PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	A prospective, randomized, controlled study on the efficacy and safety of different strategies of tranexamic acid with total blood loss, blood transfusion rate and thrombogenic biomarkers in total knee arthroplasty: study protocol
AUTHORS	Jin, Qunhua; Yang, Yong; Wang, Zheng; Wang, Faxuan; Zhao, Xin; Yang, Kaijie; He, Jinlong; Jin, Yun; Yang, Haibo; Ding, Dong

VERSION 1 – REVIEW

REVIEWER	Sachiyuki TSUKADA, MD Hokusuikai Kinen Hospital, Japan
REVIEW RETURNED	27-Mar-2020

GENERAL COMMENTS	Major Concerns: 1. This RCT includes many primary outcomes. In principle, one RCT should have one primary outcome. 2. In Limitation part, the authors mentioned that their sample size was small. The main reason of this limitation was that their RCT included five groups. Too many study groups were not appropriate to RCT. Minor concerns: 1. Line 18- total knee arthroplasties -> TKAs 2. Line 22- Combined topical and intravenous administration was proved to be one of the most effective routes [Nielsen, et al. Combined intra-articular and intravenous tranexamic acid reduces blood loss in total knee arthroplasty: a randomized, double-blind, placebo-controlled trial. J Bone Joint Surg Am. 2016;98:835–41.]. Clarify why the authors did not include this administration route. 3. Line 29- Why 5 groups? Too many. 4. Line 30- The authors must determine one primary outcome. 5. Line 58- For detecting VTE, the reviewer considers that d-dimer would be used most frequently. Please clarify why d-dimer is not appropriate in this RCT. 6. Table 1. Recently published guideline [Fillingham YA, et al. Tranexamic acid use in total joint arthroplasty: the clinical practice guidelines endorsed by the American Association of hip and Knee Surgeons, American Society of Regional Anesthesia and Pain Medicine, American Academy of Orthopaedic surgeons, hip society, and knee society. J Arthroplasty. 2018;33:3065–9.] raised problems about the lack of high-level evidence to support the use of tranexamic acid in patients with a history of thromboembolic events. Please clarify more concisely whether these patients were included or not. 7. Line 148- Recently, many institutions do not use drain for TKA. Why does your institution use it? Clarify the reason.

REVIEWER	Jiri Gallo Palacky University, University Hospital Olomouc, Czech Republic
REVIEW RETURNED	28-Mar-2020

GENERAL COMMENTS	Thank you for the opportunity to review the manuscript "A prospective, randomized, controlled study on the efficacy and safety of different strategies of tranexamic acid with total blood loss, blood transfusion rate and thrombogenic biomarkers in total knee arthroplasty: study protocol" by Yang et al. I agree with the need to specify the protocols of TXA administration in patients undergoing TKA implantation. The authors suggest a comparison of the intravenous regimen (IV) with three ways of topical TXA administration (traditional irrigation before closing the wound, periarticular soft tissue injection of TXA and delivery of TXA to the joint through drainage). Beside evaluation of the utility in the sense of reduction of blood loss (total, or from drains), lower decrease in haemoglobin, or a lower number of blood transfusion, they plan the application of sensitive thrombogenesis tests – Plasminogen activator inhibitor – 1 (PAI-1), Prothrombin fragment F1+2 (F1+2), thrombin-antithrombin complexes (TAT) – all these in plasm (in 2 hours, on the 1st and 3rd days postoperatively); or in waste in drains 2 hours postoperatively. The sample size was calculated with the objective to determine noninferiority of tested interventions for the transfusion rate against the standard (i.e. IV). The RCT protocol is clearly described (including quality control); it largely meets the SPIRIT criteria (the list of fulfilled items is attached); the strengths of proposed protocol are obvious. Major comments: 1) There is a certain concern over the undesirable influence of high TXA concentrations on synovialocytes, tenocytes and perhaps other cells; this is why it is necessary for the authors to clarify how it is exactly planned to namely perform the periarticular application of TXA because they plan closing the joint first and then apply TXA into the periarticular tissues – where exactly? to what depth? into retinacula, or into the back part of the joint too? 2) Beside evaluation of benefits of TXA administration, the authors plan to evaluate harms too – however, I am not sure whether a sample of 50 patients (in each shoulder of RCT) receiving DVT prevention allows identification of differences in wound healing impairments, DVT, or even pulmonary embolism – studies focused on these events must be performed on much greater numbers of patients. 3) The authors should also state on which postoperative day the patients are usually discharged as they state that ultrasound examination of lower extremity blood vessels will be performed before discharge. If a thrombus takes several days to develop, it is possible that it will not be detected by ultrasound at all under the designed protocol, or that the researchers will learn about its symptomatic forms during a check-up 1 or more months later, and the asymptomatic forms will pass completely unnoticed. 4) I do not like the introduction in Discussion, the 1st paragraph – blood management (BM) is part of all the TKA surgery protocols (and those of many other big surgeries); I cannot see a reason for emphasizing the unique relation between fast-track surgery and blood management because it does simply not exist. I would start with "BM is an inseparable part of all big orthopaedic surgeries, including TKA implantation …" 5) I cannot see why the authors do not mention the combined way of

	TXA administration designed and tested for the potential synergy effect and decrease of the risk of complications. This procedure is also supported by a rationale and backed by literature (both RCTs and RTC meta-analyses). Minor comments: 1) Some grammatical and typographical mistakes. 2) VTE is explained in the Introduction (it is misspelt as TVE on page 4, line 73) and then again introduced on line 222. 3) I cannot agree with the sentence on the 240: "... the results of this study are not physician or implant dependent." I believe that surgery outcomes are always dependent on the surgeon – the correct formulation should be that variability of surgery outcomes is lowered by a single surgeon ... 4) Will the samples really be frozen at 280°C? (lines 177–78).

VERSION 1 – AUTHOR RESPONSE

Replies to Reviewer 1

Major Concerns:

1. This RCT includes many primary outcomes. In principle, one RCT should have one primary outcome.

Answer: Thanks for your valuable suggestion.

The outcome measures in this study have been revised. Primary outcome: Total blood loss (TBL); secondary outcomes: *blood transfusion rate (BTR), drainage volume, plasma D-dimer, plasma and drainage PAI-1, Plasma and drainage TAT, plasma and drainage F1+2, wound complications, venous thromboembolism (VTE) and length of hospital stay (LOS). Line 230-231.*

Thank you!

2. In Limitation part, the authors mentioned that their sample size was small. The main reason of this limitation was that their RCT included five groups. Too many study groups were not appropriate to RCT.

Answer: Thanks for your valuable suggestion.

The purpose of this study is to investigate the most effective administration from intravenous and three topical applications. Therefore, five groups were included in this trial. Considering the volume of patients in our hospital and the duration of this study, the sample size of each group was calculated to be 50 participants. Too many study groups bring much heavier work to carry out and statistically analyze, but we have confidence to finish this study. Line 45-57.

Thank you!

Minor concerns:

1. Line 18- total knee arthroplasties -> TKAs

Answer: Thanks for your valuable suggestion. The "total knee arthroplasties" in line 18 has been replaced with "TKAs". Line 18.

Thank you!

2. Line 22- Combined topical and intravenous administration was proved to be one of the most effective routes [Nielsen, et al. Combined intra-articular and intravenous tranexamic acid reduces blood loss in total knee arthroplasty: a randomized, double-blind, placebo-controlled trial. J Bone Joint Surg Am. 2016;98:835–41.]. Clarify why the authors did not include this administration route.

Answer: Thanks for your valuable suggestion.

Nielsen, et al. have demonstrated the combined intravenous and topical administration of TXA to be effective. In Nielsen's study, only one topical administration (capsule injection) combined with intravenous（IV） administration was investigated comparing to intravenous administration [J Bone Joint Surg Am. 2016 May 18;98(10):835-41.]. Other topical administrations combined with IV administration still need to be investigated. Our study includes IV and three topical administrations of TXA, if we define the most effective topical administration among three topical administrations, the result may provide a direction for further research on combined administration of TXA.

Thank you!

3. Line 29- Why 5 groups? Too many.

Answer: Thanks for your valuable suggestion.

The objective of this study is to investigate the most effective administration from intravenous and three topical administrations of TXA. Thus, five groups were included in this study. Sarzaeem MM et al. performed a clinical, randomized and double-blind study to compare the efficacy of three administrations of TXA in TKAs, there were 4 groups included in Sarzaeem's study[J Arthroplasty. 2014 Aug;29(8):1521-4.]. In order to investigate four administrations of TXA, we added the PI group in our study.

Thank you!

4. Line 30- The authors must determine one primary outcome.

Answer: Thanks for your valuable suggestion.

The outcome measures in this study have been revised. Primary outcome: Total blood loss (TBL); secondary outcomes: *blood transfusion rate (BTR), drainage volume, plasma D-dimer, plasma and drainage PAI-1, Plasma and drainage TAT, plasma and drainage F1+2, wound complications, venous thromboembolism (VTE) and length of hospital stay (LOS). Line 230-231.*

Thank you!

5. Line 58- For detecting VTE, the reviewer considers that d-dimer would be used most frequently. Please clarify why d-dimer is not appropriate in this RCT.

Answer: Thanks for your valuable suggestion.

D-Dimer, as a breakdown product of cross-linked fibrin resulting from fibrinolysis, has been considered as an indicator for VTE. However, some other conditions, including an increased age, tissue injury, infection, and acute respiratory distress syndrome also lead to the elevation in D-dimer levels. The current cut-off value of D-dimer, 0.5mg/L, leads to a high sensitivity and low specificity [Clin Appl Thromb Hemost. 2017 Jan;23(1):78-83.]. Several studies have defined different threshold values in different fields [Clin Appl Thromb Hemost. 2017 Jan;23(1):78-83. & J Neurosurg. 2013 Nov;119(5):1340-6.]. However, there is no study on the predictive value of D-dimer in VTE after total knee arthroplasty. Therefore, our further objective of this study is to investigate the predictive value of D-dimer, PAI-1, TAT and F1+2 after TKAs.

6. Table 1. Recently published guideline [Fillingham YA, et al. Tranexamic acid use in total joint arthroplasty: the clinical practice guidelines endorsed by the American Association of hip and Knee Surgeons, American Society of Regional Anesthesia and Pain Medicine, American Academy of Orthopaedic surgeons, hip society, and knee society. J Arthroplasty. 2018;33:3065–9.] raised problems about the lack of high-level evidence to support the use of tranexamic acid in patients with a history of thromboembolic events. Please clarify more concisely whether these patients were included or not.

Answer: Thanks for your valuable suggestion.

The participants of this study will be enrolled according to the inclusion/exclusion criteria. Patients with a coronary artery stent placement or bypass history or a prothrombotic condition will be excluded.

Thank you!

7. Line 148- Recently, many institutions do not use drain for TKA. Why does your institution use it? Clarify the reason.

Answer: Thanks for your valuable suggestion.

To our best knowledge, there is no guideline to prompt the abandon of drain for TKA. Previously, a prospective study has indicated that drains do not reduce joint effusion but do reduce haematoma formation [J Bone Joint Surg Br. 2010 Jan;92(1):51-5.]. Recently, a systematic review and meta-analysis indicated that TXA plus drain-clamping is an efficient method for controlling blood loss after TKA [Int J Surg. 2018 Apr;52:334-341.]. Therefore, our hospital still use drainages for TKA.

Thank you!

Replies to Reviewer 2

1) There is a certain concern over the undesirable influence of high TXA concentrations on synovialocytes, tenocytes and perhaps other cells; this is why it is necessary for the authors to clarify how it is exactly planned to namely perform the periarticular application of TXA because they plan closing the joint first and then apply TXA into the periarticular tissues – where exactly? to what depth? into retinacula, or into the back part of the joint too?

Answer: Thanks for your valuable suggestion.

The doses of TXA were comparable to concentrations reported in previous studies of IV-TXA in TKA, showing concentration of 10–100 mg/mL for topical TXA solutions [Int Orthop. 2011;35:1639–45. & J Bone Joint Surg Am. 2013;95:1961–8.]. In our study, the concentration of TXA will be 40mg/mL (3000mg diluted in 75mL NS), thus the TXA dose is safe for periarticular injection. There is a mistake in our manuscript, we are so sorry for this. The TXA will be injected prior to capsule closure. The injected area includes the medial and lateral capsule, the quadriceps muscle tendon, and the infrapatellar fat pad. This part has been revised in manuscript. Line 192-194.

Thank you!

2) Beside evaluation of benefits of TXA administration, the authors plan to evaluate harms too – however, I am not sure whether a sample of 50 patients (in each shoulder of RCT) receiving DVT prevention allows identification of differences in wound healing impairments, DVT, or even pulmonary embolism – studies focused on these events must be performed on much greater numbers of patients.

Answer: Thanks for your valuable suggestion.

Previous study indicated that mean time from surgery to postdischarge symptomatic VTE was 17.7 days for the TKA [Vasc Health Risk Manag. 2018 May 8;14:81-89.]. To our knowledge, most studies use clinical observation to screen VTE. In our study, all patients will receive vascular ultrasonography to screen VTE before discharge, during follow-up period, patients will be invited to the outpatient clinic 2 weeks, 1, 3 and 6 months after the operation to assess and record the complications and mortality. The vascular ultrasonography will also be performed at the follow-up time point. Therefore, our study will identify all symptomatic and asymptomatic VTEs and complications. Line 226-228.

Thank you!

3) The authors should also state on which postoperative day the patients are usually discharged as they state that ultrasound examination of lower extremity blood vessels will be performed before discharge. If a thrombus takes several days to develop, it is possible that it will not be detected by ultrasound at all under the designed protocol, or that the researchers will learn about its symptomatic forms during a check-up 1 or more months later, and the asymptomatic forms will pass completely unnoticed.

Answer: Thanks for your valuable suggestion.

Previous study indicated that mean time from surgery to postdischarge symptomatic VTE was 17.7 days for the TKA [Vasc Health Risk Manag. 2018 May 8;14:81-89.]. In our study, all patients will be discharged within 3 to 5 days, postoperatively, if no complication occurs, all patients will receive vascular ultrasonography to screen VTE before discharge, during follow-up period, patients will be invited to the outpatient clinic 2 weeks, 1, 3 and 6 months after the operation to assess and record the complications and mortality. The vascular ultrasonography will also be performed at the follow-up time

point. Therefore, our study will identify all symptomatic and asymptomatic VTEs and complications. Line 224-228.

Thank you!

4) I do not like the introduction in Discussion, the 1st paragraph – blood management (BM) is part of all the TKA surgery protocols (and those of many other big surgeries); I cannot see a reason for emphasizing the unique relation between fast-track surgery and blood management because it does simply not exist. I would start with "BM is an inseparable part of all big orthopaedic surgeries, including TKA implantation …"

Answer: Thanks for your valuable suggestion.

This part has been changed to "Blood management is an inseparable part of all big orthopaedic surgeries, including TKA procedure". Line 252-254.

Thank you!

5) I cannot see why the authors do not mention the combined way of TXA administration designed and tested for the potential synergy effect and decrease of the risk of complications. This procedure is also supported by a rationale and backed by literature (both RCTs and RTC meta-analyses).

Answer: Thanks for your valuable suggestion.

Nielsen, et al. have demonstrated the combined intravenous and topical administration of TXA to be effective. In Nielsen's study, only one topical administration (capsule injection) combined with intravenous（IV） administration was investigated comparing to intravenous administration [J Bone Joint Surg Am. 2016 May 18;98(10):835-41.]. Other topical administrations combined with IV administration still need to be investigated. Our study includes IV and three topical administrations of TXA, if we define the most effective topical administration among three topical administrations, the result may provide a direction for further research on combined administration of TXA.

Thank you!

Minor comments:

1) Some grammatical and typographical mistakes.

Answer: Thanks for your valuable suggestion.

The quality of the language has been improved and revised under the assistant of one of our English speaking friends.

Thank you!

2) VTE is explained in the Introduction (it is misspelt as TVE on page 4, line 73) and then again introduced on line 222.

Answer: Thanks for your valuable suggestion.

This part in discussion has been deleted and revised carefully. And the misspelt mistake has been revised, we are very sorry for this mistake! Line 104 and line 262-266.

Thank you!

3) I cannot agree with the sentence on the 240: "... the results of this study are not physician or implant dependent." I believe that surgery outcomes are always dependent on the surgeon – the correct formulation should be that variability of surgery outcomes is lowered by a single surgeon ...

Answer: Thanks for your valuable suggestion.

We agree with your suggestion! The sentence has been revised to "All TKA procedures will be performed by one surgeon, and total knee arthroplasty is performed using a single arthroplasty system; therefore, the results of this study are not implant dependent, and the surgery outcomes is lowered by a single surgeon." Line 50-53.

Thank you!

4) Will the samples really be frozen at 280°C? (lines 177–78).

Answer: We are so sorry for this mistake! The plasma will be frozen and stored at -80 ℃. Line 211-212.

Thank you!

VERSION 2 – REVIEW

REVIEWER	Sachiyuki Tsukada Hokusuikai Kinen Hospital, JAPAN
REVIEW RETURNED	11-Jul-2020

GENERAL COMMENTS	The reviewer is pleased to inform the author that no further questions have arisen.

REVIEWER	Gregory J. Stoddard University of Utah School of Medicine, USA
REVIEW RETURNED	31-Oct-2020

GENERAL COMMENTS	The authors have a very nicely written manuscript. Comparing five groups in the same study is acceptable. The authors state, "For statistical analyses, a professional statistician is being consulted." The lack of a statistician's input at the study design stage is apparent. The following five issues regarding their statistical approach need to be addressed to have a properly designed study. (1) The study is randomizing TKA patients to one of five TXA applications (5 groups). The authors state the study compares the 5 groups to determine the most effective strategy. They state that their primary outcome is total blood loss (TBL), and that analysis of variance (ANOVA) will be used to compare TBL between the groups. The ANOVA approach is not helpful. It gives a single p value, which if significant, supports the conclusion that there is a difference in mean TBL somewhere among the groups. It provides no useful information, however, since it fails to identify which group differs from which other group. So, it does not help the authors determine the most effective strategy, which the authors state is their goal. The authors need to present a statistical approach that matches the goal of the study. (2) For their sample calculation, the authors provide a noninferiority analysis approach for transfusion rate, which appears to be the proportion, or percent, of the sample who required a blood transfusion. They state, "For this primary end point…" The sample size determination should be based on the primary endpoint. So, they need to decide which is really their primary endpoint. If they want multiple primary endpoints, a sample size determination for each should be provided. If the study is underpowered for a primary endpoint, this will have to be mentioned when they report their results, so they should have that power analysis prepared at the study design stage. (3) The authors cite Abdel et al. (2018) [their reference] in their sample size section. Abdel et al. (2018) based their sample size on the outcome TBL, where they powered for a relative 10% difference, ending up with a required N=320 per group. It would appear the authors of this manuscript under review are going to be underpowered for their primary outcome of TBL, since they are only using N=50 per group. -------- 29. Abdel MP, Chalmers BP, Taunton MJ, et al. Intravenous Versus Topical Tranexamic Acid in Total Knee Arthroplasty: Both Effective in a Randomized Clinical Trial of 640 Patients. J Bone Joint Surg Am 2018;100(12):1023-29. (4) In their sample size section, they provide a sample size based on a noninferiority analysis, where their noninferiority margin is an absolute 10% difference. An absolute difference of 10% on a binary outcome is very different from a relative 10% on a continuous

	outcome (Abdel et al. used relative 10% for TBL). Given that transfusion rates in the Abdel et al. (2018) paper where 1.6% and 0.6%, so 1/100 patients, it seems unlikely that 10%, or 1/10 patients, which is a 10-fold increase, would be an acceptable transfusion rate by their profession. They should provide a justification for this noninferiority margin, or realize it is too wide to be accepted as a noninferiority criteria. The justification would be along of the lines of "knowing they could use a method that has 1/100 risk for blood transfusion, choosing a method that increases the risk to 1/10 is of no clinical importance, since a blood transfusion has only minimal risk of harm to begin with." If they do not think their profession would agree with that statement, then their noninferiority comparison is invalid. (5) The authors state they will use the chi-square test to compare transfusion rates. The chi-square test is not a noninferiority test—it only tests for superiority of one method to another. So, they clearly do not have a noninferiority testing approach to go along with their noninferiority sample size determination.

REVIEWER	David C. Hoaglin University of Massachusetts Medical School, USA
REVIEW RETURNED	06-Nov-2020

GENERAL COMMENTS	Reviewer's Comments on BMJOpen-2020-038399 A prospective, randomized, controlled study on the efficacy and safety of different strategies of tranexamic acid with total blood loss, blood transfusion rate and thrombogenic biomarkers in total knee arthroplasty: study protocol By Yong Yang, Zheng Wang, Xin Zhao, Kaijie Yang, Jinlong He, Yun Jin, Haibo Yang, Dong Ding, and Qunhua Jin The authors nicely summarize the gaps in evidence on the efficacy and safety of the available strategies for using TXA in TKA, and they present a well-structured protocol, but I am concerned about four aspects of their trial and protocol. First, a sample size of 50 patients in each of the five groups (actually 39 per group, after allowing for loss to follow-up) seems rather small, especially if the aim is 99% power. I have not done a sample-size calculation for a noninferiority trial with multiple groups, and I was disappointed that the manuscript did not include the details of the sample size calculation or even a reference to an article in the statistical literature describing the method used. Also, I did not see an explanation of why the sample size calculation is based on transfusion rate, a secondary outcome, rather than on the primary outcome, total blood loss. Second, blinding seems problematic. Lines 166 and 167 say that various personnel will be blinded to group allocation. However, after reading the description of the interventions (lines 176 to 190), I do not understand how that can be accomplished. For the placebo group and the IV group, labeling the container as "study solution" would suffice; the rest of the procedure is the same. But the procedures in the TI, PI, and DI groups differ (from one another and from the placebo and IV groups) in essential ways that would seem to make blinding impossible. It may be necessary to avoid describing the study as "blinded." The terms "randomized" and "controlled" still

apply. For the randomization, I suggest that the authors use some type of blocking to provide balance of the sample sizes among the groups as the study proceeds.

Third, a complete protocol includes a reasonably detailed statistical analysis plan. The present protocol, however, does not include an SAP, nor does it say where one can be found, as required by Item 20a in the SPIRIT Checklist.

A solid SAP, however, is only the start of the statistical component. In the interest of transparency and reproducibility, authors should follow the long-standing advice of the International Committee of Medical Journal Editors: Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. The protocol should also include a commitment to this level of documentation in manuscripts resulting from the study. For example, they should have a supplemental file that includes the SPSS commands and the resulting output.

In summarizing continuous variables, the SAP could go beyond the customary (but inadequate) mean and standard deviation: routinely provide the median, the quartiles (not simply their difference, the IQR), and the minimum and maximum. Graphical displays, such as a figure showing parallel boxplots, may be helpful.

The authors plan to use ANOVA to compare various continuous variables (including plasma PAI-1, TAT, and F1+2 levels) among the groups and, separately, to use Student's t tests to compare plasma PAI-1, TAT, and F1+2 before and after the operation in each group. This approach is unnecessarily fragmented. A better analysis of plasma PAI-1, TAT, and F1+2 would use analysis of covariance with the values after the operation as the outcome variable and the values before the operation as a covariate (along with other covariates, if relevant). A similar approach might be helpful for other outcomes. Also, some variables may produce a better analysis in a transformed scale (e.g., a logarithmic scale).

The possibility of including covariates arises also for the variables for which the authors plan to use chi-square tests. The analysis would then be based on logistic regression.

The SAP should acknowledge the issues of multiple comparisons that arise in comparing the various interventions against one another (lines 23 and 24).

Fourth, the authors should re-examine the implications of the statements in lines 47 to 50. Because all the TKA procedures are performed by the same surgeon (QJ, line 171), the study does not have to consider differences among surgeons as a source of variability (e.g., a variance component in the ANOVAs). On the other hand, the relation of the results to other surgeons will not be clear (e.g., those who have less experience than QJ may have higher complication rates). The same comment applies to the use of a single prosthetic system. Why will the results of the study not be "implant dependent" (line 49)? These features of the study are necessary and unavoidable limitations, but, to an extent, they are also strengths. They should be discussed more clearly.

I also have a number of minor comments.

	Line 22: Shouldn't "effective" be "efficacious"? Line 28 has "efficacy". Lines 31 to 33: In Table 2 the secondary outcomes include plasma D-dimer. Line 49: I do not understand the meaning of "lowered" here. Line 88: "potency" seems to be the wrong word here. Line 97: "comprised of" should be "composed of" (incorrect use of "comprised"). Line 100: The methods in Reference 22 have serious flaws. It is not acceptable to choose between a fixed-effect analysis and a random-effects analysis on the basis of an estimate such as I2 or the result of a test for heterogeneity. Also, the DerSimonian-Laird method for random-effects meta-analysis can produce biased estimates with falsely high precision (Cornell et al. 2014), and its confidence intervals have below-nominal coverage and are inferior to those produced by another method (IntHout et al. 2014). A meta-analysis published in 2019 should not have ignored those shortcomings. Line 103: Should "TVE" be "VTE"? Line 264: "are not unrelated" should be "are not related" (or "are unrelated"). Starting in the title and continuing throughout, the manuscript frequently uses the word different. This weakness in the writing (shared with many other writers) should be corrected. Overuse of the word different has reached epidemic proportions. It often conveys no more information than saying that the present manuscript has nine different authors. For example, removing different from line 24 and line 28 would not change the meaning of those sentences. The topical applications are clearly different, and so are the strategies. The authors should review all instances of different and keep only the ones that are clearly necessary. In some instances, they can change different to various or to the specific number. References Cornell JE, Mulrow CD, Localio R, et al. (2014). Random-effects meta-analysis of inconsistent effects: a time for change. Annals of Internal Medicine 160:267-270. IntHout J, Ioannidis JPA, Borm GF (2014). The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Medical Research Methodology 14:25.

VERSION 2 – AUTHOR RESPONSE

Reply to Reviewer 1

The reviewer is pleased to inform the author that no further questions have arisen.

Answer: Thank you for your great work on our manuscript. Thank you very much!

Replies to Reviewer 2

The authors state, “For statistical analyses, a professional statistician is being consulted.” The lack of a statistician’s input at the study design stage is apparent.

Answer: Thank you for your valuable suggestion.

Dr. Faxuan Wang, from School of public health, Ningxia Medical University, as a professional statistician, is being consulted. The information of Dr. Faxuan Wang has been added as one of the authors. Line 4, line 8-9 and line 287.

Thank you!

(1) The study is randomizing TKA patients to one of five TXA applications (5 groups). The authors state the study compares the 5 groups to determine the most effective strategy. They state that their primary outcome is total blood loss (TBL), and that analysis of variance (ANOVA) will be used to compare TBL between the groups. The ANOVA approach is not helpful. It gives a single p value, which if significant, supports the conclusion that there is a difference in mean TBL somewhere among the groups. It provides no useful information, however, since it fails to identify which group differs from which other group. So, it does not help the authors determine the most effective strategy, which the authors state is their goal. The authors need to present a statistical approach that matches the goal of the study.

Answer: Thank you for your valuable suggestion.

We have modified the statistical method of our study protocol. Normality of data will be tested by Kolmogorov-Smirnov test. One- way analysis of variance (one- way ANOVA) will be used to compare the TBL, drainage volume, plasma PAI-1, TAT and F1+2 levels, maximum hemoglobin drop and length of hospital stay (LOS) between groups. If there is a significant difference, the comparison between groups will be performed by Scheffe test post hoc analysis. Repeated-measures analysis of variance will be used to compare plasma PAI-1, TAT, F1+2 levels according to time points. Chi square test will be used to compare the blood transfusion rate, VTE and wound complications between groups. This part has been changed in our manuscript. Line 228-240.

Thank you!

(2) For their sample calculation, the authors provide a noninferiority analysis approach for transfusion rate, which appears to be the proportion, or percent, of the sample who required a blood transfusion. They state, “For this primary end point…” The sample size determination should be based on the primary endpoint. So, they need to decide which is really their primary endpoint. If they want multiple primary endpoints, a sample size determination for each should be provided. If the study is underpowered for a primary endpoint, this will have to be mentioned when they report their results, so they should have that power analysis prepared at the study design stage.

Answer: Thank you for your valuable suggestion.

The sample size determination should been based on the primary outcome, the total blood loss (TBL). Therefore, we modified the determination of sample size.

Previous study based on an analysis of a national database with 7133 primary TKA procedures has indicated the total blood loss of intravenous TXA application, topical TXA application and control group were 830±410ml, 970±470ml, and 1200±640ml, respectively [Thromb Res. 2019 Jan;173:96-101.]. Based on this data, the sample size was calculated by PASS 15.0 (NCSS, LLC, Kaysville, UT, USA) with one-way analysis of variance F-tests. We calculated that a total of 39 patients per group to provide a power of 80% to detect it at a significance level of 5%. Considering factors such as loss of follow-up, we expanded the sample size by about 30%, therefore, there will be a minimum of 50 patients in each group. The section of Sample size calculation has been revised in our manuscript. Line 138-146.

Thank you!

(3) The authors cite Abdel et al. (2018) [their reference] in their sample size section. Abdel et al. (2018) based their sample size on the outcome TBL, where they powered for a relative 10% difference, ending up with a required N=320 per group. It would appear the authors of this manuscript under review are going to be underpowered for their primary outcome of TBL, since they are only using N=50 per group.

---

29. Abdel MP, Chalmers BP, Taunton MJ, et al. Intravenous Versus Topical Tranexamic Acid in Total Knee Arthroplasty: Both Effective in a Randomized Clinical Trial of 640 Patients. J Bone Joint Surg Am 2018;100(12):1023-29.

Answer: Thank you for your valuable suggestion.

The sample size was recalculated by PASS 15.0 (NCSS, LLC, Kaysville, UT, USA) with one-way analysis of variance F-tests. The section of Sample size calculation has been revised in our manuscript. Line 138-146.

Thank you!

(4) In their sample size section, they provide a sample size based on a noninferiority analysis, where their noninferiority margin is an absolute 10% difference. An absolute difference of 10% on a binary outcome is very different from a relative 10% on a continuous outcome (Abdel et al. used relative 10% for TBL). Given that transfusion rates in the Abdel et al. (2018) paper where 1.6% and 0.6%, so 1/100 patients, it seems unlikely that 10%, or 1/10 patients, which is a 10-fold increase, would be an acceptable transfusion rate by their profession. They should provide a justification for this noninferiority margin, or realize it is too wide to be accepted as a noninferiority criteria. The justification would be along of the lines of "knowing they could use a method that has 1/100 risk for blood transfusion, choosing a method that increases the risk to 1/10 is of no clinical importance, since a blood transfusion has only minimal risk of harm to begin with." If they do not think their profession would agree with that statement, then their noninferiority comparison is invalid.

Answer: Thank you for your valuable suggestion.

The primary end point of this study is total blood loss, therefore, the sample size calculation was modified based on our primary outcome. The sample size was recalculated by PASS 15.0 (NCSS, LLC, Kaysville, UT, USA) with one-way analysis of variance F-tests. The section of Sample size calculation has been revised in our manuscript. Line 138-146.

Thank you!

(5) The authors state they will use the chi-square test to compare transfusion rates. The chi-square test is not a noninferiority test—it only tests for superiority of one method to another. So, they clearly

do not have a noninferiority testing approach to go along with their noninferiority sample size determination.

Answer: Thank you for your valuable suggestion.

The sample size was recalculated by PASS 15.0 (NCSS, LLC, Kaysville, UT, USA) with one-way analysis of variance F-tests. Therefore, the chi-square test is available to compare transfusion rates. Line 138-146.

Thank you!

Replies to Reviewer 2

-First, a sample size of 50 patients in each of the five groups (actually 39 per group, after allowing for loss to follow-up) seems rather small, especially if the aim is 99% power. I have not done a sample-size calculation for a noninferiority trial with multiple groups, and I was disappointed that the manuscript did not include the details of the sample size calculation or even a reference to an article in the statistical literature describing the method used. Also, I did not see an explanation of why the sample size calculation is based on transfusion rate, a secondary outcome, rather than on the primary outcome, total blood loss.

Answer: Thank you for your valuable suggestion.

The primary outcome of this study is the total blood loss, and the sample size determination should been based on the primary outcome. Therefore, we modified the determination of sample size as following:

Previous study based on an analysis of a national database with 7133 primary TKA procedures has indicated the total blood loss of intravenous TXA application, topical TXA application and control group were 830±410ml, 970±470ml, and 1200±640ml, respectively [Thromb Res. 2019 Jan;173:96-101.]. Based on this data, the sample size was calculated by PASS 15.0 (NCSS, LLC, Kaysville, UT, USA) with one-way analysis of variance F-tests. We calculated that a total of 39 patients per group to

provide a power of 80% to detect it at a significance level of 5%. Considering factors such as loss of follow-up, we expanded the sample size by about 30%, therefore, there will be a minimum of 50 patients in each group.

The section of Sample size calculation has been revised in our manuscript. Line 138-146.

Thank you very much!

-Second, blinding seems problematic. Lines 166 and 167 say that various personnel will be blinded to group allocation. However, after reading the description of the interventions (lines 176 to 190), I do not understand how that can be accomplished. For the placebo group and the IV group, labeling the container as “study solution” would suffice; the rest of the procedure is the same. But the procedures in the TI, PI, and DI groups differ (from one another and from the placebo and IV groups) in essential ways that would seem to make blinding impossible. It may be necessary to avoid describing the study as “blinded.” The terms “randomized” and “controlled” still apply. For the randomization, I suggest that the authors use some type of blocking to provide balance of the sample sizes among the groups as the study proceeds.

Answer: Thank you for your valuable suggestion.

In present study, the patients in the TI, PI and DI groups will also receive 100 ml normal saline intravenously ten minutes prior to skin incision, 3h and 6h postoperatively. Therefore, the blinding of this study would be possible. This part has been added in our manuscript. For the randomization, a computer-based randomization system will be used to screen and randomize the patients, and the recruiting of patients will be ended until all groups meet at least 50 patients. Line187-188.

Thank you very much!

Third, a complete protocol includes a reasonably detailed statistical analysis plan.

The present protocol, however, does not include an SAP, nor does it say where one

can be found, as required by Item 20a in the SPIRIT Checklist.

Answer: Thank you for your valuable suggestion.

The section Statistical analysis has been modified in our manuscript, which includes a statistical analysis plan as following:

Normality of data will be tested by Kolmogorov-Smirnov test. One- way analysis of variance (one- way ANOVA) will be used to compare the TBL, drainage volume, plasma PAI-1, TAT and F1+2 levels, maximum hemoglobin drop and length of hospital stay (LOS) between groups. If there is a significant difference, the comparison between groups will be performed by Scheffe test post hoc analysis. Repeated-measures analysis of variance will be used to compare plasma PAI-1, TAT, F1+2 levels according to time points. Chi square test will be used to compare the blood transfusion rate, VTE and wound complications between groups. Line 228- 240.

Thank you!

Fourth, the authors should re-examine the implications of the statements in lines 47

to 50. Because all the TKA procedures are performed by the same surgeon (QJ, line

171), the study does not have to consider differences among surgeons as a source of

variability (e.g., a variance component in the ANOVAs). On the other hand, the

relation of the results to other surgeons will not be clear (e.g., those who have less

experience than QJ may have higher complication rates). The same comment applies

to the use of a single prosthetic system. Why will the results of the study not be

“implant dependent” (line 49)? These features of the study are necessary and

unavoidable limitations, but, to an extent, they are also strengths. They should be

discussed more clearly.

Answer: Thank you for your valuable suggestion.

Although there is no consensus regarding the influence of surgeon experience on TKA [J Arthroplasty. 2018 Apr;33(4):1231-1234. & J Arthroplasty. 2001 Aug;16(5):635-40. & J Long Term Eff Med Implants. 2003;13(5):389-97. & Knee Surg Relat Res. 2020 Jan 1;32(1):3.]. To our understanding and experience, more experienced surgeons lead to shorter operative time, shorter duration of tourniquet. An increased operative time leads to higher blood loss and transfusion rates [Orthop Clin North Am.]. A short duration tourniquet during TKA gives better symptomatic pain relief in the early postoperative period as compared to long duration use of tourniquet [J Clin Orthop Trauma. Jan-Mar 2018;9(1):46-50.]. Previous study has demonstrated that the use of closed- and open-box knee prostheses resulted in a significant difference in blood loss in simultaneous bilateral total knee arthroplasty [Clin Orthop Surg. 2019 Dec;11(4):409-415.], the single prosthetic system may lower this effect. Therefore, we still consider the surgery outcomes are lowered by a single surgeon, and the results of this study are not implant dependent.

Thank you very much!

- Line 22: Shouldn't "effective" be "efficacious"? Line 28 has "efficacy".

Answer: Thank you for your valuable suggestion.

The sentence "The aim of this trial is to investigate the most effective delivery method….." has been changed to "The aim of this trial is to investigate the most efficacious delivery method…..". Line 24.

Thank you very much!

-Lines 31 to 33: In Table 2 the secondary outcomes include plasma D-dimer.

Answer: Thank you for your valuable suggestion.

D-dimer, as a VTE detection marker, is been added in our manuscript. Line 27, 32, 61, 67, 102, 107, 113, 122, 203, 234, 254 and line 260.

Thank you very much!

-Line 49: I do not understand the meaning of "lowered" here.

Answer: The surgery outcomes are always dependent on the surgeon, thus the variability of surgery outcomes is lowered by a single surgeon.

Thank you very much!

-Line 88: “potency” seems to be the wrong word here.

Answer: Thank you for your valuable suggestion.

The word “potency” has been changed to “possibility”. Line 88.

Thank you very much!

-Line 97: “comprised of” should be “composed of” (incorrect use of “comprised”).

Answer: Thank you for your valuable suggestion.

The manuscript has been revised. Line 97.

Thank you very much!

-Line 100: The methods in Reference 22 have serious flaws. It is not acceptable to choose between a fixed-effect analysis and a random-effects analysis on the basis of an estimate such as I2 or the result of a test for heterogeneity. Also, the DerSimonian-Laird method for random-effects meta-analysis can produce biased estimates with falsely high precision (Cornell et al. 2014), and its confidence intervals have below-nominal coverage and are inferior to those produced by another method (IntHout et al. 2014). A meta-analysis published in 2019 should not have ignored those shortcomings.

Answer: Thank you for your valuable suggestion.

The Reference 22 has been removed from References. Line 102.

Thank you very much!

-Line 103: Should “TVE” be “VTE”?

Answer: Thank you for your valuable suggestion.

The manuscript has been revised. Line 103.

Thank you very much!

-Line 264: “are not unrelated” should be “are not related” (or “are unrelated”).

Answer: Thank you for your valuable suggestion.

The manuscript has been changed. Line 263.

Thank you very much!

-Starting in the title and continuing throughout, the manuscript frequently uses the word different. This weakness in the writing (shared with many other writers) should be corrected. Overuse of the word different has reached epidemic proportions. It often conveys no more information than saying that the present manuscript has nine different authors. For example, removing different from line 24 and line 28 would not change the meaning of those sentences. The topical applications are clearly different, and so are the strategies. The authors should review all instances of different and keep only the ones that are clearly necessary. In some instances, they can change different to various or to the specific number.

Answer: Thank you for your valuable suggestion.

We have checked our manuscript carefully, and many “different” has been changed or removed.

Thank you!

VERSION 3 – REVIEW

REVIEWER	Sachiyuki Tsukada Hokusuikai Kinen Hospital
REVIEW RETURNED	05-Jan-2021

GENERAL COMMENTS	The reviewer considers that the manuscript has been revised appropriately.

REVIEWER	Gregory J. Stoddard University of Utah, United States
REVIEW RETURNED	11-Jan-2021

GENERAL COMMENTS	You have responded favorably to all of my previous concerns. Thank you for making that effort. I have no further suggestions.

REVIEWER	David C. Hoagllin University of Massachusetts Medical School, USA
REVIEW RETURNED	20-Jan-2021

GENERAL COMMENTS	Reviewer's Comments on BMJOpen-2020-038399.R2 A prospective, randomized, controlled study on the efficacy and safety of different strategies of tranexamic acid with total blood loss, blood transfusion rate and thrombogenic biomarkers in total knee arthroplasty: study protocol By Yong Yang, Zheng Wang, Faxuan Wang, Xin Zhao, Kaijie Yang, Jinlong He, Yun Jin, Haibo Yang, Dong Ding, and Qunhua Jin I thank the authors for their responses to my comments on the previous version. Before commenting on the current version, I would like to correct a misunderstanding introduced by another reviewer, who wrote, "The ANOVA approach is not helpful. It gives a single p value, which if significant, supports the conclusion that there is a difference in mean TBL somewhere among the groups. It provides no useful information, however, since it fails to identify which group differs from which other group." No careful analysis focuses solely on a single p-value. One-way analysis of variance provides a framework for comparing the data from the groups, more broadly than by calculating the mean in each group. Modern applications of ANOVA begin by examining the data (e.g., in parallel plots for the groups). Among other features, that step can reveal potential outliers, patterns that may make group means an inappropriate summary, and a tendency for the spread of the data to vary among the groups (sometimes in ways that suggest analyzing the data in a transformed scale). If appropriate, a one-way ANOVA is a classic example of "borrowing strength"—combining the deviations from the groups to obtain a more-stable estimate of the within-group variance. If the analysis points to differences among the group means, a rich variety of procedures are available (under the general heading of "multiple comparisons") for determining which differences should be considered "significant." Some of those procedures produce "simultaneous confidence intervals" in addition to the usual "individual confidence intervals." I turn now to my comments on the current version. One of my comments on the previous version dealt with the sample size calculation. Fortunately, the authors now base that calculation on the primary outcome (TBL) and are not trying to use some sort of noninferiority approach. I asked for details of the sample size calculation. The current version (line 139ff) provides more information, but it still does not include important details. In another comment, I stressed the need to describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. I am disappointed that the authors did not show their understanding of this important advice by applying it to the sample size calculation. The data from a previous study are valuable information. The authors say (line 142), "Based on this data, the sample size was calculated by PASS 15.0." They should have reported, specifically, how they used those data as an input for PASS (for example, what command or routine in PASS did they use, with what arguments and options?). The result of the calculation was "a total of 39 patients per group to provide a power of 80% to detect it at a significance level of 5%." They do not,

	however, define a seemingly minor but key word in that sentence: it. What did they actually ask PASS to do? Further, under Statistical analysis they say (lines 237 and 238), “If there is a significant difference, the comparison between groups will be performed by Scheffe test post hoc analysis.” Did the sample size calculation take this possibility into account? A related issue is whether the authors intend to determine “the most efficacious delivery method of TXA” (lines 24 and 25, in the Abstract; also “most effective” in line 259) or only “compare the efficacy of various strategies of TXA” (line 110). Picking the most efficacious method will require statistical techniques specifically designed for ranking and, probably, a different sample size calculation. The article by Gibbons et al. (1979) gives an introduction to ranking and selection. I am not yet convinced that the blinding will be satisfactory. In response to my comment on the previous version, the authors explained that “the patients in the TI, PI and DI groups will also receive 100 ml normal saline intravenously ten minutes prior to skin incision, 3h and 6h postoperatively.” That information is helpful. However, the description of the three topical applications does not explain who will carry out the topical irrigation in TI, the injection into periarticular tissue in PI, or the injection into joint cavity through drainage (after wound closure) in DI. It seems that these steps will reveal the patient’s group assignment. In another comment on the previous version, I mentioned the need to include “a reasonably detailed statistical analysis plan.” The current section on Statistical analysis (line 228ff) is only a summary; it is far from a reasonably detailed SAP. As one small example, lines 253 and 254 mention “comparison of intravenous and topical applications, and comparison of three topical applications,” but the section on Statistical analysis does not explain how the topical applications will be represented in the comparison of intravenous and topical applications or give the details of the comparison of the three topical applications. Will the first of these use the average of the three topical applications? Will the second use all pairwise comparisons? Such details (and many others) should be stated in advance. In line 219, what is “the follow-up time point”? Line 217 mentions four times. In Table 2 it would be helpful to include the units for each measure. Reference Gibbons JD, Olkin I, Sobel M (1979). An introduction to ranking and selection. The American Statistician, 33(4):185-195.

VERSION 3 – AUTHOR RESPONSE

Reply to Reviewer 4

- One of my comments on the previous version dealt with the sample size calculation. Fortunately, the authors now base that calculation on the primary outcome (TBL) and are not trying to use some sort of noninferiority approach. I asked for details of the sample size calculation. The current version (line 139ff) provides more information, but it still does not include important details. In another comment, I stressed the need to describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results. I am disappointed that the authors did not show their understanding of this important advice by applying it to the sample size calculation. The data from a previous study are valuable information. The authors say (line 142), "Based on this data, the sample size was calculated by PASS 15.0." They should have reported, specifically, how they used those data as an input for PASS (for example, what command or routine in PASS did they use, with what arguments and options?). The result of the calculation was "a total of 39 patients per group to provide a power of 80% to detect it at a significance level of 5%." They do not, however, define a seemingly minor but key word in that sentence: it. What did they actually ask PASS to do? Further, under Statistical analysis they say (lines 237 and 238), "If there is a significant difference, the comparison between groups will be performed by Scheffe test post hoc analysis." Did the sample size calculation take this possibility into account? A related issue is whether the authors intend to determine "the most efficacious delivery method of TXA" (lines 24 and 25, in the Abstract; also "most effective" in line 259) or only "compare the efficacy of various strategies of TXA" (line 110).Picking the most efficacious method will require statistical techniques specifically designed for ranking and, probably, a different sample size calculation. The article by Gibbons et al. (1979) gives an introduction to ranking and selection.

I am not yet convinced that the blinding will be satisfactory. In response to my comment on the previous version, the authors explained that "the patients in the TI, PI and DI groups will also receive 100 ml normal saline intravenously ten minutes prior to skin incision, 3h and 6h postoperatively." That information is helpful. However, the description of the three topical applications does not explain who will carry out the topical irrigation in TI, the injection into periarticular tissue in PI, or the injection into joint cavity through drainage (after wound closure) in DI. It seems that these steps will reveal the patient's group assignment. In another comment on the previous version, I mentioned the need to

include “a reasonably detailed statistical analysis plan.” The current section on Statistical analysis (line 228ff) is only a summary; it is far from a reasonably detailed SAP. As one small example, lines 253 and 254 mention “comparison of intravenous and topical applications, and comparison of three topical applications,” but the section on Statistical analysis does not explain how the topical applications will be represented in the comparison of intravenous and topical applications or give the details of the comparison of the three topical applications. Will the first of these use the average of the three topical applications? Will the second use all pairwise comparisons? Such details (and many others) should be stated in advance.

Answer: Thank you for your great work on our manuscript.

(1) The sample size of this study was calculated by PASS 15.0 (NCSS, LLC, Kaysville, UT, USA) with one-way analysis of variance F-tests. The input information is following:

Design Tab

Solve For:	*Sample Size*
Power:	*0.8*
Alpha:	*0.05*
G (Number of Groups):	*5*
Group Allocation Ratios:	*Equal*
Input σm Using:	*List of means (μi's) from which σm is calculated*
Means (μ1, μ2, ..., μG):	*1.2 0.97 0.97 0.97 0.83*
K (Means Multiplier):	*1*
σ (Standard Deviaton):	*0.64 0.47 0.47 0.47 0.41*

and the calculated output is following:

Numeric Results
Means: 1.2 0.97 0.97 0.97 0.83

Power	*Average n*	*G*	*Total N*	*K*	*Std Dev of Means σm*	*Standard Deviation σ*	*Effect Size*	*Alpha*
0.8113	*30.00*	*5*	*150*	*1.00*	*0.12*	*0.41*	*0.2904*	*0.0500*
0.8100	*39.00*	*5*	*195*	*1.00*	*0.12*	*0.47*	*0.2533*	*0.0500*
0.8100	*39.00*	*5*	*195*	*1.00*	*0.12*	*0.47*	*0.2533*	*0.0500*
0.8100	*39.00*	*5*	*195*	*1.00*	*0.12*	*0.47*	*0.2533*	*0.0500*
0.8005	*70.00*	*5*	*350*	*1.00*	*0.12*	*0.64*	*0.1860*	*0.0500*

References
Desu, M. M. and Raghavarao, D. 1990. Sample Size Methodology. Academic Press. New York.
Fleiss, Joseph L. 1986. The Design and Analysis of Clinical Experiments. John Wiley & Sons. New York.
Kirk, Roger E. 1982. Experimental Design: Procedures for the Behavioral Sciences. Brooks/Cole. Pacific Grove,
California.

Report Definitions
Power is the probability of rejecting a false null hypothesis. It should be close to one.

n is the average group sample size.
G is the number of groups.
Total N is the total sample size of all groups combined.
K is the group means multiplier.
σm is the standard deviation of the group means under the alternative hypothesis.
σ is the within group standard deviation.
The Effect Size is the ratio of σm and σ.
Alpha is the probability of rejecting a true null hypothesis. It should be small.

Summary Statements
In a one-way ANOVA study, sample sizes of 30, 30, 30, 30, and 30 are obtained from the 5 groups whose means are to be compared. The total sample of 150 subjects achieves 81% power to detect differences among the means versus the alternative of equal means using an F test with a 0.0500 significance level. The size of the variation in the means is represented by their standard deviation which is 0.12. The common standard deviation within a group is assumed to be 0.41.

(2) For the blinding, this study is a single-blinded trial. The manuscript has been revised. Line 162.

Thank you very much!

(3) The statistical analysis plan of this study is following:

Statistical Analysis Plan

Objectives

The primary objective of this study is to investigate the most effective administration from intravenous and three topical applications of tranexamic acid (TXA).

The secondary objective of this study is to compare the efficacy and safety of three topical applications of TXA.

Design and methods

Design

This study is a prospective, single-center, parallel-group, single-blinded randomized controlled trial. The study compares total blood loss, blood transfusion rate and drainage volume between different administrations of TXA (comparison of intravenous and topical applications, and comparison of three topical applications), and to determine the most effective strategy of TXA. The study also investigates the safety of TXA strategies in terms of the effect of TXA on the plasma D-dimer, PAI-1, TAT and F1+2 levels and wound complications, length of hospital stay, deep vein thrombosis (DVT) and pulmonary embolism (PE). Randomization will be performed with a 1:1:1:1:1 allocation into five groups: placebo group, intravenous group (IV); topical irrigation group (TI); periarticular tissue injection group (PI); and drainage injection group (DI).

Patient eligibility criteria

Inclusion criteria:

- *Undergoing primary TKA, of both genders*
- *>18 and <100 years at time of inclusion*

Exclusion criteria:

- *Allergy to TXA*
- *Preoperative hepatic or renal dysfunction*
- *Serious cardiac or respiratory disease, including coronary artery stent placement or bypass*
- *Congenital or acquired coagulopathy, as evidenced by an international normalized ratio (INR) of >1.4 or a partial thromboplastin time (PTT) of >1.4 times normal*
- *A preoperative platelet count of <150,000/mm^3*
- *History of a prothrombotic condition*
- *Pregnancy or breastfeeding*
- *Diagnosis of inflammatory arthritis*
- *A preoperative hemoglobin level of <10 g/dL*

Randomization and blinding

A computer-based randomization system will be used to screen and randomize the patients one day before operation. Randomization allocation forms will be sealed in non-transparent envelopes by a person not involved in the study. The envelopes will be kept in a locked cabinet at the surgical unit. Single envelopes will be opened and thus patients will be randomized by the anaesthesia nurse no earlier than 2 hours prior to the surgery. Patients, anesthesiologists, and research assistants collecting data will be blinded to group allocation.

Intervention

All TKA procedures will be performed by one surgeon (QJ), under general anaesthesia, with tourniquets. A standard midline skin incision and medial parapatellar arthrotomy approach will be used. Standard surgical techniques for intraoperative hemostasis will be performed. All output will be measured and recorded in milliliters. All drains will be clamped for 2 hours and removed 24 hours after placement.

For patients in the placebo group, 100 ml normal saline (NS) will be applied intravenously ten minutes prior to skin incision, and 100 ml NS will be applied intravenously at 3h and 6h postoperatively.

For patients in IV group, 1g of TXA in 100ml NS will be applied intravenously ten minutes prior to skin incision, and 1g of TXA in 100ml NS will be applied intravenously at 3h and 6h postoperatively.

For patients in TI group, after cementation of the implant, 3 g of TXA diluted in 75 mL NS solution will be irrigated topically to the open joint surfaces five minutes prior to tourniquet release. The surgeon subsequently suctioned away excess study solution without touching the surrounding tissue surfaces.

For patients in PI group, 3 g of TXA diluted in 75 mL NS solution will be injected to periarticular tissue, including the medial and lateral capsule, the quadriceps muscle tendon, and the infrapatellar fat pad, prior to capsule closure.

For patients in the DI group, after wound closure, 3 g of TXA diluted in 75 mL NS will be injected into joint cavity through drainage.

As placebo, the patients in the TI, PI and DI groups will also receive 100 ml NS intravenously ten minutes prior to skin incision, 3h and 6h postoperatively.

Outcomes

Primary outcome

- *Total blood loss (TBL)*

  Calculation of TBL: according to Nadler's formula (Surgery, 1962,51(2):224-232.)

  TBL（ml）=1000 hemoglobin（loss）/ hemoglobin (pre-op)；*

*hemoglobin（loss）= blood volume *（hemoglobin (pre-op) - hemoglobin (post-op)）*0.001+blood transfusion volume；*

blood volume（L）=[K_1 height^3]+[K_2* weight]+K_3，K_1=0.3669（male）or 0.3561（female），K_2=0.03219（male）or 0.03308（female），K_3=0.6041（male）or 0.1833（female）*

Secondary outcomes

- *Blood transfusion rate (BTR)*

- *Drainage volume*
- *Plasma D-dimer, Plasminogen activator inhibitor-1(PAI-1), Thrombin-antithrombin complexes (TAT) and Prothrombin fragment F1+2 (F1+2)*
- *Wound complications*
- *Venous thromboembolism (VTE)*
- *Length of hospital stay (LOS)*

Sample size calculation

Previous study based on an analysis of a national database with 7133 primary TKA procedures has indicated the total blood loss of intravenous TXA application, topical TXA application and control group were 830±410ml, 970±470ml, and 1200±640ml, respectively (Thromb Res 2019;173:96-101.). Based on this data, the sample size was calculated by PASS 15.0 (NCSS, LLC, Kaysville, UT, USA) with one-way analysis of variance F-tests. We calculated that a total of 39 patients per group to provide a power of 80% to detect it at a significance level of 5%. Considering factors such as loss of follow-up, we expanded the sample size by about 30%, therefore, there will be a minimum of 50 patients in each group.

statistical analysis

The statistical analysis will be performed by using SPSS19.0 software (Statistic Package for Social Science, SPSS, Inc., Chicago, IL, USA). Normality of data will be tested by Kolmogorov-Smirnov test. One - way analysis of variance (one - way ANOVA) will be used to compare the TBL, drainage volume, plasma D-dimer, PAI-1, TAT and F1+2 levels, maximum hemoglobin drop and length of hospital stay (LOS) between groups. If there is a significant difference, the comparison between groups will be performed by Scheffe test post hoc analysis. Repeated-measures analysis of variance will be used to compare plasma PAI-1, TAT, F1+2 levels according to time points. Chi square test will be used to compare the blood transfusion rate, VTE and wound complications between groups.

Significance levels of tests and confidence intervals

All statistical tests will use a two-sided p value of 0.05, unless otherwise specified. There will be no formal adjustment of p values for any interim analyses performed. Two-sided 95% confidence intervals will be presented for all estimates.

Baseline comparability

The baseline continuous variables will be summarized using mean ± standard deviation or median (interquartile range) for continuous variables as appropriate. Categorical variables will be summarized using frequency (percentage). No statistical tests of differences in baseline characteristics between groups will be done, as any differences between treatment arms must be due to chance rather than bias.

Thank you very much!

- In line 219, what is “the follow-up time point”? Line 217 mentions four times.

In Table 2 it would be helpful to include the units for each measure.

Answer: Thank you for your valuable suggestion.

(1) Patients will be invited to the outpatient clinic 2 weeks, 1, 3 and 6 months after the operation to assess and record the complications and mortality. The vascular ultrasonography will also be performed at the follow-up time points. Therefore, the follow-up time point will be 2 weeks, 1, 3 and 6 months postoperatively. Line 218

Thank you very much!

(2) The units for each measure have been added in Table 2. Line 220- 221.

Thank you very much!