze nefopam as we only identified five studies. When re-analyzed using Bayesian meta-regression, control group morphine consumption remained a significant cause of heterogeneity and parameter estimates were very similar (Table 1). Mean control group consumption in each meta-
analysis varied between 28mg to 50mg (Table 1). Sensitivity analyses did not change parameter estimates.

Other significant causes of between-study heterogeneity when added to the univariate model (Table 2 and 3) included route of administration and allocation concealment for acetaminophen ($R^2=94\%$; $p<0.001$). Intravenous acetaminophen was more effective than other routes ($-4mg; 95\% CI -7.55mg to -0.42mg$). For ketamine, the final model included blinding and allocation concealment, which explained the majority of the between-study heterogeneity ($R^2=56\%$; $p<0.001$). For alpha-2 agonists, the addition of attrition bias and route of administration significantly improved the model, with intravenous ($-5.4mg; 95\% CI -11.68mg to 0.89mg$) and epidural/spinal ($-4.2mg; 95\% CI -10.89mg to 2.48mg$) administration (omnibus $p=0.07$) the most effective ($R^2=75\%$; $p<0.001$). The gabapentin model was improved by the addition of peri-operative dose ($R^2=93\%$; $p<0.001$). For pregabalin, the final model included allocation concealment, which significantly improved the model ($R^2=78\%$; $p<0.001$). For lidocaine, the final model included route of administration and attrition bias ($R^2=87\%$; $p<0.001$). Intravenous administration was more effective than subcutaneous patch ($-6.7mg; 95\% CI -12.62mg to -0.87mg$). For magnesium, the addition of allocation concealment significantly improved the final model ($R^2=31\%$; $p=0.008$). We did not include dose, as this did not exaggerate effect estimates. Dexamethasone was the only analysis where type of surgery was a significant predictor. The final model included type of surgery and blinding ($R^2=100\%$; $p<0.001$), with larger morphine reductions in spinal and ENT surgery (although only
based on single studies). However, analysis could not performed with type of surgery and allocation concealment due to multicollinearity.

When assuming a fixed consumption of 50mg of postoperative morphine, we observed moderate clinically significant reductions (in order of efficacy; Figure 3) with gabapentin, acetaminophen, alpha-2 agonists, NSAIDS and COX-2 inhibitors, pregabalin, tramadol, magnesium and lidocaine. We observed small clinically significant reductions with ketamine and dexamethasone, although we could not exclude a moderate effect with ketamine based on 95% credible intervals. When adjusting conventional estimates for confounders, gabapentin (1200mg) demonstrated a large clinically significant reduction and the results for magnesium adjusted for allocation concealment resulted in a small clinical effect (Table 3).
Discussion

We identified baseline risk as an important cause of statistical heterogeneity in meta-analyses of analgesics. We exemplified the improvement in the precision and accuracy of evidence synthesis with our proposed novel methods for several frequently used adjuvant analgesics. Our results can be used by clinicians to estimate the effects of most adjuvant analgesics in their population based on simple audit data. In addition, we found evidence that methodological limitations explained some of the residual heterogeneity. Type of surgery did not appear to be an independent cause of between-study heterogeneity. Furthermore, our novel models have the additional advantage of being able to adjust estimates for other clinical and methodological confounding variables in the included studies.

Recent meta-analyses have attempted to explore heterogeneity using clinical covariates such as dose and type of surgery. However, these often report a low proportion of variation explained when compared to our results using baseline risk. We derived this covariate from previous empirical studies suggesting larger reductions in pain scores following analgesic treatment with higher baseline pain scores. One study examined around 500 participants following dental extraction and found those with severe pain (3/3) had greater reductions in pain scores following treatment with ibuprofen compared to those with moderate pain (2/3). Another study found acetaminophen and codeine treatment following Caesarean section was only effective in those participants with severe pain (>6/10). Although it should be noted other factors in addition to degree of pain may also influence postoperative opioid
consumption such as access to patient-controlled analgesia, concurrent analgesic
protocols, patient characteristics and the prescribing practices of attending medical
professionals (which may be region dependent).

A previous study of postoperative pain reviews has found widespread statistical
heterogeneity and suggested that this should be explored based on type of surgery or
pain scores. This review recommended future meta-analyses should include only
trials from the same surgical procedures or those with close acute postoperative pain
levels and explore this using subgroup analysis. We would argue that baseline risk is a
more appropriate covariate than type of surgery and meta-regression a more useful
analysis than subgroup analysis as it allows reporting of the proportion of
heterogeneity explained by the model ($R^2$) as well as the ability to adjust for other
confounding variables. Importantly, baseline risk encompasses many different
participant and trial characteristics such as age, sex, concurrent analgesic protocols
and other factors shown to influence postoperative morphine consumption. In our
previous meta-analysis with gabapentin, morphine consumption varied even within
procedure-specific subgroups and type of surgery was a small determinant of
heterogeneity between studies in relation to morphine consumption and pain scores.

Our results suggest that expected postoperative morphine consumption (as a surrogate
for pain, participant characteristics and concurrent analgesia) is a large determinant of
heterogeneity between studies.
Our results demonstrate that with baseline risk held constant, type of surgery did not explain the differences in effect estimates in different studies for almost all adjuvant analgesics examined, meaning the adjuvant was equally effective in all surgeries. Previous groups have argued that procedure-specific evidence is necessary when evaluating evidence derived from trials of analgesic agents. Our results suggest that the efficacy of analgesic agents is determined more by the degree of morphine consumption during the postoperative period rather than the type of surgery. Indeed, procedure-specific meta-analyses still suffer from considerable statistical heterogeneity. Therefore, we could find little empirical basis for conducting such procedure-specific reviews for adjuvant analgesics. However, we accept that this finding contradicts both previous research and clinical intuition. We can infer from our results that evidence synthesis of analgesic adjuncts should be conducted in all forms of surgery then type of surgery can be entered in a meta-regression model with baseline risk to see if differences exist. However, future clinical studies may be required to identify whether degree of pain is a better predictor of analgesic efficacy than type of surgery. In addition, we could not exclude an effect of type of surgery mediated via differences in baseline risk (some procedures having higher morphine consumption). Furthermore, we acknowledge that other interventions such as regional anesthesia may have more relevance to procedure-specific evidence.

When reporting the results from analgesics using a fixed consumption of postoperative morphine, we found the most effective analgesics were gabapentin, acetaminophen, alpha-2 agonists, NSAIDS and COX-2 inhibitors, pregabalin,
tramadol, magnesium and lidocaine, all with moderate clinically significant effects. Ketamine and dexamethasone had small clinically significant effects. However, these rankings should be interpreted with caution due to the uncertainties surrounding the point estimates, which may mean analgesics lower down the table are statistically equivalent. Furthermore, efficacy is not the only consideration when considering use of these agents. Adverse effects should also be considered when selecting an analgesic agent. Agents such acetaminophen, which have a low incidence of adverse events may be preferable to agents that induce peri-operative adverse effects such as sedation with gabapentin, especially as the differences between these agents is negligible.

In terms of the implications of our work for clinical practice, as meta-analyses are often used to inform clinical practice, reviews should present opioid reductions using a fixed consumption of morphine to more accurately reflect efficacy, as quoting the mean difference will be heavily influenced by the mean control morphine consumption from the included trials. In addition, indiscriminate use of analgesic adjuncts around the peri-operative period should be avoided. Instead, clinicians can use information from small audits of mean opioid consumption and the regression parameters in our analysis to estimate the likely mean reduction in morphine consumption for samples of patients in that particular clinical situation. For example, if an anesthesiologist carries out an audit of patients undergoing surgery on their list and finds a mean morphine consumption of 70mg then this information combined
with Table 1 parameter estimates can be used to assess the likely mean reduction in 24-hour morphine consumption with acetaminophen:

Mean reduction in 24-hour morphine (mg) = intercept + beta coefficient x 70mg

= 0.77 + -0.38 x 70

= 25.83mg

As all agents are associated with adverse effects, this more targeted use of analgesic adjuncts may help improve clinical significance and avoid inappropriate use of multiple agents when expected opioid reductions are small. For example, gabapentin is known to cause excessive sedation\textsuperscript{10} so the benefits would need to balanced against this risk. In the instance where average reductions in morphine are 3mg (low baseline risk) versus 20mg (high baseline risk), a clinical decision can be made that in the first instance, the risk may not be worth the estimated benefits. It should also be noted that at the extremes of baseline risk, confidence intervals are wider (Figure 2) which needs to be considered when using these estimates. Confidence intervals can be extrapolated from this figure, which at lower baseline risk may include 0 (p>0.05).

In terms of randomized controlled trial design, when studying analgesic agents for postoperative pain, trials should be conducted in surgeries where expected postoperative morphine consumption is anticipated to be high. For example, for intravenous acetaminophen, where the expected postoperative morphine consumption is either 70mg or 20mg in the first 24-hours postoperatively, the anticipated reduction
in morphine would be 26mg and 6mg respectively. Relying solely on the mean difference (8mg) may underestimate clinical significance in the context where postoperative morphine consumption is high. Furthermore, such larger reductions in morphine consumption may have a more pronounced effect on opioid adverse effects,\textsuperscript{139} which have additional clinical relevance. In terms of trial conduct, as with previous studies, we have found evidence that methodological limitations, in particular allocation concealment, were associated with larger reductions in morphine for many adjuncts.\textsuperscript{140} Given that only a third of the included studies reported adequate allocation concealment, this is a particular area of internal validity future studies should aim to address.

In terms of secondary research studies, future meta-analyses of postoperative analgesic agents should aim to explore heterogeneity using control group morphine consumption, in addition to other sources of clinical heterogeneity such as dose or route of administration. Such explanation of statistical heterogeneity would lead to higher quality evidence derived from these reviews as per GRADE.\textsuperscript{2} Effect estimates from these reviews should be reported and displayed using our methods from a fixed consumption of morphine to avoid confounding by the variable baseline risk in the included primary studies (and confirmed using Bayesian analysis). To illustrate this issue using a simple example, if we take an identical intervention (gabapentin) and perform random-effects meta-analysis of the ten studies with the lowest (MD -1.8mg) and highest (MD -36.8mg) baseline risk we get widely disparate effect estimates.
As an extension to this, incorporating other clinical and methodological covariates into these regression models to adjust estimates can reduce further confounding. As systematic reviews are inherently observational (despite deriving data from randomized studies), more advanced and appropriate statistical methods are required (regression) that allows more accurate prediction than using mean differences, while having the additional advantage of controlling for known confounders. For these reasons, future reviews of postoperative analgesics should avoid univariate subgroup analyses (due to confounding) and move towards multivariate regression models, which include control group morphine consumption (as is common practice in observational primary research studies). Future reviews could also investigate whether heterogeneity in other outcomes such as pain scores or length of stay can be explained by baseline risk. Our previous work has indicated that this may indeed be the case and other research has shown this may apply to other outcomes such as depression. Such explanation may have clinical implications, as well as allow use of the methods presented in this paper to improve reporting of effect estimates.

There are several limitations with this review. Firstly, meta-regression analysis should be regarded as observational despite deriving data from randomized studies. Such analyses are prone to both residual confounding and aggregation bias (as results are based on aggregated study estimates rather than from individual patients). For this reason, our implications for clinical practice focus on aggregated patient outcomes (from audits) rather than applying these to individual patients. Nevertheless, our
models are based on studies including only a certain type of surgery or patient characteristic (from inclusion criteria), which makes extrapolation of our models to clinical circumstances outside these clinical circumstances problematic. Secondly, we cannot rule out type I errors in our analyses. Although conventional to set a lower level of significance to covariate adjustment in regression models (p<0.1), this may also increase false positive results.

Thirdly, although our models can adjust for confounding variables, our analyses are limited to published primary studies and are therefore still susceptible to publication bias. Although identification of imprecise study effects is possible in systematic reviews, it is impossible to know if this is secondary to true publication bias and therefore this limits our findings. Fourthly, as we generally derived our studies from reviews of active versus placebo groups, we were unable to perform network meta-analysis, which may be a more appropriate method to directly compare analgesics in future reviews. Finally, we included trials from a variety of patient populations with differing clinical characteristics (co-morbidity for example), differing analgesic protocols and healthcare systems, which may all influence morphine consumption. However, the use of a fixed consumption of morphine should have controlled for these differences when analyzing baseline risk.

In conclusion, we have demonstrated that baseline risk (control group morphine consumption) explained a large proportion of between-study variance and controlling for this allows pooling of heterogeneous trials in meta-analyses of adjuvant
analgesics. In addition, we have shown how the use of baseline risk can allow more accurate estimation of the likely clinical benefit of analgesic adjuncts using local departmental audit data. Moreover, we have developed a novel method of presenting and an amended method of graphically displaying effect estimates which mitigate variable baseline risk in included trials and are able to control for other confounding variables. Despite the limitations of our analysis, we recommended use of these methods in future research studies and clinical practice.
Acknowledgements

We would like to thank the staff at the Royal Derby Hospital library for accessing articles.
References


32 STATA Version 14. *College Station, Texas: StataCorp LP.*


36 Scott LJ, Perry CM. Tramadol. *Drugs* 2000; 60: 139-76.


40 Baidya DK, Agarwal A, Khanna P, Arora MK. Pregabalin in acute and chronic

41 Girard P, Chauvin M, Verleye M. Nefopam analgesia and its role in multimodal
2016; 43: 3-12.

42 Jibril F, Sharaby S, Mohamed A, Wilby KJ. Intravenous versus oral acetaminophen
for pain: Systematic review of current evidence to support clinical decision-making.

43 Jebaraj B, Maitra S, Baidya DK, Khanna P. Intravenous paracetamol reduces
postoperative opioid consumption after orthopedic surgery: a systematic review of

44 Macario A, Royal MA. A literature review of randomized clinical trials of
intravenous acetaminophen (paracetamol) for acute postoperative pain. *Pain Pract*

45 Hyllested M, Jones S, Pedersen JL, Kehlet H. Comparative effect of paracetamol,
NSAIDs or their combination in postoperative pain management: a qualitative review.
46 Zemmel MH. The role of COX-2 inhibitors in the perioperative setting: efficacy and safety-a systematic review. *AANA J* 2006; 74: 49.


58 Chan AKM, Cheung CW, Chong YK. Alpha-2 agonists in acute pain management. 
*Expert Opin Pharmacother* 2010; 11: 2849-68.

59 Ben-Abraham AAWR. Dextromethorphan and dexmedetomidine: new agents for 

60 Armand S, Langlade A, Boutros A *et al.* Meta-analysis of the efficacy of extradural 
clonidine to relieve postoperative pain: an impossible task. *Br J Anaesth* 1998; 81: 
126-34.

61 Zakkar M, Frazer S, Hunt I. Is there a role for Gabapentin in preventing or treating 

62 Clivatti J, Sakata RK, Issy AM. Revisão sobre o uso de gabapentina para controle 

63 Remérand F, Couvret C, Baud A, Laffon M, Fusiardi J. Balance bénéfique risque 
de la prégabaline en périopératoire: revue systématique de la littérature. *Ann 
Françaises d'anesthésie et de Reanimation* 2011; 30: 569-77.

64 McCarthy GC, Megalla SA. Impact of intravenous lidocaine infusion on 
postoperative analgesia and recovery from surgery. *Drugs* 2010; 70: 1149-63.


95 Subramaniam K, Subramaniam B, Steinbrook RA. Ketamine as adjuvant analgesic to opioids: a quantitative and qualitative systematic review. *Anesth Analg* 2004; 99: 482-95.


112 Seib RK, Paul JE. Preoperative gabapentin for postoperative analgesia: a meta-

113 Hurley RW, Cohen SP, Williams KA, Rowlingson AJ, Wu CL. The analgesic

114 Mathiesen O, Møiniche S, Dahl JB. Gabapentin and postoperative pain: a
qualitative and quantitative systematic review, with focus on procedure. *BMC

115 Mishriky BM, Waldron NH, Habib AS. Impact of pregabalin on acute and
persistent postoperative pain: a systematic review and meta-analysis. *Br J Anaesth*
2015; 114: 10.

116 Yao Z, Shen C, Zhong Y. Perioperative pregabalin for acute pain after

117 Zhang J, Ho KY, Wang Y. Efficacy of pregabalin in acute postoperative pain: a


129 Murphy JD, Paskaradevan J, Eisler LL et al. Analgesic efficacy of continuous intravenous magnesium infusion as an adjuvant to morphine for postoperative


141 Smith AF, Carlisle J. Reviews, systematic reviews and Anaesthesia. *Anaesthesia* 2015; 70: 644-50.

Figure Legends

**Figure 1:** Flowchart of included reviews. The left chart describes the initial search strategy where we identified reviews from electronic databases with 105 included in the final review. The right hand chart describes the identification of randomized controlled trials, which were obtained from the 105 included reviews.

**Figure 2:** The nine amended scatterplots show for each adjuvant analgesic separately, how each study estimate (green dots) for mean morphine-equivalent reduction (y-axis) is linearly related to the baseline morphine consumption (x-axis) in the study population (STATA code to produce this in Supplementary Text Document 2). Diagonal red line represents the regression line and blue dotted area is the 95% CI. The vertical red line represents the 50 mg consumption at which to estimate the dose reduction with each adjuvant analgesic. The horizontal red line represents the effect estimate from a fixed consumption of 50mg where these lines meet the Y-axis (effect estimate equivalent to the middle of the diamond on a Forest plot). The dashed horizontal red lines represent the 95% CI from a fixed consumption of 50mg (equivalent to the width of the diamond on a Forest plot). A more detailed plot with the size of each study estimate representing the weight given to the study for the evidence synthesis, is available on request.

**Figure 3:** The forest plot of effect estimates for each analgesic in order of efficacy is based on our Bayesian meta-regression analysis and subsequent modelling for each
analgesic and suggests for example, a much greater reduction in mean morphine consumption for gabapentin compared with dexamethasone. Effect size (ES) is the median and 95% CrIs from Bayesian meta-regression estimates. X-axis is the mg reduction in morphine consumption from a fixed consumption of 50mg.