Pain management strategies after total knee arthroplasty: A protocol for a network meta-analysis of randomized controlled trials

Abdullah S. Terkawi, MD,1,2,3 Dimitris Mavridis, PhD,4,5 Daniel I. Sessler, MD,3,6 Megan S. Nunemaker, MSLS,7 Khaled S. Doais, MD,2 Rayan S. Terkawi, MD,2 Yazzed S. Terkawi, MS,8 and Edward C. Nemergut, MD1.

1. Department of Anesthesiology, University of Virginia, Charlottesville, VA, USA
2. Department of Anesthesiology, King Fahad medical city, Riyadh, Saudi Arabia
3. OUTCOMES RESEARCH Consortium, Cleveland Clinic, Cleveland, OH, USA
4. Department of Primary School Education, University of Ioannina, Greece
5. Department of Hygiene and Epidemiology, University of Ioannina, Greece
6. Department of OUTCOMES RESEARCH, Cleveland Clinic, Cleveland, OH, USA
7. Claude Moore Health Sciences Library, University of Virginia, Charlottesville, VA, USA
8. Umm Durman University, Faculty of Medicine, Khartoum, Sudan.
Background

Knee replacement, or total knee arthroplasty (TKA), is among the most common orthopedic procedures. TKA aims to relieve pain, improve quality of life, and improve or maintain knee function.\textsuperscript{1}

In the United States, about 4 million adults have had a TKA, representing 4.2\% of the population exceeding fifty years of age. Over half of adults in the U.S. diagnosed with knee osteoarthritis will eventually have a total knee replacement. The prevalence of TKA among older adults in the U.S. far exceeds that of rheumatoid arthritis, and is almost as prevalent as congestive heart failure.\textsuperscript{2} Nearly 700,000 TKA procedures are performed annually in the US. This number is expected to increase to 3.5 million procedures per year by 2030.\textsuperscript{3}

Inadequate postoperative analgesia impairs rehabilitation, prolongs hospitalization, and increases the risk of adverse events including myocardial ischemia and infarction, pulmonary dysfunction, paralytic ileus, urinary retention, thromboembolism, impaired immune functions, and anxiety.\textsuperscript{4} Importantly, inadequate postoperative pain control is strongly associated with development of persistent postsurgical pain.\textsuperscript{4} Postoperative pain may also worsen postoperative blood loss after TKA.\textsuperscript{5}

Chang and Cho\textsuperscript{6} conducted a survey to evaluate various analgesic approaches to TKA, comparing pain intensity and analgesic efficacy in 424 patients who had a TKA in 14 hospitals. They found that pain management protocols and pain intensity varied greatly, particularly during the initial two postoperative days. Differences in pain intensity were greatest the first postoperative night, with mean visual analog scores ranging from 17 to 94 mm on a 100-mm scale. Combined use of periarticular infiltration and femoral nerve blocks provided better analgesia than other methods during the first two postoperative days. Furthermore, patients who had either periarticular injection along with a femoral nerve block, or epidural analgesia, reported being most satisfied two weeks after TKA.
Perioperative management of TKA pain has evolved rapidly in recent decades. Before the 1990s, nurse-administered systemic opioid was practically the only analgesic approach. Nurse-administered opioids were gradually replaced by patient-controlled opioids (PCA) in 1990s. Soon thereafter, epidural analgesia became more common. In early 21st century, peripheral nerve blocks and peri-articular local anesthetic infiltration gained popularity.\(^4\),\(^7\),\(^8\) In recent decades, there has been great interest in defining the optimal peripheral nerve block and periarticular (and intra-articular) local anesthetic infiltration techniques — although the best approach remain unclear. Multimodal analgesia (that is, combining a peripheral nerve block with peri-articular local anesthetic infiltration or oral analgesia and pregabalin has also proven effective.

Options for peripheral nerves blocks include: lumbar (psoas) plexus, femoral, sciatic, obturator, 3-in-1, fascia iliaca, and adductor canal. Each can be performed as a single injection or provided as a continuous infusion.\(^4\)

A recent Cochrane review (2014) found that femoral nerve blocks (with or without concurrent treatments including PCA opioid) after TKA provided better analgesia than PCA opioid alone, similar analgesia to epidural blocks, and less nausea/vomiting than PCA alone or PCA with epidural analgesia.\(^9\) The review also found that continuous femoral nerve blocks provided better analgesia than single-shot blocks. The authors did not found sufficient evidence to support definitive conclusions regarding the comparison between femoral nerve block and local infiltration analgesia or oral analgesia.\(^9\)

Another 2014 pairwise meta-analysis evaluated the efficacy of local anesthetic infiltration versus placebo, no infiltration, or femoral nerve block. These investigators found significantly improved analgesia in the initial 24 postoperative hours in patients who given local anesthetic infiltration instead of placebo, but similar analgesia with local anesthetic infiltration and femoral nerve blocks.\(^10\)

Another systematic review in 2014 evaluated the efficacy of high-volume multimodal wound (peri-articular) infiltration (single dose or continuous infusion) versus no infiltration, femoral nerve block, or epidural analgesia. They observed that better acute analgesia after wound infiltration, without definitive evidence that infiltration
reduced opioid consumption, achievement of early milestones, or shortened hospitalization. The authors could not come to definitive conclusions regarding the precise role of individual agents or in the use of a percutaneous wound catheter for postoperative administration in providing pain relief.11

Finally, Anderson and Kehlet12 conducted another systematic review in 2014 assessing the analgesic efficacy of local infiltration analgesia in TKA. They found, with sparse evidence, that local infiltration analgesia provided better analgesia than placebo, equal effect to femoral nerve block, and similar or better efficacy than epidural analgesia. Most of the assessed trials had a high risk of bias and did not use sufficient pain management protocol in the control group, which restrict firm conclusions.

Many surgeons and anesthesiologists avoid femoral nerve blocks for fear of associated motor weakness and consequent risk of patient falls. But whether femoral nerve blocks actually increase the risk of patient falls remains unclear. In a retrospective study of 2,197 patients, Wasserstein et al.13 found only that a continuous femoral nerve block, but not single shot block, was an independent risk factor for falls. In another retrospective study that involved 191,570 patients from the national Premier Perspective database, Memtsoudis et al. did not found an association between femoral nerve blocks and falls.14

Recently, a long-acting liposomal formulation of bupivacaine (EXPERAL®) was approved by the U.S. Food and Drug Administration for single-dose injection into the surgical site.15 While this may seem a to be an optimal drug in TKA, a recent review found insufficient evidence to support its efficacy.4 Furthermore, a recent retrospective study showed liposomal bupivacaine provided inferior pain control than a traditional multimodal analgesic approach in patients recovering from TKA.16

The gold standard for postoperative analgesia remains unclear. The optimal modality should achieve effective pain control with less opioid consumption and the best rehabilitation profile.17 There are now more than ten competing pain management strategies for TKA. It would be prohibitively expensive and impractical to conduct a randomized trial simultaneously comparing them all. We therefore propose to compare
available interventional analgesic methods using a network meta-analysis approach. The advantage of this approach is that network meta-analysis extends the concept of the traditional meta-analysis to produce pairwise comparisons and relative treatment effects across a range of interventions.\textsuperscript{18}

**Objectives**

1. To assess the available interventional pain management modalities for TKA in terms of:
   a) Efficacy: analgesia, opioid consumption, and rehabilitation;
   b) Safety: side effects and duration of hospitalization.
2. To generate a clinically useful ranking of available pain management modalities according to their efficacy and safety.
Methodology

The study registered in PROSPERO 2015: CRD42015015870. Available at: [http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015015870](http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015015870)

Criteria for considering studies for this review

**Inclusion criteria:**
We will only include randomized clinical trials that evaluated pain management efficacy, quality of recovery (e.g., nausea and vomiting), and rehabilitation profile after unilateral total knee arthroplasty using any of the following interventional techniques:

1) *Neuraxial analgesia*: epidural and spinal analgesia.

2) *Peripheral nerve blocks* (single dose or continuous infusion):
   a. 3-in-1 nerve block
   b. Femoral nerve
   c. Fascia iliaca compartment block
   d. Sciatic nerve
   e. Obturator nerve
   f. Lumbar plexus (psoas) block
   g. Adductor-canal-block

3) *Intra-articular and periarticular local anesthetic infiltration*. All intra-articular, subcutaneous, and peri-articular infiltration has been referred in the literature as “local anesthetic infiltration”.

4) *Auricular Acupressure*; although it is considered non-invasive, but we intended to include it as it is an interesting growing eastern pain management modality.

5) *Intravenous patient control analgesia (PCA)*

6) *Placebo* (systematic opioid will be considered placebo when not given via PCA).

**Exclusion criteria:**
1) Oral pain medications.

2) Retrospective studies, case reports, case series, abstracts, pilot studies, and non-randomized prospective studies.

3) Arthroscopy.

4) Studies that included both knee and hip patients, without separate presentation of the results for each.

5) Combinations of more than one intervention category: epidural with peripheral nerve block, epidural with local infiltration, or peripheral nerve block with local infiltration. The combination of more than one intervention from the same category (e.g., femoral with sciatic nerves blocks) is not an exclusion criterion.

We assume that patients who meet inclusion criteria are, in principle, equally likely to be randomized to any of the eligible interventions as a starting point. We will also explore deviations from this assumption.

**Type of outcome measures:**

- **Primary outcomes:** (1) acute postoperative pain (during rest and movement); (2) intra- and post-operative opioid consumption; and, (3) quality of early postoperative rehabilitation (functional assessments).

- **Secondary outcomes:** postoperative complications (e.g., nausea, vomiting, falls), duration of hospitalization, amount of blood loss, incidence of procedure failure, and patient withdrawal.

**Definition of relevant outcome**

1. Postoperative acute pain during hospitalization.
2. Intraoperative and postoperative opioid consumption during hospitalization.
3. Postoperative nausea and vomiting during hospitalization.
4. Duration of hospital stay; usually counted after discharge from post-anesthesia care unit (PACU).
5. Patient satisfaction
6. Blood loss
7. Complications; e.g., nausea, vomiting
8. Incidence of postoperative in-hospital falls
9. Quality of early rehabilitation which can include any of the following validated physical-performance-based measures:\[19\]
   a) Range of motion (ROM): these might be reported as continuous or dichotomous variables. ROM is a measure of combined flexion/extension of the operated knee, either actively or passively with assistance. Ninety degrees is considered the minimum range required to navigate steps, while 110 degrees is necessary for adequate performance of activities-of-daily-living (ASLs). Typically, range is measured with a goniometer.
   b) Quadriceps strength (QS): Strength is typically measured with an electromechanical dynamometer, and often reported as maximum voluntary isovolumetric contraction (MVIC) in Newtons normalized to body mass index.
   c) Six-minute walk test (6MWT): the maximum distance ambulated on level ground (in meter) with standardized encouragement.
   d) Timed Up and Go (TUG): the time taken to stand up from a standard-height armchair, walk 3 meters, walk back to the chair, and sit down. It is meant to assess patient’s balance and risk of falling.
   e) Stair time (ST): the time necessary to ascend and descend standard-height 20-cm steps.
   f) Self-paced walk test (SPWT): timing patients walking a 20-meter course bi-directionally.
Since there is no currently evidence to show superiority for any of these measures over others, and we are not sure how consistent the studies are reporting them, we will ultimately use the ones most commonly reported to reflect the quality of early rehabilitation. If comparably reported, we will use ROM as the main parameter.

**Reporting functional outcomes after TKA:**
Choi et al. did a systematic review to assess the quality of reporting functional outcome assessment after TKA in patients underwent regional anesthesia. Table 1 summarizes their findings.

Two broad categories are commonly used to assess functional outcome (rehabilitation) after TKA: (1) Physical-Performance Measures which includes Range of Motion, Quadriceps Strength, Six-Minute Walk Test, Timed Up and Go, Stair Time, and Self-paced Walk Test; and, (2) Self-reported Measures which includes Western Ontario and McMaster Universities Osteoarthritis Index, Knee Outcomes Severity Score, and the Lower-Extremity Functional Scale.

Choi et al. defined the appropriate duration of functional outcome evaluation into early (initial 2 weeks), intermediate (6-12 weeks), and late (6-12 months) — and recommended that specific tests to be included for each period. They identified only two studies that reported long-term functional outcomes per their specifications and will thus limit our evaluation of rehabilitation to the early period. To the extent practical, we will focus on the tests that Choi et al. recommend.
<table>
<thead>
<tr>
<th>Functional Outcome</th>
<th>Suggested Assessment Intervals</th>
<th>Minimal Detectable Change</th>
<th>Time to Administer</th>
<th>Usage Restrictions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical-performance-based outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROM</td>
<td>Preoperative, 2 and 12 wk, 6 and 12 mo</td>
<td>9.6 degrees</td>
<td>&lt;1 min</td>
<td>None</td>
<td>• Encompasses both flexion and extension • Minimum range to navigate steps 90 degrees, ADLs 110 degrees • Measurement error 5 degrees • Maximal change ~12 wk, maximum range ~12 mo</td>
</tr>
<tr>
<td>MVIC</td>
<td>Preoperative, 4 wk; 3, 6, 12, 24 mo</td>
<td>0.33 N - m/kg</td>
<td>1 min</td>
<td>None</td>
<td>• Greatly reduced immediate postoperative period (~62%) • Continued dysfunction up to 2 y postoperatively</td>
</tr>
<tr>
<td>6MWT</td>
<td>Preoperative, 6 and 12 wk, 6 mo</td>
<td>61.3 m</td>
<td>6 min</td>
<td>None</td>
<td>• 500 m required to perform ADL • Maximal change between wk 6–9 with plateau at 26 wk</td>
</tr>
<tr>
<td>TUG</td>
<td>Preoperative, immediate postoperative period, 6 and 12 wk</td>
<td>2.5 s</td>
<td>&lt;2 min</td>
<td>None</td>
<td>• Assesses balance and risk of falls • Independent &lt;10 s, able to perform ADL &lt;20 s, mostly dependent &gt;30 s • Best utilized as a categorical variable</td>
</tr>
<tr>
<td>ST</td>
<td>Preoperative, 6 wk</td>
<td>5.5 s</td>
<td>&lt;2 min</td>
<td>None</td>
<td>• Not routinely administered immediately postoperatively • Increases &gt;100% at &lt;1 wk</td>
</tr>
<tr>
<td>SPWT</td>
<td>Preoperative, 6 wk</td>
<td>4.0 s</td>
<td>&lt;5 min</td>
<td>None</td>
<td>• More easily completed by patients in the immediate postoperative period than 6MWT</td>
</tr>
<tr>
<td>Self-report-based outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WOMAC</td>
<td>Preoperative, 6 and 12 wk, 6 and 12 mo</td>
<td>9.1 points</td>
<td>12 min</td>
<td>Adults</td>
<td>• Reported in three subsections (pain, stiffness, functional limitation) or summed with maximum total score of 96 • Available in &gt;90 languages • Greatest change wk 9–13 postoperatively • Five subsections (pain, symptoms, ADL, sports, QOL) each with a maximum score of 100 • Incorporates WOMAC • Superior to WOMAC in young (&lt;60 y), active patients • Available in 45 languages • Most sensitive to change between 6 and 12 mo</td>
</tr>
<tr>
<td>KOOS</td>
<td>Preoperative, 6 and 12 mo</td>
<td>10 points*</td>
<td>10 min</td>
<td>Adults</td>
<td>• Comparable validity and reliability to WOMAC • Maximal change between 4 and 12 wk • Plateaus at 6 mo</td>
</tr>
<tr>
<td>LEFS</td>
<td>Preoperative, 6 and 12 wk, 6 mo</td>
<td>9 points*</td>
<td>5 min</td>
<td>Adults</td>
<td></td>
</tr>
</tbody>
</table>

*Minimal clinically significant difference. MVIC indicates maximum voluntary isovolumetric contraction (quadriceps strength); OA, osteoarthritis; QOL, quality of life.

Table 1. Functional outcome measures validated for use after TKA. From Choi et al.\textsuperscript{19}

**Search strategy:**

The search will be conducted as recommended by the (ISPOR) International Society for Pharmacoeconomics and Outcomes Research 2011 Task Force.\textsuperscript{20}
The following databases will be searched: MEDLINE via PubMed, Embase, the Cochrane Library, and Web of Science’s Core Collection (excluding MEDLINE) and SciELO Citation Index. The search will not be limited by language or date. We will search www.clinicaltrials.gov for ongoing studies and contacted the authors of the ongoing studies. We will also search the major anesthesiology and orthopedic journals for online first publications after the date of conducting the literature search.

The following search terms will be used: (PubMed search example)

(((Arthroplasty, Replacement, Knee[mesh]) OR Knee replacement*) OR Knee arthroplast*)) AND (((Injections, Intra-Articular[mesh] OR Femoral Nerve[mesh] OR Nerve Block[mesh] OR Sciatic Nerve[mesh] OR Lumbosacral Plexus[mesh] OR Analgesia, Epidural[mesh] OR Analgesia, Patient-Controlled[mesh] OR Obturator Nerve[mesh] OR Acupuncture, Ear[mesh] OR Anaesthesia, Local[mesh] OR Anesthetics, Local[mesh] OR Analgesics[mesh] OR Analgesia[mesh] OR (Peri articular injection* OR Periarticular injection* OR "Local infiltration analgesia" OR "Local anesthetic infiltration analgesia" OR "Local infusion analgesia" OR "Local infiltration" OR Femoral nerve block* OR Femoral catheter* OR Adductor canal block* OR Sciatic nerve block* OR Three in one block* OR Three in one femoral nerve block* OR 3 in 1 nerve block* OR Lumbar plexus block* OR Lumbosacral plexus block* OR Lumbar plexus nerve block* OR Lumbar plexus infusion* OR Lumbar plexus catheter* OR "Epidural analgesia" OR "Patient controlled analgesia" OR Fascia iliaca compartment block* OR Obturator nerve block* OR "Auricular acupressure" OR "Auricular acupuncture" OR "Local anaesthesia" OR Local anesthetic* OR Local anaesthe* OR Analgesic* OR Regional analgesia* OR "Regional analgesia/anesthesia" OR "Regional analgesia/anaesthesia" OR "Regional anaesthesia" OR "Regional anaesthesia anaesthesia" OR Peripheral nerve block* OR Spinal anesthe* OR Spinal anaesthe* OR Opioid analgesia*))).
Selection of studies: Two independent authors will screen the resultant search for eligible studies.

Assessment for risk of bias
Two independent authors will assess the risk of bias for each study using the Cochrane Collaboration’s risk of bias assessment tool. A third reviewer will adjudicate disagreements.

The areas that will be evaluated are:

- **Random sequence generation**: Was there adequate sequence generation (selection bias)?
- **Allocation concealment**: Was allocation adequately concealed (selection bias)?
- **Blinding**: Was knowledge of the allocated intervention adequately prevented during the study (detection bias)?
  - Participants and personnel
  - Outcome assessors
- **Incomplete outcome data**: Were incomplete outcome data adequately addressed (attrition bias)?
- **Selective outcome reporting**: Are reports of the study free of possible selective outcome reporting (reporting bias)?

We will evaluate the aforementioned items within each study, and in each pairwise comparison. We will classify each piece of direct evidence in the network as having low, moderate, or high risk of bias. If significant discrepancies are found between treatment comparisons, we will illustrate these assessments in the network plot for the primary outcome with colored edges according to the risk of bias. We will also produce the contribution matrix which gives the percentage contribution of each direct estimate to the network meta-analysis estimates. This will help to delineate the contribution of direct and indirect evidence to each network meta-analysis estimate.
Data collection

Data extraction and management
Using a standardized data collection form, four researchers will review and extract data from the filtered articles. We will collect the following data: first author; year of publication; study title; journal; study country and language; randomization (e.g., parallel, or crossover); type of intervention with details on each group; number of patients in each group; demographic characteristics of each group; whether patients on chronic opioid use were included or not and their number; anesthesia, analgesia, and anti-emetic protocol used; technique of surgery; pain scores at rest and movement; intraoperative and postoperative opioid consumption; in-hospital rehabilitation profile; incidence of fall, nausea, and vomiting; duration of hospital stay; and Cochrane Collaboration’s risk of bias assessment.

Data extraction:
A team from four investigators will extract the data independently. Two investigators will extract data from each article, independently, and a third investigator will confirm the extracted data.

Pain scores are usually presented in a numeric rating scale (NRS), ranging from 0 to 10, but it may, less frequently, presented as visual analogue scale (VAS), ranging from 0 to 100, in this situation we will convert the VAS to NRS by dividing the results by 10.

Opioids will be converted to morphine equivalent in mg using a standardized conversion calculator [http://clincalc.com/Opioids/](http://clincalc.com/Opioids/).

If the pain scores and opioid consumptions are not reported numerically, they will be estimated from manual measurements of the corresponding figures.

For studies in which incidences of nausea and/or vomiting is not reported separately, but reported as incidences of postoperative nausea and vomiting (PONV), we will consider PONV to represent the nausea incidence since nausea is about ten times as common as vomiting, and vomiting without concomitant nausea is rare. In studies in which the
number of anti-emetics used is reported instead of incidences of nausea and/or vomiting, we will use that as the incidences of vomiting.

Data analysis

Measures of treatment effect
We will estimate the pairwise relative treatment effects of the competing interventions using standardized mean differences (SMD) for continuous outcomes and odds ratios (OR) for dichotomous outcomes. Effect sizes will always be accompanied by 95% confidence intervals.

Results from NMA will be presented as summary relative effect sizes (SMD or OR) for each possible pair of treatments. We will interpret the results and place confidence in the output of the network meta analysis using the methods suggested by Salanti et al. that are based on the methodology developed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group for pairwise meta-analyses.

Relative treatment ranking
We will estimate the ranking probabilities for all treatments of being at each possible rank for each intervention using the network command in STATA. We will obtain a hierarchy of the competing interventions using rankograms. We use the surface under the cumulative ranking curve (SUCRA) and mean ranks to obtain a treatment hierarchy. We will produce the relevant plots using the suite of STATA commands by Chaimani et al.

Unit of analysis issues
We expect that some studies will not report mean values and standard deviations (SD) but instead report quintiles or similar measures. If a study reports the median, minimum, and maximum values, we will use the methodology from Hozo et al. to estimate the respective mean and SD of the study population. We will include also studies that report the median and the interquartile range (IQR), assuming that data are normally distributed
and the standard deviation would be SD=IQR/1.35 with the mean equaling the median. However, reporting of medians and IQRs usually indicates a non-normal data distribution.\textsuperscript{28} We will therefore repeat the analysis excluding these studies as a sensitivity analysis for the main network. If we find studies reporting effect sizes but not standard deviation, we estimate the unreported standard deviation, if possible, using the methods found in the Cochrane Handbook.\textsuperscript{28} Similarly, if means and sample sizes are reported in each arm, but not the standard deviations, we will estimate the standard deviations, if possible, using the methodology described in the Cochrane handbook.\textsuperscript{28}

**Studies with multiple treatment groups**
We will take into account the correlations between effect sizes measured within a single trial.

**Assessment of reporting biases**
For each pairwise comparison that includes at least 10 trials, we will draw contour-enhanced funnel plots and compute Egger’s test to test visually and statistically for small-study effects.\textsuperscript{29} For these comparisons we will draw contour-enhanced funnel plots to distinguish small study effects from publication bias.\textsuperscript{30} We will also draw a comparison-adjusted funnel plot to explore for small study effects assuming that small study effects favor the novel treatment.\textsuperscript{22}

**Dealing with missing data**
Missing data and dropouts will be assessed in all included studies. Details and characteristics of dropouts will be investigated and reported. We will explore if reasons for missing data are related to the actual outcome and if missing data are balanced in the intervention arms. If it appears that data may not be missing-at-random, we will use pattern mixture models to allow for uncertainty in the summary estimate due to missing data.\textsuperscript{31}
Assessment of clinical and methodological heterogeneity within treatment comparisons

There are three different types of heterogeneity, namely clinical, methodological, and statistical. To evaluate the presence of clinical heterogeneity we will generate descriptive statistics for trial and study population characteristics across all eligible trials that compare each pair of interventions. We will assess the presence of clinical heterogeneity within each pairwise comparison by comparing these characteristics (mentioned in details on the “Investigation of heterogeneity and inconsistency via subgroup analysis and meta-regressions” section). We will assess methodological heterogeneity by evaluating the design of the studies. Statistical heterogeneity refers to differences in true effect sizes.

Assessment of transitivity across treatment comparisons

Although participants are randomized within a study, treatment strategy comparisons are not randomized across studies. We assume that an intervention is missing from a trial for reasons not associated with its relative effectiveness and any patient that meets the inclusion criteria is, in principle, equally likely to be randomized to any of the eligible interventions. This is a key assumption in network meta-analysis called transitivity. It states that we can genuinely learn about the relative effectiveness between two treatments via an indirect route.

If, for example, the treatments “femoral nerve block” and “peri-articular infiltration” are both directly compared to “epidural analgesia”, then we can assess “femoral nerve block” vs. “peri-articular infiltration” indirectly through “epidural analgesia”. This assumption is that “epidural analgesia” is similar when it appears in “epidural analgesia” vs. “femoral nerve block” and “epidural analgesia” vs. “peri-articular infiltration” trials, and also that the distribution of effect modifiers is similar in “epidural analgesia” vs. “femoral nerve block” and “epidural analgesia” vs. “peri-articular infiltration” trials.

The assumption of transitivity will be evaluated for all treatment comparisons. Specifically, we will assess the transitivity assumption by comparing the distribution of the potential effect modifiers across various pairwise comparisons. We suspect that year of study publication and sample size of the trial will be effect modifiers, and we will
explore whether the distribution of these potential modifiers differs across treatment comparisons. We will assume that the most common treatment (epidural analgesia, in this example) used for indirect comparisons is itself similar when it appears in different trials.

**Data synthesis**

**Methods for direct treatment comparisons**
We will conduct pairwise meta-analyses in STATA, using random effects models\textsuperscript{34} for each treatment comparison with at least two studies.

**Methods for indirect and mixed comparisons**

We will use network meta-analysis to compare various pain management interventions for TKA. Network meta-analysis synthesizes both direct and indirect evidence, estimates the relative effectiveness amongst pairs of interventions, even if specific interventions have never been compared directly in RCT’s, and provides a ranking of interventions.\textsuperscript{35-38} For example, femoral nerve block vs. epidural analgesia, direct evidence would be provided by trials directly comparing these two interventions whereas indirect evidence would be provided by an indirect path linking these two treatments.

By combining direct and indirect evidence we obtain estimates with increased precision. We will perform network meta-analysis in STATA using the \textit{network} command\textsuperscript{25} and self-programmed STATA routines available at http://www.mtm.uoi.gr/index.php/stata-routines-for-network-meta-analysis.\textsuperscript{22}

**Assessment of statistical heterogeneity**

**Assumptions when estimating heterogeneity**

In standard pairwise meta-analyses, we assume different heterogeneity estimates for different comparisons. In network meta-analysis we assume that heterogeneity is the same
for all treatment comparisons. We estimate heterogeneity using restricted maximum likelihood both in pairwise and network meta-analysis.

**Measures and tests for heterogeneity**

We will assess statistical heterogeneity visually by inspecting the forest plot for each pairwise comparison. We will compute the $I^2$ index and the chi-square statistic within each pairwise comparison. Both these two measures can be unreliable and the chi-square statistic has low power to detect heterogeneity. For dichotomous outcomes, we will compare the estimated values for heterogeneity to their empirical distribution derived by Turner et al.

**Assessment of statistical inconsistency**

**Local approaches for evaluating inconsistency**

To evaluate the presence of inconsistency locally [i.e., within a specific closed loop of evidence (e.g. if there are studies comparing A vs B, B vs C and A vs C then treatments A, B and C form a closed loop of evidence)] we will use the loop-specific approach. This method evaluates the consistency assumption in each closed loop of the network separately as the difference between direct and indirect estimates for a specific comparison in the loop (inconsistency factor). Then, the magnitude of the inconsistency factors and their uncertainty expressed by 95% confidence intervals are used to infer about the presence of inconsistency in each loop. We will assume a common heterogeneity estimate within each loop. We will present the results of this approach graphically in a forest plot using the *ifplot* command in STATA. We will use the node-splitting approach to evaluate if there is a difference, particular comparison, between ‘direct’ and ‘indirect’ evidence.

**Global approaches for evaluating inconsistency**

We will use the ‘design-by-treatment’ model as described by Higgins and colleagues to check the assumption of inconsistency in the entire network for each outcome. This method accounts for different source of inconsistency that can occur when studies with
different designs (two-arm trials vs. three-arm trials) give different results as well as disagreement between direct and indirect evidence. The design-by-treatment model will be performed in STATA using the `mvmeta` command.

**Investigation of heterogeneity and inconsistency via subgroup analysis and meta-regressions**

If we find important heterogeneity or/and inconsistency, we will explore possible sources. If sufficient studies are available, we will perform meta-regression or subgroup analyses for the primary outcome by using the following effect modifiers as possible sources of inconsistency and or heterogeneity:

1. Year of publication: older studies were done at the time where more invasive surgical techniques were used.
2. Patients age
3. Patients gender
4. Surgical technique (minimal invasive versus standard): minimally invasive surgery (MIS) total knee arthroplasty (TKA) approaches were introduced as an alternative approaches than the standard TKA approach. These include; the limited parapatellar, limited midvastus, limited subvastus, and quadriceps-sparing approaches.\(^{45}\)
5. Whether the study was funded or not, and whether it was founded by a pharmaceutical company.
6. Duration of hospitalization, which might be shorter in hospitals with fast-track discharge protocols.\(^{46}\)
7. Variations in the drug type, dose, concentration that used
8. Whether or not non-local anesthetic drugs were used as an adjuvants in the mixture of the medications (e.g., opioids, ketamine)
9. The use of concomitant analgesic regimen (e.g., nonsteroidal anti-inflammatory drugs, acetaminophen, or gabapentin)
10. Type of anesthesia: general (total intravenous anesthesia vs. volatile anesthesia with or without nitrous oxide), and neuroaxial anesthesia (spinal vs. epidural).
11. Preoperative chronic pain treated chronically with opioids.
12. Whether the procedure is for the first or second knee in staged bilateral total knee arthroplasty. Recent study suggest that patients having staged bilateral TKA experience more postoperative pain with the second procedure, perhaps because of hyperalgesia extending beyond the initial injury site and/or central sensitization.\textsuperscript{47}

**Sensitivity analyses**

For the primary outcome and the main network we will repeat the analysis excluding those studies that are at high or unclear risk of bias. If there are large missing rates and suspicions that data are missing not at random, we will apply pattern mixture models to account for missing outcome data.\textsuperscript{31} We will also repeat the analysis including studies with bilateral one-stage total knee arthroplasty.
References


