



HAL
open science

Évaluation de la qualité de l'acte chirurgical dans les prothèses totales de genoux

Pierre-Alban Bouché

► **To cite this version:**

Pierre-Alban Bouché. Évaluation de la qualité de l'acte chirurgical dans les prothèses totales de genoux. Santé publique et épidémiologie. Université Paris Cité, 2021. Français. NNT : 2021UNIP5270 . tel-04617888

HAL Id: tel-04617888

<https://theses.hal.science/tel-04617888v1>

Submitted on 19 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université de Paris

École doctorale 393, Pierre Louis de Santé Publique
Inserm, UMR1153 Centre de Recherche Épidémiologie et
Statistiques Sorbonne Paris Cité (CRESS), équipe ECSTRRA

Évaluation de la qualité de l'acte chirurgical dans les prothèses totales de genoux

Evaluation of the quality of surgical intervention in total knee arthroplasty

Par Pierre-Alban Bouché

Thèse de doctorat de santé publique
Spécialité : épidémiologie clinique

Dirigée par le Pr Pr Rémy Nizard
Et par le Pr Matthieu Resche-Rigon

Présentée et soutenue publiquement le 10 décembre 2021

Devant un jury composé de :

M. le Professeur Didier HANNOUCHE - Université de Genève - Rapporteur
Mme. le Docteur Laurence LE CLEACH (MCU-PH) - Université Paris-Est Créteil - Rapporteur
M. le Professeur Jean-Noël ARGENSON - Université Aix-Marseille - Examineur
Mme. la Professeur Agnès DECHARTRES - Sorbonne université - Examineur
M. le Professeur Rémy NIZARD - Université de Paris - Directeur
M. le Professeur Matthieu RESCHE-RIGON - Université de Paris - Codirecteur



Except where otherwise noted, this work is licensed under
<https://creativecommons.org/licenses/by-nd/3.0/fr/>

Remerciements

À M. le Professeur Matthieu Resche-Rigon, directeur de thèse : de m'avoir poussé et soutenu dans la recherche clinique. Merci pour ton expertise et de tes conseils autant professionnels que personnels. Nous avons encore plein de projets en cours et futurs.

À M. le Professeur Rémy Nizard, directeur de thèse : pour votre confiance accordée très vite, votre loyauté et votre sensibilité. Vous m'avez initié et donné goût à la recherche clinique. Vous m'avez ouvert l'esprit sur la dimension humaine de notre métier et fait explorer différent domaine que j'ignorais. Je suis honoré de votre présence, d'avoir effectué ce travail sous votre direction et d'être votre élève.

À M. le Professeur David Biau : pour votre soutien et votre confiance autant chirurgicale que scientifique. Depuis le début de mon internat en chirurgie, j'entendais parler de vous notamment au SBIM. Vous êtes un modèle de parcours.

À M. le Professeur Moussa Hamadouche : pour vos conseils avisés et votre accompagnement aussi bien chirurgical que scientifique. Vous m'avez fait prendre conscience que le plus important était la famille. Je tâcherais de m'efforcer d'appliquer au quotidien vos principes.

À Mme. la Professeur Sylvie Chevret : pour votre enseignement et votre expertise pendant. Merci de m'avoir accueilli dans votre service et initié à la recherche clinique.

À M. le Professeur Didier Hannouche, Mme. la Docteur Laurence Le Cleach, Mr. le Professeur Jean-Noël Argenson et Mme. la Professeur Agnès Dechartres d'avoir accepté de faire parti de ce jury de thèse.

Au docteur Biard, de m'avoir encadré en master I et donner goûts aux biostatistiques. Merci pour ta gentillesse.

Au docteur Guillon pour votre vision de la chirurgie et de la vie. Je vous dois mes premières interventions et mon premier article scientifique. Merci d'avoir été présent toutes ces années et sans défection. Merci pour cette relation d'amitié.

Au Docteur Zahi pour ce semestre à Montfermeil, une de mes plus belles rencontres humaines de mon internat.

Aux Docteurs Auberger, Coré, Lehanneur et Sabatier qui ont su me transmettre leur passion chirurgicale grâce à leurs marques d'amitié et de confiance, que j'espère pouvoir transmettre à mon tour.

Aux étudiants, doctorants et membres du SBIM,

À tous mes compagnons de stage et particulièrement à Adrien Chaud, Jules Descamps, Édouard Gaillot-Drevon, Nicolas Gaujac, Aurélien Hallé, Elena Lang et Diane Petit.

Au futur docteur Simon Corsia. Merci pour tout. Je pense que tu dois être la personne avec qui j'ai passé le plus d'heures ces 6 dernières années. Merci d'avoir autant été présent et disponible notamment pour toutes ces relectures de mails et d'anglais. Mais t'inquiètes pas je ne pense pas m'arrêter là. Je ne pourrais jamais assez te remercier.

À Anna Fain, Henri Flament et Luce Mouchez, amis de toujours. Merci pour votre soutien, vos conseils et tous ces moments passés ensemble. Merci d'être tout simplement présent.

À ma famille et belle famille,

À mes grands-parents, merci d'avoir toujours cru en moi et de me pousser depuis mon plus jeune âge. Merci pour votre éducation et transmission.

À ma sœur, mon beau-frère et mes nièces.

À mes parents, qui sans eux, je n'aurais rien réussi. Merci pour votre éducation et la vie que vous m'avez offert.

À Brune, brubru, brugnon, ma fille, ma fierté. Ma plus grande réussite.

À Mahaut avec qui je partage ma vie depuis plus de 10 ans. Tu m'as toujours soutenu de manière inconditionnelle et sans faille. Merci d'être autant compréhensible et de réussir à accepter tous ces sacrifices. Cette réussite est la tienne.

Sommaire

Listes abrégées	5
Table des figures	6
Résumé	7
Abstract	10
Publication	12
Préambule	15
Partie 1 : Introduction	17
Chapitre 1 - Prothèse totale de genoux	18
Chapitre 2 - Problématique	24
a. Évaluation clinique des prothèses totales de genoux	24
b. Évaluation de l'apprentissage	26
Partie 2 : Évaluation à l'aide de méta-analyses en réseau	28
Chapitre 3 - Principe des méta-analyse	29
Chapitre 4 - Particularité des méta-analyses en réseau	32
a. Représentation du réseau	32
b. Hypothèses à valider	33
c. Analyse statistique	34
d. Classement des traitements	38
e. Méthodologies utilisées dans les deux méta-analyses en réseau effectuées	38
Chapitre 5 - Évaluation des voies d'abord dans les prothèses totales de genoux	41
Chapitre 6 - Évaluation des guides de coupe dans les prothèses totales de genoux	53
Conclusion	69
Partie 3 : Évaluation de l'apprentissage	74
Chapitre 7 - Méthode de modélisation des courbes d'apprentissage	75
Chapitre 8 - Méthode de sommation cumulée	104
a. CUSUM	104
b. LC-CUSUM	106
Chapitre 10 - Comparaison de courbes d'apprentissage à partir du LC-CUSUM	108
Conclusion	130
Partie 4 : Perspective	132
Bibliographie	136
Appendices	153

Liste des abréviations

AAOS : American Academy of Orthopedic Surgeons

AAKS : American Association of Hip and Knee Surgeons

COMET: Core Outcome Measures in Effectiveness Trials

CUSUM (Cumulative Summation)

HKA Hip Knee Angle

ICTRP international Clinical Trials Registry Platform

IMC : Indice de Masse Corporel

KSS : Knee Society Score

LC CUSUM (Learning Curve cumulative Summation)

MCID: minimal Clinically Important Difference

MCMC: Monte carlo Markov Chains

NMA network metana-analysis

PCA Porous Coated Anatomic

PSI PATient specific Instrumentation

PTG prothèse totale de genou

ROB : Risk Of Bias

SF36 : Short Form 12 health survey

SUCRA: SUrface Under the cululative Ranking

TC total Condylar

TKA total knee arthroplasty

UHWPE (Ultra-high-molecular-weight polyethylene

WOMAC : Western Ontario et McMaster

95% CrI : Intervalle de crédibilité à 95%

Table des figures:

Figure 1: Frise chronologique de l'histoire des PTG.	20
Figure 2: Exemple réseau ouvert et fermé.	33
Figure 3. Représentation schématique des différentes voies d'abord analysées dans l'article.	42
Figure 4: Représentation schématique des différentes guides de coupes analysées dans l'articleBi	53

Résumé

Ce travail de thèse porte sur l'évaluation de la qualité de l'acte chirurgical dans les prothèses totales de genoux. Elle s'articule en deux parties.

La première s'intéresse à l'évaluation de deux aspects techniques dans les prothèses totales de genoux à l'aide de méta-analyses en réseaux. La seconde à l'évaluation de l'apprentissage de nouvelles procédures en orthopédie, notamment dans la technique de l'arthroplastie totale du genou.

D'un point de vue général, les méta-analyses en réseau permettent d'extraire des données relatives à une population recevant un traitement particulier, et de comparer entre elles, deux interventions qui ne l'avaient pas encore été dans la littérature internationale.

La première méta-analyse en réseau a évalué les voies d'abord dans les prothèses totales de genoux. Au total cinq voies d'abord ont été comparées : la voie parapatellaire, la voie subvastus, la voie midvastus, la voie mini-parapatellaire et la voie quadriceps sparring. Nous avons observé que la voie subvastus améliorait les mobilités dans les six premiers mois postopératoires sans différence sur les scores fonctionnels à terme. La voie parapatellaire interne diminuait la durée opératoire, comparée aux autres voies d'abord chirurgicales. La durée d'hospitalisation des patients opérés par voie mini-parapatellaire interne était inférieure à celle des patients opérés par une autre technique. Au vu des résultats, aucune voie d'abord chirurgicale médiale pour les prothèses totales de genoux ne semble clairement supérieure à une autre. Le choix de la voie d'abord revient au chirurgien en fonction de ses habitudes, de ses préférences et de sa maîtrise technique. La deuxième méta-analyse en réseau a évalué les guides de coupes dans les prothèses totales de genoux. Au total cinq guides de coupe ont été comparés : le guide standard, la navigation assistée par ordinateur, les guides de coupe réalisés sur mesure (Patient Specific Instrumentation [PSI]), les accéléromètres et la robotique. Nous avons observé que la navigation permettait de diminuer le taux de valeurs aberrantes d'alignement du membre inférieur et de positionnement des implants prothétiques dans les plans frontal et sagittal comparée au guide de coupe standard, mais augmentait le temps opératoire d'environ 20 minutes.

Au vu de ces résultats, nous ne pouvons pas recommander l'utilisation d'un ancillaire de coupe

particulier, puisque les différences observées sur les critères radiologiques n'ont pas eu de répercussions sur les scores cliniques.

La deuxième partie de ce travail s'est intéressée à l'évaluation de l'apprentissage de nouvelles procédures en orthopédie et notamment dans les prothèses totales de genoux. En effet de nouvelles technologies en chirurgie prothétique du genou se développent et l'évaluation et la validation de leur apprentissage est nécessaire.

Une revue de la littérature sur les différentes méthodologies de courbes d'apprentissage utilisées en orthopédie a été menée. Six méthodologies différentes ont été identifiées dans les différentes publications afin de traiter les courbes d'apprentissage en orthopédie : comparaison des premiers patients opérés par un chirurgien avec les suivants, comparaison des actes pratiqués par des chirurgiens expérimentés à ceux réalisés par des débutants, le CUSUM test, le LC-CUSUM test, la méthode utilisant les analyses de survie et la méthode basée sur les régressions linéaires.

Le LC-CUSUM semble être la méthode la plus simple, la plus flexible et la plus robuste pour modéliser une courbe d'apprentissage. De plus, elle peut être utilisée dans différentes spécialités médicales. La seconde observation de cette revue de la littérature était qu'il n'existait aucun test statistique permettant de comparer deux méthodes d'apprentissage.

Enfin, nous présentons dans la dernière partie de cette thèse le développement d'un test statistique afin de pouvoir comparer deux méthodes d'apprentissage entre elles. Après avoir validé les performances du test statistique, une application sur des données réelles d'apprentissage de connaissances a été proposée. Afin que ce test soit validé de manière formelle, une application clinique sur un volume de données plus important est nécessaire.

Concernant l'évaluation clinique des prothèses totales de genoux, la méta-analyse en réseau est une méthodologie adaptée. En effet, il est difficile de mettre en place un essai randomisé comportant quatre ou cinq bras d'une intervention chirurgicale. Elles permettent d'extraire des données et de réaliser des comparaisons jamais effectuées auparavant dans la littérature en réutilisant ces mêmes données et en les synthétisant afin de répondre à une question précise.

Concernant l'évaluation de l'apprentissage en orthopédie, il n'y a aucune recommandation méthodologique. Le LC-CUSUM test semble être la meilleure méthode pour modéliser une courbe d'apprentissage. C'est la technique la plus simple, la plus flexible et la plus robuste. Aucun test statistique n'existe permettant de comparer deux courbes d'apprentissage entre elles. Nous en

avons élaboré un mais qui nécessite une validation lors d'une application clinique plus importante.

Mots-clés : Prothèse totale de genou, voie d'abord, guide de coupe, méta-analyse en réseau, courbes d'apprentissages, test statistique, LC-CUSUM.

Abstract

This thesis work focused on the evaluation of the quality of the surgical operation in total knee arthroplasties. It is organized on the basis of two parts.

The first part was centered on the evaluation of two fundamental technical aspects in total knee arthroplasty : knee approach and bone cutting guides. The second part focused on the learning process of new procedures in orthopedics, especially in knee arthroplasty.

Network meta-analyses (NMA) allow investigators to extract from multiples therapeutic studies data from each treatment group and to compare them with each other, even when a study comparing two specific treatments had not been designed or carried out yet.

Our first NMA evaluated the different surgical approaches to the knee. A total of five different approaches were assessed: the medial parapatellar approach, the subvastus, midvastus, mini parapatellar and quadriceps sparing approaches. We observed that up to six postoperative months, the subvastus approach improved the patients' knee range of motion. At long term, no significant differences were found. The standard parapatellar approach was associated with a reduced operative time, compared to every other approach. Finally, patients operated by a mini parapatellar approach exhibited a shorter duration of hospital stay. In the light of these results, none of the compared knee approaches exhibited a clear superiority to the others. We advocate that this choice belongs to the surgeon, who should use the technique he masters the most.

The second NMA evaluated the different cutting guides available in knee arthroplasty. A total of five guides were evaluated: standard cutting guides, computer assisted navigation surgery, patient custom fit cutting guides (patient specific instrumentation [PSI]), accelerometer, and robotically assisted surgery. We observed that, compared to standard guides and to PSI, navigation was associated with a lower rate of outliers regarding the knee alignment in the frontal plane regarding the femoral and tibial prosthetic components' positions in the frontal and sagittal planes. However, navigation increases the surgical time by an average of twenty minutes. Finally, none of the five cutting guides was associated with a significant superiority regarding midterm clinical results. Having regards to these elements, we cannot strongly recommend the use of an ancillary equipment rather than another.

The second part of this research work focused on the evaluation of the learning process of new orthopedic procedures, especially total knee replacements. As new technologies arise in the knee replacement's landscape, it is crucial to evaluate the capacity of surgeons to learn how to use and master them. A systematic review of literature has been conducted to evaluate the different learning curve methodologies used in orthopedic surgery. A total of six different methods were identified : comparison of experienced against beginners' results, comparison of first and most recent procedures' results, CUSUM tests, LC-CUSUM tests, general linear regression models and survival tests. Finally, the LC-CUSUM appeared to be the simplest, most flexible and strong method at the same time regarding the modeling of a learning curve. In addition, this test can be used in many medical and surgical specialties.

Furthermore, we observed that there was no statistical test to compare to each other two learning curves. Thus, we focused the last part of this research work on the development of a new statistical method to compare to learning curves. After validating the performance of this test, we proposed a practical application on real learning data with encouraging results. In order to formally validate this test, clinical application on a greater amount of data is required.

With regard to the clinical evaluation of total knee replacements, network meta-analyses appear to be a sophisticated, reliable, and strong device. Indeed, it is virtually impossible to organize even a single randomized control trial comparing 4 to 5 different surgical techniques. NMA provide data that allow comparisons that, for some of them, were not even made by previous trials. Those data can therefore be gathered, synthesized to answer specific questions.

Regarding the assessment of learning processes in orthopedics, there is to this day no methodological recommendation. The LC-CUSUM seems to be the best method we have to model surgeons' learning curves. We finally propose a test to allow the comparison of learning curves. Despite good initial results, we look forward to validate this test by with its application on larger populations.

Key words: Total knee arthroplasty, surgical approach, cutting guide, network meta-analysis, learning curve, statistical test, LC-CUSUM.

Publications

Publications en rapport avec la thèse :

Bouché P-A, Corsia S, Dechartres A, Resche-Rigon M, Nizard R. Are There Differences in Accuracy or Outcomes Scores Among Navigated, Robotic, Patient-specific Instruments or Standard Cutting Guides in TKA? A Network Meta-analysis. *Clin Orthop Relat Res.* 2020;478(9):2105-2116. doi:10.1097/CORR.0000000000001324

Bouché P-A, Corsia S, Nizard R, Resche-Rigon M. Comparative Efficacy of the Different Surgical Approaches in Total Knee Arthroplasty: A Systematic-Review and Network Meta-Analysis. *J Arthroplasty.* 2021;36(3):1187-1194.e1. doi:10.1016/j.arth.2020.09.052

Bouché P-A, Resche-Rigon M, Nizard R. Learning curves in orthopedic and trauma surgery: a systematic review. Submitted to *BMC Medical Education.*

Autres publications:

Bouche P-A, Gaujac N, Corsia S, Leclerc P, Anract P, Auburger G. Ankle CT scan allows better management of posterior malleolus fractures than X-rays. *Eur J Orthop Surg Traumatol.* Published online September 1, 2021. doi:10.1007/s00590-021-03104-y

COVIDSurg Collaborative. Delaying surgery for patients with a previous SARS-CoV-2 infection. *Br J Surg.* 2020;107(12):e601-e602. doi:10.1002/bjs.12050

Glasbey JC, Nepogodiev D, Simoes JFF, et al. Elective Cancer Surgery in COVID-19- Free Surgical Pathways During the SARS-CoV-2 Pandemic: An International, Multicenter, Comparative Cohort Study. *J Clin Oncol.* 2021;39(1):66-78. doi:10.1200/JCO.20.01933

Bouché P-A, Guillon P. Infiltrations par corticoïdes des lésions méniscales : efficacité, tolérance et survie à 6 mois minimum. *Journal de Traumatologie du Sport.* 2019;36(3):165-171. doi:10.1016/j.jts.2019.03.003

Vallet H, Seve P, Biard L, et al. Infliximab Versus Adalimumab in the Treatment of Refractory Inflammatory Uveitis: A Multicenter Study From the French Uveitis Network. *Arthritis Rheumatol.* 2016;68(6):1522-1530. doi:10.1002/art.39667

Bouche P-A, Corsia S, Boukebous B, Boutroux P, Zahi R, Guillon P. Is the position of dual-mobility cup in THA for femoral neck fractures optimal? A retrospective study. *Eur J Orthop Surg Traumatol.* 2020;30(2):275-280. doi:10.1007/s00590-019-02576-3

COVIDSurg Collaborative. Outcomes from elective colorectal cancer surgery during the SARS-CoV-2 pandemic. *Colorectal Dis.* Published online November 15, 2020. doi:10.1111/codi.15431

Waldstein W, Bouché PA, Pottmann C, et al. Quantitative and individualized assessment of the learning curve in preoperative planning of the acetabular cup size in primary total hip arthroplasty. *Arch Orthop Trauma Surg.* Published online March 11, 2021. doi:10.1007/s00402-021-03848-6

COVIDSurg Collaborative, GlobalSurg Collaborative. SARS-CoV-2 infection and venous thromboembolism after surgery: an international prospective cohort study. *Anaesthesia.* Published online August 24, 2021. doi:10.1111/anae.15563

COVIDSurg Collaborative, GlobalSurg Collaborative. SARS-CoV-2 vaccination modeling for safe surgery to save lives: data from an international prospective cohort study. *Br J Surg.* Published online March 24, 2021. doi:10.1093/bjs/zxab101

Sigonney F, Lopes R, Bouché P-A, et al. The ankle ligament reconstruction-return to sport after injury (ALR-RSI) is a valid and reproducible scale to quantify psychological readiness before returning to sport after ankle ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc.* 2020;28(12):4003-4010. doi:10.1007/s00167-020-06020-6

COVIDSurg Collaborative. Timing of surgery following SARS-CoV-2 infection: an international prospective cohort study. *Anaesthesia.* n/a(n/a). doi:https://doi.org/10.1111/anae.15458

Bouche P-A, Valteau B, Dumaine V, et al. Were protective procedures against SARS-CoV-2 effective in an orthopaedic and trauma centre during the lockdown period? A retrospective study. *Int Orthop.* 2020;44(12):2493-2498. doi:10.1007/s00264-020-04729-0

Communications:

Belas M, Gaujac N, Pierre-Alban B, Charousset M. Is the learning curve for arthroscopic Latarjet procedure facilitated by a specific step-by-step guiding system for the surgical technique? *Orthopaedic Journal of Sports Medicine*. 2021;9:2325967121S0000. doi:10.1177/2325967121S00003

Sigonney F, Steltzlen C, Pierre-Alban B, Pujol N. A quality analysis of the YouTube video content on meniscus repair surgery. *Orthopaedic Journal of Sports Medicine*.

Préambule

L'arthrose est une pathologie articulaire. Elle affecte et détruit le cartilage. Il s'agit de la maladie articulaire la plus courante (Felson, 1988; Heijink et al., 2012). L'arthrose touche 65% de la population après 65 ans et 80% au-delà de 80 ans (Inserm arthrose, 2012). En 2050, près de 130 millions de personnes souffriront d'arthrose dans le monde (Wittenauer et al., 2013). Dans les pays industrialisés, la prévalence de l'arthrose, quelle qu'en soit la localisation, est estimée entre 8 et 15% (Rabenda et al., 2007; WHO Scientific Group on the Burden of Musculoskeletal Conditions at the Start of the New Millennium, 2003). Parmi toutes les localisations de l'arthrose, la gonarthrose touche 30% de la population âgée de 65 à 75 ans. Elle serait symptomatique jusqu'à 16,7% des personnes de plus de 45 ans (Lawrence et al., 2008). Son incidence annuelle est estimée à 240/100 000 personnes-années (Oliveria et al., 1995). En France, l'arthrose du genou touche 4,7% des hommes et 6,6% des femmes (Guillemin et al., 2011).

Son expression clinique associe une douleur du genou d'intensité variable (essentiellement à la marche, puis parfois au repos), une altération fonctionnelle (limitation de la distance de marche et de l'amplitude des mouvements) et une dégradation de la qualité de vie. Le but du traitement est d'atténuer les symptômes et de ralentir sa progression. Différentes thérapies ont tenté de permettre ces objectifs : mesures générales, physiothérapie, pharmacothérapie et chirurgie. Il existe des traitements chirurgicaux conservateurs (ostéotomie tibiale distale et fémorale proximale), des traitements chirurgicaux palliatifs dans les cas extrêmes (amputation et arthrodèse du genou) et enfin des traitements prothétiques.

La chirurgie prothétique consiste à retirer le cartilage altéré et le tissu osseux sous-chondral des compartiments altérés et à les remplacer par des implants prothétiques. Le genou est composé de trois compartiments : le compartiment fémoro-tibial interne, le compartiment fémoro-tibial externe et le compartiment fémoro-patellaire. Il existe deux grandes catégories de prothèses de genou : les prothèses unicompartmentales (PUC) et les prothèses totales du genou (PTG).

L'arthroplastie totale du genou est une chirurgie courante et efficace dans la prise en charge des gonarthroses symptomatiques évoluées. En 2008, l'incidence des arthroplasties totales du genou était proche de 175 interventions/100 000 dans la population mondiale (Kurtz et al., 2011). En

2013, en France, l'ANSM estimait qu'environ 80 000 PTG étaient pratiquées chaque année (ASNM, 2015). D'ici 2030, l'arthroplastie totale primaire du genou devrait représenter 3,48 millions d'interventions par an dans le monde (Kurtz et al., 2007). Cette intervention implique le remplacement des deux compartiments tibio-fémoraux et parfois du compartiment fémoro-patellaire.

Les prothèses totales de genoux étaient historiquement proposées en dernier recours pour les arthroses extrêmement évoluées. Ces évolutions provoquaient une grande invalidité fonctionnelle et de très grandes déformations des membres inférieurs. L'objectif principal des prothèses était de redonner de la mobilité aux patients et de durer dans le temps. Du fait, de l'augmentation de la population mondiale, de l'obésité et des innovations en médecine, les prothèses totales de genoux sont posées chez des patients de plus en plus jeunes et ayant des objectifs fonctionnels supérieurs. Cependant le taux de satisfaction actuel des patients à la suite d'une prothèse totale de genou est de l'ordre d'environ 80%. C'est une procédure chirurgicale qui est encore perfectible. Il y a actuellement très peu d'innovations sur le design des prothèses. Les points d'amélioration possible portent sur l'apport des nouvelles technologies dans la chirurgie prothétique et sur la gestion péri-opératoire des patients (analgésies, réhabilitation précoce).

Le développement de nouvelles technologies et de nouvelles techniques dans le domaine médical et notamment prothétique est toujours attrayant autant pour le praticien que pour le patient. Malgré la pression parfois commerciale et de la mode, il est toujours important d'évaluer la balance bénéfice/risque des innovations et leur apport réel pour le chirurgien et surtout pour le patient.

Pour cela, il est important de savoir évaluer ces avancées aussi bien en termes de résultats fonctionnels, radiologiques et de survie des implants qu'en terme d'apprentissage.

Nous nous sommes intéressés, dans ce travail, à l'évaluation de la qualité de l'acte chirurgicale dans les prothèses totales de genoux aussi bien en termes de résultats cliniques qu'en terme d'apprentissage.

Dans un premier temps, nous ferons un bref rappel des principes généraux des prothèses totales de genoux. S'en suivra, une présentation des problématiques de méthodologies sur l'évaluation des prothèses totales de genoux.

Partie I

Introduction

Chapitre I

Prothèse totale de genoux

A. Histoire des prothèses totale de genou :

Les premières interventions chirurgicales sur l'articulation du genou remontent au 18^{ème} siècle. Il s'agissait de résections arthroplastiques. Les premières techniques rapportées sont celles effectuées par le Dr Filkin en 1702 (Cooper Samuel, 1838) et par le Dr Henri Park en 1781 (Ollier et al., 1885). À la suite de ces différentes publications, l'intervention connut une grande diffusion. Il faudra attendre 50 ans, en 1827, pour voir le développement d'une nouvelle technique chirurgicale sur le genou : les ostéotomies. Les premiers travaux décrits étaient ceux de J. Rhea Barton qui réalisa deux ostéotomies de réaxation dont une sur le genou en 1837. Cette réaxation permettait de redonner un peu de mobilité dans la néo-articulation créée par l'ostéomie. Rapidement l'idée de mettre un tissu d'interposition pour faciliter le mouvement tout en limitant l'ankylose osseuse, a émergé. La notion, d'interposition tissulaire a été retranscrite pour la première fois par Léopold Ollier dans son livre intitulé « Traité des résections et des opérations conservatrices qu'on peut pratiquer sur le système osseux » publié en 1891 (Ollier et al., 1885). Les premiers tissus d'interposition utilisés étaient du périoste (Louis Ollier 1886), de la peau (Gluck), de la graisse (Lexer 1918) ou parfois du fascia lata (Lexer, 1914).

Les premières PTG ont été implantées dans les années 1890 par le docteur Theophilus Gluck (Robinson, 2005). Il s'agissait d'une prothèse à charnière en ivoire avec une fixation à l'aide de plâtre de Paris et de Colophane. Bien que révolutionnaire, ces implants ont rapidement été abandonnés en raison d'un taux élevé d'infection et d'une fixation inadéquate.

Il faudra attendre les années 1950 pour voir apparaître les évolutions majeures dans le design des PTG. Le Dr Walldius proposa, en 1958, une prothèse charnière en Chrome Cobalt (Jackson and Elson, 1973). Cette prothèse articulée a été largement utilisée jusqu'au début des années 1970.

En 1971, la prothèse de Bosquet-Trillat a fait son apparition. Cette prothèse était novatrice dans le fait qu'elle permettait des mouvements rotatoires (Trillat et al., 1973). La popularité de cette prothèse n'a pas duré longtemps car l'implant nécessitait de grande résection osseuse et avait une tendance à se desceller précocement.

Les prothèses à charnière ont ensuite été réservées aux révisions et à la chirurgie tumorale à cause des taux élevés de reprises. De nouveaux concepts de prothèses ont été explorés. Les années 1970 ont été une période d'innovation dans le domaine de la chirurgie prothétique du genou. C'est dans ces années que le concept de l'arthroplastie condylienne du genou a été développé. Cette approche consistait à remplacer les surfaces porteuses du fémur et du tibia par des composants non liés, refaisant complètement la surface de l'articulation tibio-fémorale. Ces implants dépourvus de contraintes permettraient une meilleure préservation du stock osseux (Robinson, 2005). La première prothèse de ce type fut la prothèse polycentrique développée par Frank Gunston. Cette prothèse était cimentée et composée de deux prothèses unicompartmentales intéressant le fémur et le tibia. Un insert en polyéthylène était présent entre les implants fémoraux et tibiaux. C'est en 1968 que la première prothèse polycentrique fut implantée (Gunston, 2006). Les premiers résultats à long terme furent peu satisfaisants et cette prothèse fut abandonnée.

En 1971, les docteurs Chitranjan Ranawat, Allan Inglis, John Insall et Peter Walker ont développé les prothèses Duocondylar et Unicondylar. La Duocondylar était composée d'implants tibiaux séparés ce qui permettait de conserver les deux ligaments croisés. En 1974, elle a été modifiée afin d'inclure l'articulation fémoropatellaire et afin de préserver le ligament croisé postérieur. Cette prothèse était connue sous le nom de Duopatella.

La prothèse Total Condylar (TC) a été implantée pour la première fois la même année que la prothèse Duopatella. Elle est connue pour être le premier implant à être vendu à l'échelle mondiale et couramment utilisé pour les PTG. Créé par Insall, Ranawat, Walker et Scott, la Total Condylar était une prothèse ne conservant pas le ligament croisé postérieur. L'implant fémoral était en chrome cobalt et les implants rotuliens et tibiaux étaient en full polyéthylène. Elle a été conçue afin de reproduire les mouvements naturels du genou dans le plan sagittal. Elles permettaient une flexion limitée à environ 90° et 100°. Pour résoudre ce problème, Insall et Burstein ont développé une prothèse postéro-stabilisée en 1978, qui était une adaptation de la prothèse Total Condylar (Insall et al., 1982). L'objectif de la postéro-stabilisation était de résoudre les problèmes d'instabilité antérieure et de flexion insuffisante. Dans une période similaire entre 1970 et 1973, une autre prothèse condylienne a été développée par le Dr Mark Coventry :

la prothèse Geomedic (Ranawat and Ranawat, 2012). Elle avait été créée afin de diminuer l'usure du polyéthylène. Les deux ligaments croisés étaient conservés. Une des dernières innovations sur le design des implants a été le développement de prothèse fixée sans ciment avec le développement de l'implant Porous Coated Anatomic (PCA) dans les années 1980 par Hungerford et Kenna (Hungerford et al., 1984). Enfin, dans l'optique de diminuer l'usure du polyéthylène, des plateaux tibiaux mobiles ont été développés. Une des premières prothèses qui intégra ce système était la prothèse LCS. L'ensemble des principales avancées sur les prothèses totales de genou sont résumées dans la figure 1.

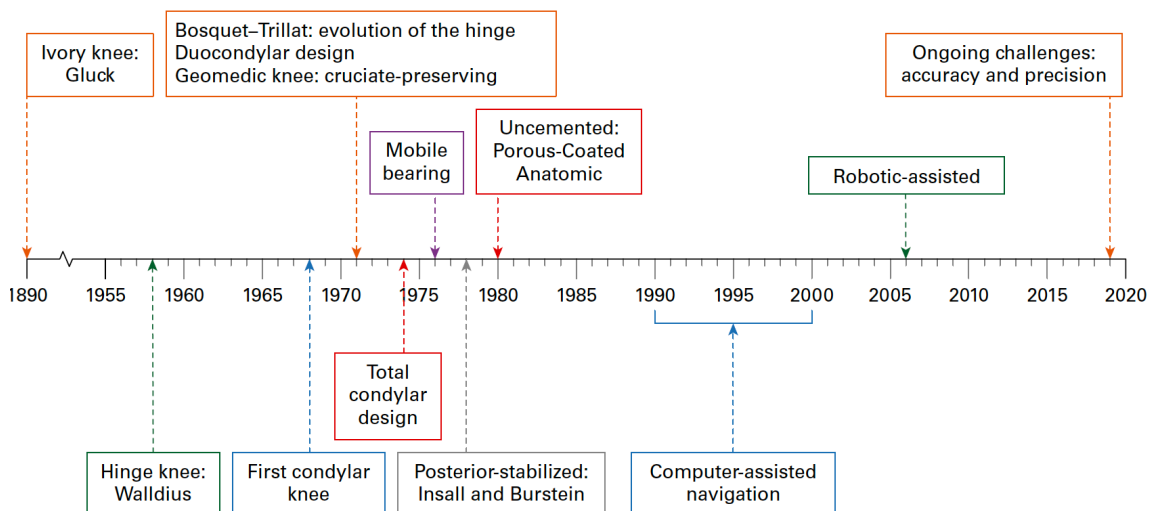


Figure 1. Frise chronologique de l'histoire des PTG tirée de l'article : « Patel NG, Waterson HB, Phillips JRA, Toms AD. 50 years of total knee arthroplasty. Bone & Joint 360. 2019;8(3):3-7. doi:10.1302/2048-0105.82.360688 »

B. Choix dans les prothèses totales de genou :

L'objectif d'une prothèse totale du genou est d'obtenir une articulation prothétique indolore, stable et pérenne dans le temps. Pour répondre à ce cahier des charges, le chirurgien a plusieurs types d'implant à sa disposition. Un des premiers aspects où une décision doit être prise en préopératoire est le type de contrainte intraprothétique de l'implant. La conservation ou non des ligaments croisés et/ou périphériques jouent un rôle prépondérant dans le choix de la contrainte. Il existe trois types de contraintes : les prothèses non-contraintes, les semi-contraintes et les contraintes. Les prothèses non-contraintes sont les prothèses de première intention lorsqu'il n'existe pas d'insuffisance ligamentaire.

On distingue plusieurs types de prothèses non-contraintes : les prothèses à conservation des deux ligaments croisés, à conservation du ligament croisé postérieur et celles ne conservant aucun des deux ligaments croisés. Les prothèses sacrifiant le pivot central sont stabilisées par la forme de l'implant (postéro stabilisation par came tibiale, plateaux tibiaux ultracongruents (deep dish) ou troisième condyle). Aucune différence n'a été mise en avant, pour le moment, entre ces différents implants dans la littérature internationale (Sun et al., 2021). Les prothèses semi-contraintes sont des prothèses permettant de suppléer une insuffisance ligamentaire à l'aide d'une grande came entre les deux condyles fémoraux. Enfin, il existe des prothèses contraintes ou à charnières. Une quille centrale relie le composant fémoral ou composant tibial. Il existe deux types de charnières : les charnières à 1 degré de liberté ou les charnières à deux degrés de mobilité avec un plateau rotatoire.

Un autre aspect qui change en fonction des prothèses est le design de l'implant tibial. On retrouve trois grandes catégories d'implant tibiaux : les implants tout polyéthylène, les implants avec un polyéthylène fixe sur un *metal back* et les implants avec un plateau mobile sur un *metal back*. Le tout polyéthylène est une pièce monobloc de UHWPE (polyéthylène de haut poids moléculaire) généralement à surface articulaire concave qui sera cimentée directement sur la coupe tibiale. Son épaisseur minimale, afin de limiter les risques d'usure précoce et de fluage, est de 8 mm dans la zone de moindre épaisseur (fond de la cupule). Les implants tibiaux avec un *metal back* correspondent à la mise en place d'un plateau tibial métallique, généralement en chrome-cobalt (parfois en titane). L'ancrage au niveau de l'os peut se faire à l'aide de ciment ou non. De même, il existe des *metal back* avec des quilles, des plots ou des vis afin de permettre de renforcer la fixation tibiale dans l'os. Sur le *metal back* est ensuite inséré un polyéthylène fixe ou mobile. Pour les plateaux fixes, le polyéthylène est soit serti en usine ou au moment de la pose.

Un autre point qui diffère entre chirurgiens est le type de fixation choisie. En effet, il est possible de cimenter ou non les implants tibiaux et fémoraux. Certains optent même pour une fixation hybride, c'est-à-dire, qu'un des deux implants est cimenté et l'autre non. Pour l'instant la littérature actuelle ne permet pas de statuer entre les trois types de fixation (Newman et al., 2020; Onggo et al., 2020; Wang et al., 2020).

Enfin, un des derniers choix à effectuer par le chirurgien sur le type d'implant à poser est de savoir s'il resurfaçait ou non la patella. Aucun consensus n'est établi non plus dans la littérature (Chen et al., 2021; Teel et al., 2019). Les arguments pour le non-resurfaçage patellaire seraient une baisse des complications mécaniques (descellement mécanique, rupture de l'appareil extenseur). Les arguments pour le resurfaçage patellaire seraient une meilleure congruence fémoropatellaire et une plus faible incidence des douleurs antérieures postopératoires.

Il existe d'autres choix que le chirurgien doit faire lors d'une pose de prothèse totale de genou qui sont le type de guide de coupe utilisé, la voie d'abord effectuée et la stratégie d'alignement. Les deux premiers font l'objet de ce travail.

L'alignement des composants tibial et fémoral d'une PTG joue un rôle clé dans le succès de la chirurgie de remplacement du genou (Berger et al., 1998; Ritter et al., 1994; Sikorski, 2008). Un défaut d'alignement peut favoriser un descellement prématuré (Jeffery et al., 1991) ou altérer la fonction (Ritter et al., 1994). L'alignement est caractérisé à la fois par l'alignement tibial et fémoral dans le plan coronal, sagittal et transversal, et par l'alignement global du membre inférieur, apprécié par l'angle fémoro-tibial HKA (Hip-Knee-Ankle) (Cherian et al., 2014; Longstaff et al., 2009; Pickering and Armstrong, 2012; Ritter et al., 1994; Sikorski, 2008).

Historiquement, l'implantation des PTG visait la restitution d'un angle HKA à $180^\circ \pm 3^\circ$, alignement dit « neutre » (Insall et al., 1985). En 2010, les travaux de Parratte (Parratte et al., 2010) ont remis en question cet objectif d'alignement en montrant d'une part qu'un alignement neutre en post-opératoire ne diminuait pas le risque de descellement précoce, et d'autre part que les PTG involontairement posées avec un angle HKA hors des objectifs habituels d'alignement présentaient une durée de survie comparable. Depuis, la question de l'alignement des PTG est devenue essentielle et de nouveaux objectifs d'alignements ont été définis (Howell et al., 2013; Hungerford and Krackow, 1985; Oussedik et al., 2020) afin d'améliorer les résultats fonctionnels des PTG qui restent décevants pour près de 20% des opérés.

Actuellement, on définit 3 grands types d'alignement : l'alignement mécanique (Insall et al., 1985), anatomique (Hungerford and Krackow, 1985) et cinématique (Howell et al., 2014). Il existe des variantes de ces différents types d'alignement avec l'alignement constitutionnel, cinématique inversée, cinématique restreint (Hutt et al., 2016).

L'état de l'art actuel montre qu'il n'existe aucun consensus international sur l'objectif d'alignement.

L'alignement mécanique (Insall et al., 1985) correspond à un angle HKA post-opératoire à $180^{\circ} (+/-3^{\circ})$ avec un angle mécanique fémoral distal à 90° et angle tibial proximal à 90° . La rotation externe fémorale est fixée à 3° par rapport au plan bicondylien postérieur. Les épaisseurs de résections sont variables.

L'alignement anatomique a été décrit par Hungerford and Krackow (Hungerford and Krackow, 1985) et prend en compte l'existence d'un valgus épiphysaire physiologique du fémur et un varus épiphysaire physiologique du tibia. Il correspond donc à angle HKA à $180^{\circ} +/- 1,5^{\circ}$ avec un angle mécanique fémoral distal à 93° et angle tibial proximal à 87° . La rotation externe fémorale est parallèle au plan bicondylien postérieur, donc à 0° . Les épaisseurs de résections sont variables.

L'alignement cinématique d'une PTG décrit par Howell (Howell et al., 2014) consiste à restituer la cinématique native du genou en réalisant un resurfaçage du genou. Les coupes fémorales distales et postérieures doivent être strictement compensées par l'épaisseur de l'implant, déduction faite de l'usure liée à l'arthrose sur le condyle fémoral distal usé. La résection osseuse distale sera donc planifiée à $9\text{mm} +/- 2\text{ mm}$ en fonction de l'usure et la résection fémorale postérieure à 9 mm . Dans le cas d'un varus, la coupe fémorale distale médiale est comprise entre 7 et 9 mm du condyle fémoral latéral, idem pour le plateau tibial. Dans le cas d'un valgus, la coupe fémorale distale latérale est comprise entre 7 et 9 mm du condyle fémoral médial, idem pour le plateau tibial. Le choix entre 7 et 9 mm sur la zone d'usure ira dans le sens du meilleur équilibrage possible. Les coupes postérieures sont toujours fixées à 9 mm donc 0° de rotation externe par rapport au plan bicondylien postérieur.

Aucune différence n'a pour l'instant été mise en évidence en termes d'alignement dans la littérature internationale (Luo et al., 2020).

Chapitre 2

Problématique

A. Évaluation clinique des prothèses totales de genoux :

L'orthopédie est une spécialité chirurgicale. Elle est composée, en France, d'une chirurgie programmée et d'une chirurgie d'urgence avec la traumatologie. C'est une spécialité fonctionnelle qui existe dans le monde entier. La recherche clinique en orthopédie s'intéresse à évaluer différentes techniques chirurgicales ou différents matériels utilisés. La méthodologie des études en orthopédie est variée. En cherchant sur Pubmed, entre 2015- 2020, en utilisant les mots clés orthopedic surgery, 126,764 articles ont été indexés. Sur ces 126,764 études, 3,532 articles étaient des essais contrôlés randomisés, soit 2,8%. Ces données montrent qu'il est difficile de mettre en place des études de hauts niveaux de preuve en orthopédie. Comparativement à d'autres spécialités, sur la même période, le nombre d'essais randomisés publiés en orthopédie est environ 2 fois moins important. Par exemple, en cardiologie, 7,029 essais randomisés ont été indexés sur Pubmed et 8,317 sur les soins intensifs entre 2015 et 2020. Cette différence peut s'expliquer par le fait que les essais randomisés pour une intervention chirurgicale est plus difficile à mettre en place et nécessite plus de moyen, notamment avec des *sham* procédures.

Afin de palier à ce déficit d'études de haut niveau de preuve, une façon d'avoir de l'information est d'augmenter l'effectif des études et par conséquent la puissance. C'est pour cela que ce sont développés les registres en orthopédie (Delaunay, 2014). Les registres sont définis en France, selon l'arrêté du 6 novembre 1995 (Gouvernement français, 1995) par le Comité National des Registres, comme « un recueil continu et exhaustif de données nominatives intéressant un ou plusieurs événements de santé dans une population géographiquement définie, à des fins de recherche et de santé publique, par une équipe ayant les compétences appropriées ».

Le premier registre a été créé en 1969 à la Mayo clinic. C'était un registre uniquement utilisé pour le suivi des prothèses. date de 1975. Il existe, dans le monde entier, environ une dizaine

Le premier registre national est le registre suédois dont la création date de 1975. Il existe, dans le monde entier, environ une dizaine de registre (Delaunay, 2014): les registres Norvégien, Danois, Suédois, Finnois, Néo-zélandais, Australien, Canadien, Roumain, Hollandais et celui de Grande-Bretagne. Les registres permettent d'avoir de l'information sur des milliers de prothèses. Le nombre d'articles se basant sur des données de registres au cours des cinq dernières années est d'environ 2100 et ne cesse d'augmenter chaque année.

Cependant, les études sur registre ont des limites (Hailer, 2015).

La première est qu'il existe une grande hétérogénéité entre les pratiques des différents pays et même entre les praticiens d'un même pays. En effet, si la question posée est de voir si le taux de survie des prothèses unicompartimentales du genou est différent en fonction du type de fixation utilisée (ciment ou non) par exemple. Les registres pourront répondre à cette question mais sans forcément prendre en compte l'expérience du chirurgien pour cette intervention, alors qu'il a été prouvé que c'est un facteur de risque sur le succès ou non de cette intervention (Liddle et al., 2016).

Une autre limite possible est qu'il existe un biais de déclaration. En effet, tous les chirurgiens d'un pays, en fonction des législations, ne sont pas obligés de déclarer/ remplir ces registres (Hailer, 2015).

Une autre manière de pouvoir répondre à une question précise en orthopédie est de faire des méta-analyses. Elles permettent notamment de réutiliser les données des différents essais randomisés. Une méta-analyse en réseau est une méta-analyse qui permet d'avoir de l'information sur des comparaisons jamais effectuées dans la littérature. C'est la méthodologie que nous avons choisie pour répondre à nos différentes questions cliniques (Leopold, 2015).

En effet, il est difficile de mettre en place un essai randomisé comparant 4 ou 5 techniques différentes pour une intervention chirurgicale donnée. Cette approche méthodologique permet de comparer plusieurs traitements qui n'ont pas été comparés directement l'un à l'autre dans des précédents essais. Autrement dit, si l'approche A a été comparée à l'approche B dans un certain nombre d'essais randomisés, et l'approche B a été comparée à l'approche C dans d'autres, une méta-analyse en réseau permet de comparer l'approche A à l'approche C et d'avoir des résultats sur cette comparaison indirectement. C'est une manière flexible de répondre à des questions cliniques importantes en réutilisant des données déjà publiées dans la littérature. Pour cela, les

méta-analyses en générale doivent suivre une méthodologie rigoureuse. Les principes généraux des méta-analyses en réseau et de leurs analyses sont développés dans la première partie de ce travail.

B. Évaluation de l'apprentissage :

De nouvelles techniques et technologies (réalité virtuelle, robotique, planification 3D, prothèses sur mesure ...) se développent de plus en plus en médecine et notamment en orthopédie. Les principaux articles traitant de ces nouveautés rapportent des résultats fonctionnels, radiologiques et de survie des implants et certains commencent à évaluer leur apprentissage c'est-à-dire la manière dont un chirurgien acquière la compétence d'une technique et sa validation. Il n'existe pas de recommandation sur la méthodologie à utiliser afin de bien évaluer l'acquisition d'une connaissance/compétence.

L'acquisition de compétences théoriques ou pratiques représente un défi tant pour l'enseignant, que pour l'étudiant. Un moyen de pouvoir objectiver la progression d'un étudiant ou d'une personne en formation est l'utilisation de courbe d'apprentissage. Les courbes d'apprentissage permettent d'observer l'acquisition d'une tâche au fil du temps. Elles ont été développées et largement utilisées pendant la Seconde Guerre mondiale pour la production d'armes (Hall, 2002). Elles ont ensuite été diffusées dans différents domaines, notamment dans l'industrie et plus tard dans la santé (Ramsay et al., 2002; Wright, 1936). L'un des principaux avantages des courbes d'apprentissage est l'association d'une représentation numérique et graphique du processus d'acquisition. La médecine s'est intéressée, en particulier dans les années 1980, aux courbes d'apprentissage afin de déterminer l'efficacité d'une procédure et le temps nécessaire pour acquérir les nouvelles techniques par les étudiants (Ramsay et al., 2002; Valsamis et al., 2018a). Plusieurs techniques ont été développées pour les décrire en médecine (Valsamis et al., 2018b) mais sans aucune recommandation.

Les courbes d'apprentissage ont été principalement développées et utilisées en chirurgie pour savoir si un chirurgien est compétent pour une procédure donnée. En effet, l'acceptabilité par les patients et les autorités d'être traité par un chirurgien incapable de prouver ses capacités à exécuter une technique conventionnelle ou en développement a considérablement diminué.

De plus, avant d'autonomiser un chirurgien junior dans une équipe, une mesure objective de l'acquisition de ces compétences est souhaitable. Une nouvelle technologie qui nécessite une courbe d'apprentissage importante afin de la maîtriser est-elle une avancée pour le patient ?

Afin de pouvoir utiliser au mieux les courbes d'apprentissage en orthopédie, nous avons, dans un premier temps, effectué une revue de la littérature pour voir les différentes méthodologies existantes et définir la plus adaptée à notre spécialité. Dans un second temps, afin de voir si la courbe d'apprentissage d'une nouvelle technique n'était pas un frein à sa diffusion, nous avons développé un test statistique permettant de comparer deux courbes entre elle.

Avec le développement de nouvelles technologies et de nouvelles techniques en orthopédie, il est nécessaire d'évaluer la balance bénéfice/risque des innovations et leur apport réel pour le chirurgien et surtout pour le patient. Pour cela, il faut savoir évaluer ces avancées.

Nous nous sommes intéressés, dans ce travail, à l'évaluation de la qualité de l'acte chirurgicale dans les prothèses totales de genoux aussi bien en termes de résultats cliniques qu'en termes d'apprentissage.

La première s'intéresse à l'évaluation de deux aspects techniques dans les prothèses totales de genoux à l'aide de méta-analyses en réseaux. La seconde à l'évaluation de l'apprentissage de nouvelles procédures en orthopédie, notamment l'arthroplastie totale du genou.

Partie 2

Évaluation à l'aide de méta-analyses en réseau

Chapitre 3

Principe des méta-analyses

Une méta-analyse est une méthode statistique qui permet de synthétiser quantitativement, par le calcul d'un effet combiné, les résultats d'études indépendantes répondant à une question de recherche bien précise. C'est, avant tout, une revue critique de l'ensemble des données de la littérature reposant sur une méthodologie rigoureuse et reproductible, composée de plusieurs étapes. Il faut, dans un premier temps, identifier toutes les études (publiées ou non) répondant à cette question clinique. Ensuite, il faut évaluer la qualité de chaque étude identifiée. Enfin, il faut synthétiser les résultats de ces études en faisant une analyse statistique.

Karl Pearson fut le premier à effectuer une méta-analyse en 1904 (Karl Pearson, 1904) pour tenter de résoudre un problème de puissance statistique pour des échantillons de petite taille. La première méta-analyse dans le domaine médical date de 1955 et a été effectuée par Henry K. Beecher (Beecher, 1955). Depuis, en médecine, le nombre de méta-analyses ne cesse d'augmenter : en 1990 la base de données Pubmed recensait 272 méta-analyse, 848 en 2000, 4691 en 2010 et 17,070 en 2020. L'incidence du nombre de méta-analyse par an a été multipliée par 20 en 20 ans. Cette augmentation est due à une explosion du nombre d'études publiées et au large déploiement de logiciels en facilitant la réalisation. En 1979, 14 essais randomisés étaient publiés par jour et en 2010, 75 essais étaient publiés par jour (Bastian et al., 2010). Il existe, donc, un écart important entre la quantité d'information disponible et les contraintes de temps des médecins pour se mettre à jour sur les données de la littérature.

Les méta-analyses permettent de synthétiser les résultats des différents essais traitant d'une question précise. En soi cette méthodologie est intéressante dans l'optique d'une médecine basée sur les preuves (Sackett et al., 1996). De plus, il s'agit, lorsque la méthodologie est rigoureuse, d'une étude de haut niveau de preuve (Burns et al., 2011). Un autre avantage des méta-analyses est qu'elles permettent de répondre à des questions pour lesquelles la littérature internationale n'arrive pas à conclure, ou donne des résultats contradictoires, en augmentant la puissance sta-

tistique et en améliorant la précision de l'estimation de l'effet traitement.

Une méta-analyse est composée d'une part d'une revue systématique de la littérature et d'une analyse statistique. La qualité d'une méta-analyse dépend de cette revue systématique de la littérature. Il faut s'efforcer d'identifier la littérature « grise », la littérature non publiée. Pour cela, il faut interroger être le plus complet possible et interroger les registres d'enregistrements des essais cliniques que sont par exemple clinicaltrials.gov et l'ICTRP. Il faut aussi scruter les communications des différents congrès et lorsqu'il le faut contacter les auteurs de ces communications. Les méta-analyses sont rétrospectives par nature. Il y a la nécessité d'écrire un protocole avant le début de la recherche. En effet, le choix des critères d'inclusion des études, des critères de jugement, de la mesure de l'effet du traitement et la méthodologie statistique doivent être définis a priori à l'instar d'un essai thérapeutique randomisé. Ce protocole doit ensuite être enregistré auprès de PROSPERO. L'enregistrement du protocole permet d'avoir de la transparence et de limiter les risques biais (décisions post-hoc).

La méthodologie d'une méta-analyse suit un plan bien défini par le comité PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Page et al., 2021). Il faut tout d'abord formuler une question claire et précise définissant la population d'étude, l'intervention évaluée, les comparateurs et les critères de jugements. Suite à ça, une stratégie de recherche est élaborée pour identifier toutes les études réalisées (publiées ou non). Puis, il faut, sélectionner les études selon les critères d'éligibilité définis dans le protocole et évaluer la qualité des études. La qualité des études s'évalue à l'aide du Risk of Bias Tool (ROB). Le risque de biais est évalué à l'aide de 7 items par 2 investigateurs : évaluation de la génération de la séquence de randomisation, de la dissimulation de l'allocation, de l'aveugle ou non des patients et des praticiens, de l'aveugle ou non des investigateurs, de la gestion des données manquantes sur les résultats, de l'existence d'un protocole ou non et des autres sources potentielles de biais. Ensuite, une extraction des données est réalisée à l'aide d'une grille de recueil testée au préalable. Enfin, l'analyse statistique et l'interprétation des résultats sont faits.

Il existe différents types de méta-analyses qui dépendent du mode de recueil des données, du mode d'agrégation des données et du type d'études incluses. Les méta-analyses sur données agrégées de la littérature sont les plus fréquentes. Leurs avantages sont que les données sont facilement accessibles et qu'elles sont facilement reproductibles.

Leur principale limite est que la qualité de la méta-analyse repose sur les données disponibles dans les articles. Les méta-analyses sur données individuelles, c'est-à-dire d'après les données des patients et non sur les données résumées par groupe de patients présentés dans l'article. Leurs avantages sont que l'analyse est plus fine et plus homogène (critères d'inclusion des patients, mesure d'association), qu'il n'existe pas de problème de l'information rapportée dans les articles et que plus de critères de jugement sont analysables. Leur principale limite repose sur l'accessibilité des données avec un fort risque d'analyse incomplète. Il existe aussi des méta-analyses cumulatives. L'effet traitement est mesuré par ordre chronologique. Enfin, il y a des méta-analyses sur essais contrôlés randomisés ou sur les études observationnelles. Les avantages des essais randomisés sont qu'ils sont plus souvent publiés et que les populations sont moins hétérogènes. Les limites des méta-analyses sur essais randomisés sont que le suivi des patients est souvent limité, le nombre de patient inclus est souvent faible et sont parfois peu représentatives de la population générale (patients sélectionnés).

Chapitre 4

Particularité des méta-analyses en réseau

A. Représentation du réseau :

Une particularité des méta-analyses en réseau est la représentation graphique du réseau. Un graphique du réseau est fait pour chaque critère de jugement afin de représenter les comparaisons directes entre les différents traitements. Dans ces graphiques, les cercles représentent les différents traitements et la taille de ces cercles est proportionnelle à la taille de l'échantillon de chaque intervention respective ; les lignes représentent les comparaisons directes et leurs épaisseurs sont proportionnelles au nombre d'études disponibles.

L'analyse du réseau est importante et donne des informations. Il décrit l'agenda de recherche, la manière dont la recherche a été effectuée au cours du temps. Il permet d'identifier les différents comparateurs utilisés dans la littérature. Il existe principalement deux grandes catégories de réseaux : les réseaux fermés et les réseaux ouverts (figure 2). Les réseaux fermés correspondent à une intervention où tous les traitements ont été comparés deux à deux. À l'opposé, un réseau ouvert correspond à une intervention où certains traitements n'ont pas été comparés l'un à l'autre.

La plupart des réseaux sont des réseaux ouverts et un comparateur commun est souvent utilisé, généralement soit un placebo ou le traitement de référence. Il manque d'essais face-à-face par absence de vision globale des chercheurs et par préférences sélectives dans le choix des comparaisons par les laboratoires. Le réseau permet de voir les manques dans l'agenda de recherche (research gaps) et par conséquent les perspectives de futures recherches.

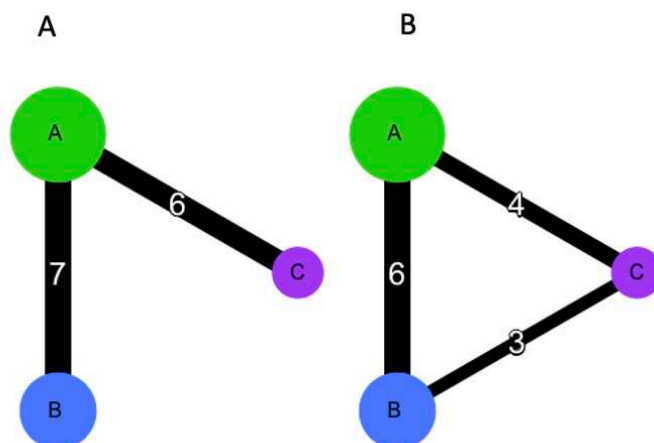


Figure 2. Exemple de réseau ouvert (2.A) et de réseau fermé (2.B). Chaque cercle représente un traitement dont la taille est proportionnelle à la taille de l'échantillon de chaque intervention respective ; les lignes représentent les comparaisons directes et leurs épaisseurs sont proportionnelles au nombre d'études disponibles.

B. Hypothèses à valider :

Afin de valider et d'évaluer la qualité d'une méta-analyse en réseau, différentes hypothèses doivent être vérifiées.

- Hypothèse de cohérence :

La première est l'hypothèse de cohérence. Elle correspond à l'hypothèse qu'il n'existe pas de divergence entre les comparaisons directes et indirectes (Higgins et al., 2003). Elle peut être évaluée à l'aide d'approches locales (loop-based tests, node splitting) et globales (Q test, I²).

- Mesure de l'hétérogénéité :

La deuxième hypothèse est la présence d'homogénéité entre les études. L'homogénéité est l'absence de variabilité excessive dans l'estimation de l'effet traitement entre les essais randomisés comparant deux interventions. Elle peut être évaluée statistiquement au sein de chaque comparaison. On peut utiliser les méthodes statistiques I² (avec son IC à 95%), qui mesure le pourcentage de variabilité non attribuable à une erreur aléatoire, et en utilisant le test de Q Cochrane (avec son IC à 95%), qui permet de distinguer l'erreur aléatoire et l'hétérogénéité.

- Hypothèse de transitivité :

Enfin la troisième hypothèse est l'hypothèse de transitivité. C'est l'absence de différence excessive entre les essais cliniques randomisés A vs. C et ceux B vs. C au niveau du comparateur commun et au niveau de la distribution des modificateurs de l'effet traitement. Par exemple, le traitement C ne doit pas marcher différemment dans A vs C et dans B versus C. L'hypothèse de transitivité part du principe qu'un patient doit pouvoir être randomisé dans n'importe quel groupe du réseau. Cette hypothèse ne peut pas être testée statistiquement. Il faut regarder la description des caractéristiques des essais (tableau 1 généralement) ; l'évaluation de la présence de modificateurs d'essai non reconnus à l'aide d'analyse en sous-groupe ou de méta régressions.

La validité des résultats d'une méta-analyse en réseau doit être évaluée à l'aide de l'outil GRADE Framework (Salanti et al., 2014).

C. Analyse statistique :

Diverses approches méthodologiques ont été proposées pour les méta-analyses en réseau. Il est possible d'utiliser un cadre fréquentiste ou bayésien. Les méthodes statistiques pour les méta-analyses classiques permettent de combiner les résultats de plusieurs études et d'estimer un effet traitement commun aux différentes études. La méta-analyse en réseau permet en plus de combiner de l'information directe et indirecte et ainsi d'effectuer la comparaison deux traitements même en l'absence d'essai en « one to one » et donc de comparaison directe existante.

L'hypothèse principale des méta-analyse en réseau est que l'efficacité relative d'un traitement est la même dans tous les essais inclus dans l'analyse.

Nous présentons dans un premier le principe de l'analyse fréquentiste puis le principe de l'analyse Bayésienne.

- Méthode fréquentiste :

La première étape est le calcul de l'effet traitement et de sa variance pour chaque étude. On obtient une estimation ponctuelle de l'effet traitement. Pour les variables binaires, les me-

sure de l'effet traitement peut être soit le calcul d'un risque relatif, soit d'un odds ratio ou soit d'une différence de risque. Pour les variables continues, les mesures de l'effet traitement sont classiquement soit le calcul d'une différence de moyenne ou d'une différence de moyenne standardisée. Les estimations sont présentées soit avec un intervalle de confiance.

Il faut ensuite combiner les estimés ponctuels afin d'obtenir une estimation combinée des effets traitement directs. L'effet traitement combiné est une moyenne pondérée des effets traitement de chaque étude. Ainsi, l'effet combiné du traitement A par rapport au traitement B s'obtient à l'aide des formules suivantes.

- Moyenne pondérée : $\hat{\theta}_{AB}^{dir} = \frac{\sum_{i=1}^k \theta_{ABi} \times w_i}{\sum_{i=1}^k w_i}$
- Variance de la moyenne pondérée : $Var\left(\hat{\theta}_{AB}^{dir}\right) = \frac{1}{\sum_{i=1}^k w_i}$

Avec $w_i = 1/\sigma_i^2$, l'inverse de la variance de l'effet traitement θ_{ABi} de la comparaison directe dans l'étude i .

L'étape suivante consiste à obtenir l'effet traitement des comparaisons indirectes. Par exemple, pour l'estimation de l'effet traitement θ_{BC} indirect entre B et C du réseau ouvert présenté dans la figure 2.A:

- $\hat{\theta}_{BC}^{ind} = \hat{\theta}_{AB}^{dir} - \hat{\theta}_{CA}^{dir}$
- $Var(\hat{\theta}_{BC}^{ind}) = Var(\hat{\theta}_{AB}^{dir}) + Var(\hat{\theta}_{CA}^{dir})$

On obtient ensuite un intervalle de confiance à 95% en utilisant une approximation normale l'IC à 95% :

$$IC\ 95\% \left[\exp \exp \left(\hat{\theta}_{BC}^{ind} - 1,96 \sqrt{Var(\hat{\theta}_{BC}^{ind})} \right); \exp \left(\hat{\theta}_{BC}^{ind} + 1,96 \sqrt{Var(\hat{\theta}_{BC}^{ind})} \right) \right]$$

Enfin, on obtient les estimations mixtes en combinant les estimations combinées directs et indirectes à l'aide de la moyenne pondérée par l'inverse de la variance. Par exemple, estimation de l'effet traitement AB de l'exemple du réseau ouvert de la figure 2.A :

$$\hat{\theta}_{AB}^{combined} = \frac{\frac{1}{Var(\hat{\theta}_{AB}^{dir})} \hat{\theta}_{AB}^{dir} + \frac{1}{Var(\hat{\theta}_{AB}^{ind})} \hat{\theta}_{AB}^{ind}}{\frac{1}{Var(\hat{\theta}_{AB}^{dir})} + \frac{1}{Var(\hat{\theta}_{AB}^{ind})}}$$

$$Var(\hat{\theta}_{AB}^{combined}) = \frac{1}{\frac{1}{Var(\hat{\theta}_{AB}^{dir})} + \frac{1}{Var(\hat{\theta}_{AB}^{ind})}}$$

Dans un cadre fréquentiste, le résultat de l'analyse des comparaisons de traitements indique si les différences observées sont statistiquement «significatives».

- Méthode Bayésienne:

Dans un cadre « fréquentiste » l'effet traitement, que l'on cherche à estimer, est considéré comme fixe. En cadre bayésienne, l'effet traitement, noté θ , est aléatoire et est défini initialement par une loi de probabilité dit a priori, notée $Pr(\theta)$. Cette loi représente l'ensemble des informations a priori disponibles sur ce paramètre θ avant l'obtention de données cliniques.

Nous obtenons, après analyse, en inférence bayésienne, une loi a posteriori $Pr(\theta|y_{obs})$ qui estime l'effet traitement en prenant en compte les données observées (y_{obs}). Cette loi fournit une information en termes d'incertitude qui est représenté par un intervalle de crédibilité.

Le calcul de la densité de la loi a posteriori $Pr(\theta|y_{obs})$ repose généralement sur des méthodes numériques, basées sur des échantillonneurs de Gibbs (méthodes dites de Monte Carlo par Chaînes de Markov (Monte Carlo Markov Chains en anglais ou MCMC)). La formule mathématique du calcul de la loi a posteriori est la suivante :

$$Pr(\theta|y_{obs}) \propto v(\theta|y_{obs}) \times Pr(\theta)$$

Avec $Pr(\theta|y_{obs})$ la loi a posteriori, $v(\theta|y_{obs})$ la vraisemblance du modèle sachant les observation y_{obs} et $Pr(\theta)$ la loi a priori d'un jeu de données y .

Le principe du MCMC est, qu'après un certain nombre de tirages, la chaîne converge vers la densité a posteriori du paramètre du modèle, $Pr(\theta|y_{obs})$.

Pour produire des estimations avec suffisamment de confiance, une période de rodage, appelée « burn in period », est réalisée. Elle consiste à exclure les premières valeurs tirées du modèle. On vérifie ensuite la convergence du modèle à l'aide de représentation graphique dont le plus fréquemment utilisé est le graphique de Brooks-Gelman-Rubin.

Le résultat de la méta-analyse bayésienne est donc une distribution conjointe a posteriori de tous les effets de traitement relatifs pour toutes les interventions incluses dans le réseau. Les effets du traitement sont généralement rendus avec leurs intervalles de crédibilité à 95% (CrI). La méthode bayésienne permet également d'estimer la probabilité qu'un traitement soit supérieur à un ou plusieurs autres.

- Effets fixes - Effets aléatoires

Il est possible de considérer, comme pour les méta-analyses standard, des modèles à effets fixes ou à effets aléatoires aussi bien dans un cadre fréquentiste que dans un cadre bayésien.

Le modèle à effet fixe suppose qu'il n'existe pas de variabilité de l'effet traitement entre les études. Sa formule mathématique est :

$$y_k \sim N(\delta, s_k^2)$$

Avec y_k l'effet traitement pour l'étude k ($k=1, \dots, K$) et δ l'effet réel du traitement et s_k^2 sa variance.

Le modèle à effet aléatoire suppose qu'il existe une hétérogénéité de l'effet traitement entre les études. Sa formule mathématique est :

$$y_k \sim N(\delta_k, s_k^2)$$

Avec y_k l'effet traitement pour l'étude k ($k=1, \dots, K$) et δ_k l'effet réel du traitement, s_k^2 sa variance.

La distribution globale de la taille de l'effet réel de l'effet traitement est définie par :

$$\delta_k \sim N(d, \tau^2)$$

Avec d l'effet moyen du traitement et τ^2 sa variabilité inter-essai.

D. Classement des traitements :

Une autre particularité des méta-analyses en réseau est qu'elles permettent d'obtenir un classement des différents traitements inclus dans l'analyse. Ce classement est possible grâce au calcul de probabilités. Pour chaque traitement du réseau, une estimation de la probabilité d'être le meilleur est calculé à la fois dans l'approche bayésienne ou dans l'approche fréquentiste.

Le classement entre les différents traitements peut être représenté soit par des courbes des probabilités de rang, soit par la distribution des probabilités de rang cumulé représentée par l'aire sous la courbe (SUCRA : Surface Under the Cumulative Ranking) ou par l'estimation du rang moyen.

Le classement des différents traitements est une information en soi mais ne peut être interpréter isolement et il ne doit pas se substituer aux effets relatifs des traitements.

E. Méthodologies utilisées dans les deux méta-analyses en réseau effectuées :

Un graphique du réseau a été fait pour chaque critère de jugement afin de représenter les comparaisons directes entre les différentes interventions.

La méta-analyse en réseau combine de l'information à partir des comparaisons directes et indirectes pour toutes les comparaisons relatives. Premièrement, pour les comparaisons directes, nous avons effectué une méta-analyse traditionnelle afin de synthétiser les résultats des essais randomisés comparant les différents traitements. Un modèle à effet aléatoire a uniquement été utilisé. Ensuite, nous avons combinés les résultats des comparaisons directes et indirectes via une méta-analyse en réseau utilisant le modèle hiérarchique de Lu et Ades à l'aide d'une approche bayésienne.

Les chaînes de Monte Carlo Markov ont été utilisées pour implémenter le modèle. Quatre chaîne de Markov ont été exécutées simultanément. Pour chaque séquence, les 5 000 premières valeurs ont été exclues (« burn-in » période). Les distributions à posteriori étaient basées sur un échantillon aléatoire de 20 000 valeurs. La convergence du modèle a été évaluée à l'aide de graphiques de Brooks-Gelman-Rubin. Dans le cadre de l'approche bayésienne, les paramètres

du modèle ont été évalués à l'aide d'échantillonneurs de Gibbs. Nous avons évalué l'ajustement via la déviance résiduelle. La cohérence du réseau (incohérence entre les comparaisons directes et indirectes) a été évaluée en utilisant les facteurs d'incohérence définis par Lu et Ades.

Les critères de jugement binaires ont été décrits à l'aide de pourcentages calculés à partir des différences de risque. Les critères de jugement continus ont été analysés à l'aide de la différence des moyennes standardisées. Des intervalles de crédibilité à 95% ont été calculés pour tous les résultats. Les données ont été analysées en intention de traiter. Une distribution a priori non informative binaire a été utilisée pour les résultats dichotomiques et une distribution a priori non informative normale pour le modèle continu.

Enfin, l'approche bayésienne nous a permis de mesurer l'effet traitement sous forme de probabilités et ainsi d'obtenir un classement des guides de coupe. Nous avons calculé la probabilité pour chaque interventions d'être classé 1, 2, 3, 4 ou 5 pour chaque critère de jugement. Pour l'analyse des données, nous avons utilisé le logiciel R version 4.3.3.

Concernant l'hypothèse de cohérence, des « Loop-based tests » et des « node splitting tests » ont été utilisés pour les approches locales). Pour identifier l'incohérence globale du réseau, nous avons utilisés la méthode statistique I^2 .

Nous avons évalué statistiquement la présence d'hétérogénéité au sein de chaque comparaison en utilisant la méthode statistiques I^2 avec son intervalle de crédibilité à 95%. Elle mesure le pourcentage de variabilité non attribuable à une erreur aléatoire. Nous l'avons aussi évaluée en utilisant le « Q test Cochran » avec son intervalle de crédibilité à 95%. Il permet de distinguer l'erreur aléatoire de l'hétérogénéité.

Nous avons effectué des méta-régressions en réseau et des analyses en sous-groupe.

Les méta-régressions ont été faites sur l'âge, l'IMC et le pourcentage de femmes dans les études. Les analyses en sous-groupe ont été faites en fonction des données recueillis selon l'intervention évaluée.

Enfin des funnel plots ont été faits pour vérifier la présence d'un biais de type small study effect. Pour évaluer, le biais de publication nous avons réalisé des contour-enhanced funnel plots pour chaque comparaison. Si un biais de publication était détecté, nous avons tenté d'établir

une relation entre l'effet traitement observé et le fait d'être publié.

La validité des résultats de ces méta-analyses en réseau a été évaluée à l'aide de l'outil GRADE Framework (Salanti et al., 2014)

Chapitre 5

Évaluation des voies d'abords dans les prothèses totales de genoux

Le type de voie d'abord chirurgical peut jouer un rôle clé dans l'alignement de la PTG (alignement des composants fémoraux et tibiaux) et dans la rééducation du patient (Goble and Justin, 2004; Vail, 2004). L'abord chirurgical est le passage créé par le chirurgien pour aller de la peau à l'articulation du genou. Cinq voies médiales chirurgicales sont actuellement utilisées (Scuderi et al., 2004) dans les prothèses totales de genoux (Figure 3) :

- Voie parapatellaire médiale
- Voie subvastus
- Voie midvastus
- Voie mini parapatellaire interne
- Voie quadriceps sparing

La différence majeure entre toutes ces voies d'abord réside dans l'incision à travers le quadriceps. La voie la plus fréquemment utilisée et la plus répandue dans le monde est l'approche parapatellaire médiale (Insall, 1971). Dans cet abord chirurgical le tendon du quadriceps est incisé longitudinalement. La mini voie d'abord parapatellaire médiale est une voie d'abord parapatellaire médiale standard dont l'extension supérieure dans le tendon du quadriceps ne dépasse pas 4 cm (Scuderi et al., 2004). Ont été ensuite développées des approches mini-invasives (Scuderi et al., 2004) qui n'ouvrent pas le tendon quadricipital. Elles permettraient une récupération plus rapide, une expérience moins douloureuse et une flexion accrue. L'abord midvastus consiste en un décollement de la partie distale des fibres obliques du vaste médial (Engh et al., 1997). L'abord subvastus consiste en une arthrotomie horizontale le long de la face inférieure du vaste médial (Hofmann et al., 1991). La voie quadriceps sparing épargne le quadriceps, c'est un abord

sous le vaste dans lequel le vaste médial est mobilisé et séparé du septum intermusculaire, permettant sa rétraction latérale (Alan and Tria, 2010; Tria and Coon 2003).

Ces différents abords chirurgicaux ont les avantages des inconvénients des autres. L'utilisation d'une voie d'abord plutôt qu'une autre pourrait améliorer la récupération postopératoire et potentiellement les mobilités.

De nombreux essais thérapeutiques ou méta-analyses comparant ces différentes voies d'abord chirurgicales de PTG ont été menés avec des résultats contradictoires (Alcelik et al., 2012; Berstock et al., 2018; Gandhi et al., 2011; Kazarian et al., 2018; Liu et al., 2014; Smith et al., 2012). Cependant, aucun des travaux précédents ne permet une comparaison simultanée de toutes les approches décrites précédemment. L'objectif de ce travail était de comparer ces différentes techniques chirurgicales à l'aide d'une méta-analyse en réseau.

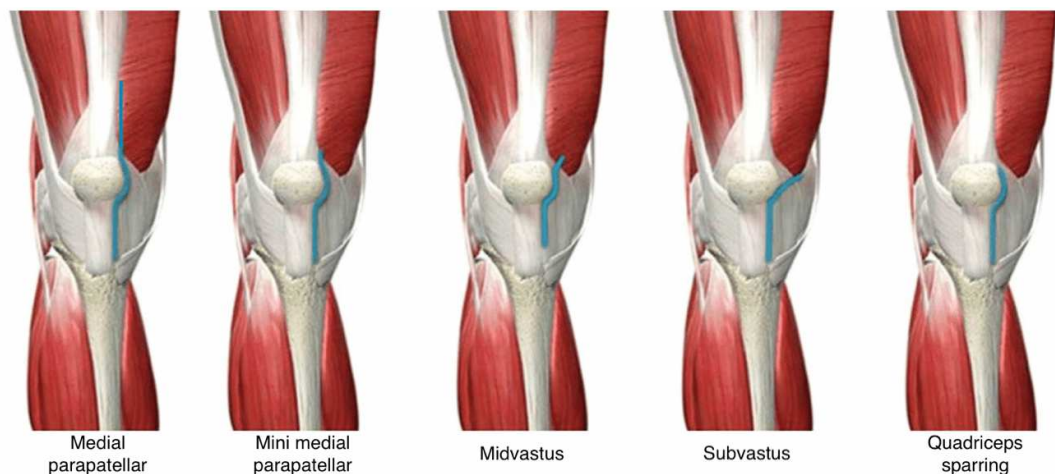


Figure 3. Représentation schématique des différentes voies d'abord analysées dans l'article. De gauche à droite, Voie parapatellaire médiale, voie mini parapatellaire médiale, voie midvastus, voie subvastus et la voie quadriceps sparring.

La référence de l'article présenté est :

Bouché P-A, Corsia S, Nizard R, Resche-Rigon M. Comparative Efficacy of the Different Surgical Approaches in Total Knee Arthroplasty: A Systematic-Review and Network Meta-Analysis. *J Arthroplasty*. 2021;36(3):1187-1194.e1. doi:10.1016/j.arth.2020.09.052

ARTICLE IN PRESS

The Journal of Arthroplasty xxx (2020) 1–8



Contents lists available at ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

Comparative Efficacy of the Different Surgical Approaches in Total Knee Arthroplasty: A Systematic-Review and Network Meta-Analysis

Pierre-Alban Bouché ^{a, b, *}, Simon Corsia ^b, Rémy Nizard, MD, PHD ^{b, 1},
Matthieu Resche-Rigon, MD, PHD ^{a, 1}

^a Service de Biostatistique et Information Médicale, Hôpital Saint-Louis, APHP; Université Paris Diderot; ECSTRA Team, UMR U1153, INSERM, Paris, France

^b Service de chirurgie orthopédique et traumatologique, Hôpital Lariboisière, Paris, France

ARTICLE INFO

Article history:

Received 4 August 2020

Accepted 30 September 2020

Available online xxx

Keywords:

total knee arthroplasty
network meta-analysis
surgical approaches
arthroplasty
review

ABSTRACT

Background: Several surgical approaches including midvastus, subvastus, mini-parapatellar, quadriceps-sparing (QS) and parapatellar are currently used to perform total knee arthroplasty (TKA). Since none of published study exhibited a simultaneous comparison of all of them, a network meta-analysis has been conducted to compare the most widely used knee surgical approaches regarding the improvement of functional outcomes and the range of motion (ROM).

Methods: Randomised controlled trials (RCTs) comparing TKA approaches were searched in electronic databases, major orthopedics journals, and oral communications, ClinicalTrials.gov and WHO International Clinical Trials Registry Platform until May 1st, 2020. Two reviewers independently selected trials and extracted data. The primary outcomes were functional scores at 6 months post-surgery evaluated by KSS and WOMAC, and the ROM.

Results: Sixty RCTs involving 5042 patients with 5107 TKA were included. No significant differences between different approaches were found for the KSS assessment or the WOMAC at 6 months. The mean of ROM at 6 months post-surgery were higher in the subvastus group than in all the others surgical approaches. The difference of ROM with subvastus approach was 7.3° (95% CI –14.1 to –0.1) with the midvastus approach, 11.1° (95% CI –18.7 to –2.8) with mini-parapatellar, 8.9° (95% CI –14.2 to –3.1) with standard parapatellar, and 9.2° (95% CI –16.1 to –1.8) with QS.

Conclusion: No differences were found in functional outcomes over short or medium terms but subvastus seemed to increase the ROM at 6 months post-surgery. Until or unless future studies can demonstrate a long-term benefit, based on these results all studied surgical approaches to perform a TKA are equal.

Level of Evidence: Network meta-analysis. Level 1.

© 2020 Elsevier Inc. All rights reserved.

The choice of the surgical approach in Total Knee Arthroplasty (TKA) could play a key role in knee's prosthetic components alignment and in the patient's recovery especially during the early postoperative period [1,2]. Five surgical approaches are currently

used [3]. The oldest and most frequent one is the medial parapatellar approach [4]. In this approach the quadriceps tendon is incised longitudinally along its medial edge, sufficiently to allow lateral patellar subluxation. The main advantages of parapatellar arthroscopy are reproducibility, and versatility, allowing its application in complex surgical situations [5]. However, it could damage the extensor mechanism and replaces medial musculotendinous area of quadriceps tendon with scar tissue [6,7]. It could also lead to higher incidence of lateral patellar release and diminished blood supply to patella [8].

Mini invasive approaches were developed, in the late 1990's, initially for unicompartmental knee replacement [3]. They aimed to allow faster recovery, less painful experience and increased post-

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to <https://doi.org/10.1016/j.arth.2020.09.052>.

* Reprint requests: Pierre-Alban Bouché, INSERM U1153 Hôpital Saint-Louis Bâtiment Lailler 3ème étage 1, Avenue Claude Vellefaux, 75010 Paris, France.

¹ Co-last author.

<https://doi.org/10.1016/j.arth.2020.09.052>

0883-5403/© 2020 Elsevier Inc. All rights reserved.

ARTICLE IN PRESS

2

P.-A. Bouché et al. / The Journal of Arthroplasty xxx (2020) 1–8

operative knee flexion. Four main mini invasive procedures are described. The mini medial parapatellar approach is a standard medial parapatellar approach with proximal extension into quadriceps tendon not exceeding 4 cm [3]. The Midvastus approach consists in a blunt dissection of the distal part of the oblique fibers of the vastus medialis that ends at the proximal medial patellar corner [9], continued by a classical medial parapatellar arthroscopy. The Subvastus approach uses a horizontal arthroscopy along the inferior aspect of the vastus medialis and an incision through the retinacular cuff, on its medial border without a formal release of the vastus medialis [10]. The quadriceps sparing (QS) approach is a subvastus approach in which the vastus medialis is mobilized and separated from the intermuscular septum, allowing its retraction laterally [11,12].

However, mini invasive approaches add surgical time, can be more difficult to master with substantial learning curves [14], and have been associated with potential serious complications [15,16]. In addition, a number of randomized trials have suggested that they may not lead to a faster recovery [13,17], nor even improve it enough to justify the added time or the associated risk of complications of their use. Finally, available literature allows only head-to-head comparisons using randomized controlled studies (RCT) or meta-analysis. They have shown contradictory results [13,17–21]. To our knowledge, no study has simultaneously compared the different approaches. Network meta-analysis (NMA) allows the comparison of treatments that were not evaluated directly against each other in individual trials [22]. Therefore, it is a powerful tool to answer an important question when multiple treatment options are available.

In order to compare the different approaches available to the surgeons for TKA approach, we performed a NMA to determine whether mini invasive approaches provided any advantages over the medial parapatellar in terms of: [1] clinical outcome scores (1989 Knee Society Score KSS, WOMAC score) measured 6 months after surgery [2] knee range of motion (ROM) measured 6 months after surgery, [3] the surgical time [4] and other secondary outcomes.

Materials and Methods

This systematic review and NMA was registered in PROSPERO (Number CRD42019139318) and is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

Search Strategy

We searched all RCTs available in electronic databases (Medline, CENTRAL and Embase) starting at the inception date up to May 1, 2020, using dedicated search equations adapted to each database (Appendix 1). We manually searched three major orthopedic journals (the *Journal of Bone and Joint Surgery American Volume*, the *Journal of Arthroplasty*, and *Clinical Orthopedics and Related Research*) and conference proceedings of orthopedic meetings (the American Academy of Orthopedic Surgeons, the American Association of Hip and Knee Surgeons, and the Société Française de Chirurgie Orthopédique et Traumatologie) from January 1, 2005 to May 1, 2020. We also searched [ClinicalTrials.gov](https://www.clinicaltrials.gov) and the World Health Organization's International Clinical Trials Registry Platform. Finally, we systematically screened reference lists of systematic reviews and meta-analyses for any additional references.

Study Eligibility and Selection Process

All RCTs that compared at least two types of surgical approaches and reported at least one of our predefined outcomes were

included. No restriction on language or date of publication was applied. Exclusion criteria were TKA for another indication than osteoarthritis, TKA after knee osteotomy or unicompartmental knee arthroplasty on the same knee, and TKA revision.

Clinical outcomes, were all assessed at 6 months postoperatively in this review: subjective outcomes, evaluated with either the 1989 Knee Society Score (KSS), or the WOMAC score, and the ROM (Appendix 2).

Secondary radiological outcomes, measured on knee X-rays at 6 months postoperatively were: the proportion of HKA outliers, defined as an HKA angle outside the 177° to 183° range; misalignment of tibial and femoral implants in sagittal plane, defined as an angle more or less than 3° from the intended value within the first (Appendix 2). Other secondary clinical outcomes were assessed: the mean operative time; blood loss; length of stay; first post-operative day knee pain with the mean VAS score; 1989 KSS and WOMAC score at 1 and 2 years postoperatively, SF36 scores at 6 months and at 1 and 2 years postoperatively and the incidence of postoperative complications.

Two reviewers independently screened titles, summaries, and full texts whenever necessary to assess the eligibility of each study. Disagreements were resolved by a third reviewer and by contacting the corresponding author in case of persisting doubt.

Selection and General Network Characteristics

Our search initially retrieved 1829 records. Finally, 60 RCTs with 5042 patients and 5107 knees were included (Fig. 1). The most frequent comparison was midvastus versus parapatellar approaches in 22 of 60 (37%). The global network diagram illustrates that parapatellar were compared with every other technique (Fig. 2). Two trials had three arms (midvastus, QS, and parapatellar) (Appendix 3). The risk of bias of each study remains heterogeneous notably concerning the “sequence generation” and “allocation concealment” (Appendix 4).

Data Extraction and Risk of Bias Assessment

The same reviewers independently extracted data and evaluated the risk of bias using the Cochrane risk-of-bias tool [23] in a dedicated data extraction form that was tested on 10 random records. In case of disagreement, a third reviewer was involved to reach consensus.

Following data was collected: general study information; characteristics of the study; population characteristics; surgical approaches; prosthesis; specific surgical; and results for outcomes of interest. Regarding continuous outcomes, means or SD were estimated from the median and the first and third quartiles and/or ranges using approximation formulas [24,25] when not available (Appendix 5). The corresponding authors were contacted in case of missing data.

Patient Involvement

Patients were not involved in any aspect of the study design, conduct, or development of the research question or outcome measures. This study is a systematic review and NMA of existing published research. Thus, there was no active patient recruitment for data collection.

Data Synthesis and Analysis

The unit of analysis was the knee. Binary outcomes were compared using risk ratios (RR). Continuous outcomes were compared using mean differences (MD) with 95% credible intervals

ARTICLE IN PRESS

P.-A. Bouché et al. / The Journal of Arthroplasty xxx (2020) 1–8

3

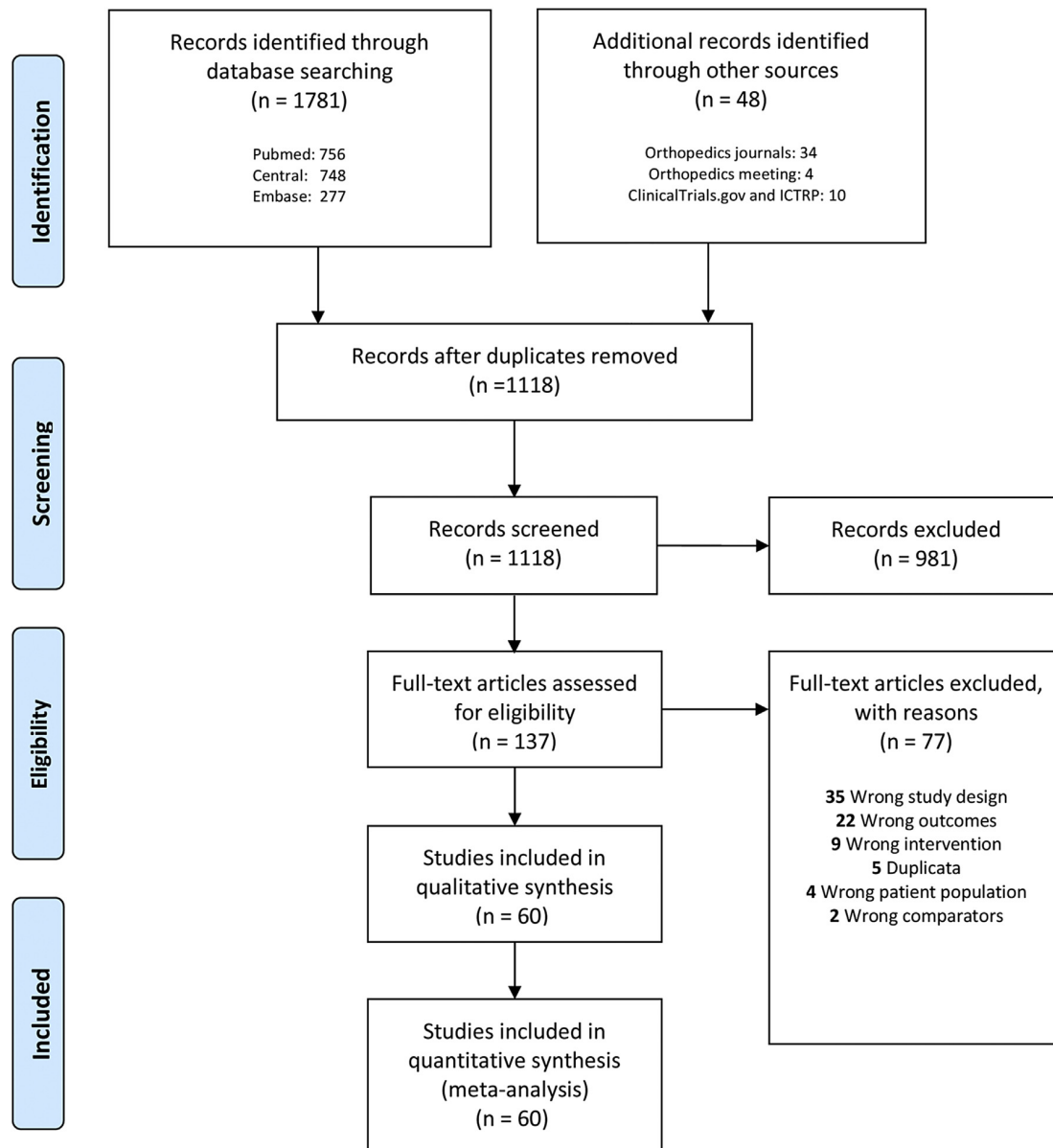


Fig. 1. This PRISMA flow chart shows the selection process of the network meta-analysis comparing the efficacy of different surgical approaches in TKA.

(CIs). For direct comparisons, we conducted a conventional meta-analysis to synthesize the results, using random-effects models and fixed-effects models as sensitivity analyses. We combined direct and indirect comparisons via a NMA using the hierarchical model of Lu and Ades [26] with a Bayesian approach. In a Bayesian framework, estimates are based on the mean of the posterior distribution of the endpoint and are called credible intervals (CIs). Usually, the 95% CI corresponding to the interval that has a posterior probability of 95% that the endpoint lies within it. Unlike the frequentist approach, the Bayesian approach does not allow the calculation of the P value. Conclusions are drawn from whether or not 1 (which corresponds to a no-difference relative risk) or 0 (which corresponds to no mean differences) belongs or not to the credible interval. We used Markov Chain Monte Carlo to implement the model. Beta non-informative prior distribution was used for

dichotomous outcomes and normal non-informative prior distribution was used for continuous outcomes [27]. The model's convergence was assessed using Brooks-Gelman-Rubin plots [28]. We evaluated the fit via residual deviance [29]. We checked the assumption of transitivity by considering similarity in the distribution of age, sex, and BMI across the different pairwise comparisons. A network diagram was used for each prespecified outcome to present direct comparisons between interventions. We obtained a treatment hierarchy using the surface under the cumulative ranking curve and mean ranks. To infer the assumption of inconsistency, we used a node-splitting approach [30], and we compared the fit between the consistency and inconsistency models. We statistically assessed the presence of heterogeneity within each network comparison using the I^2 statistic. We used the GEMTC R-package (version 0.8-2) for analyses and netmeta R-package

ARTICLE IN PRESS

4

P.-A. Bouché et al. / The Journal of Arthroplasty xxx (2020) 1–8

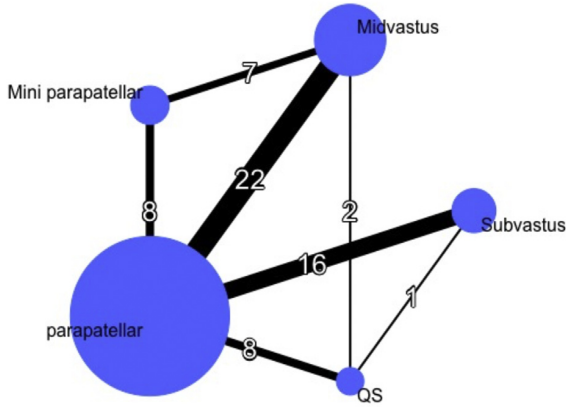


Fig. 2. In this global network diagram, the nodes (circles) represent different surgical approaches, and their sizes were proportional to the sample size of each respective intervention. The edges (lines) indicate direct comparisons, and their thicknesses were proportional to the number of studies available.

(version 0.9-8) for network diagrams in R version 3.5.2 software (R foundation for statistical Computing, Vienna, Austria). The quality of evidence for the main outcomes was evaluated using the GRADE framework [31].

Results

Functional Outcomes

We found no difference between the different surgical approaches either on KSS and WOMAC at 6 months.

Twenty-two studies (37%) reported the KSS at 6 months post-operatively (Fig. 3). The results of the pairwise comparison did not indicate any differences in functional outcomes (Appendix 6). Mini-parapatellar approach had the highest probability of leading to the highest KSS mean values (Fig. 4), although no significant differences were found (Fig. 5).

The WOMAC score at 6 months was available in six studies (10%) (Fig. 3). The results of the pairwise comparison did not indicate any differences (Appendix 6). Subvastus approach had the highest probability of leading to the highest WOMAC mean values (Fig. 4), although no significant differences were found (Fig. 6).

ROM at 6 months Post-surgery

Seventeen (28%) studies reported the ROM at 6 months post-surgery (Fig. 3). ROM from the subvastus group were superior to all other approaches. We found a difference of 7.3° (95% CI –14.1 to –0.1) with midvastus approach, of 11.1° (95% CI –18.7 to –2.8) with mini-parapatellar, of 8.9° (95% CI –14.2 to –3.1) with

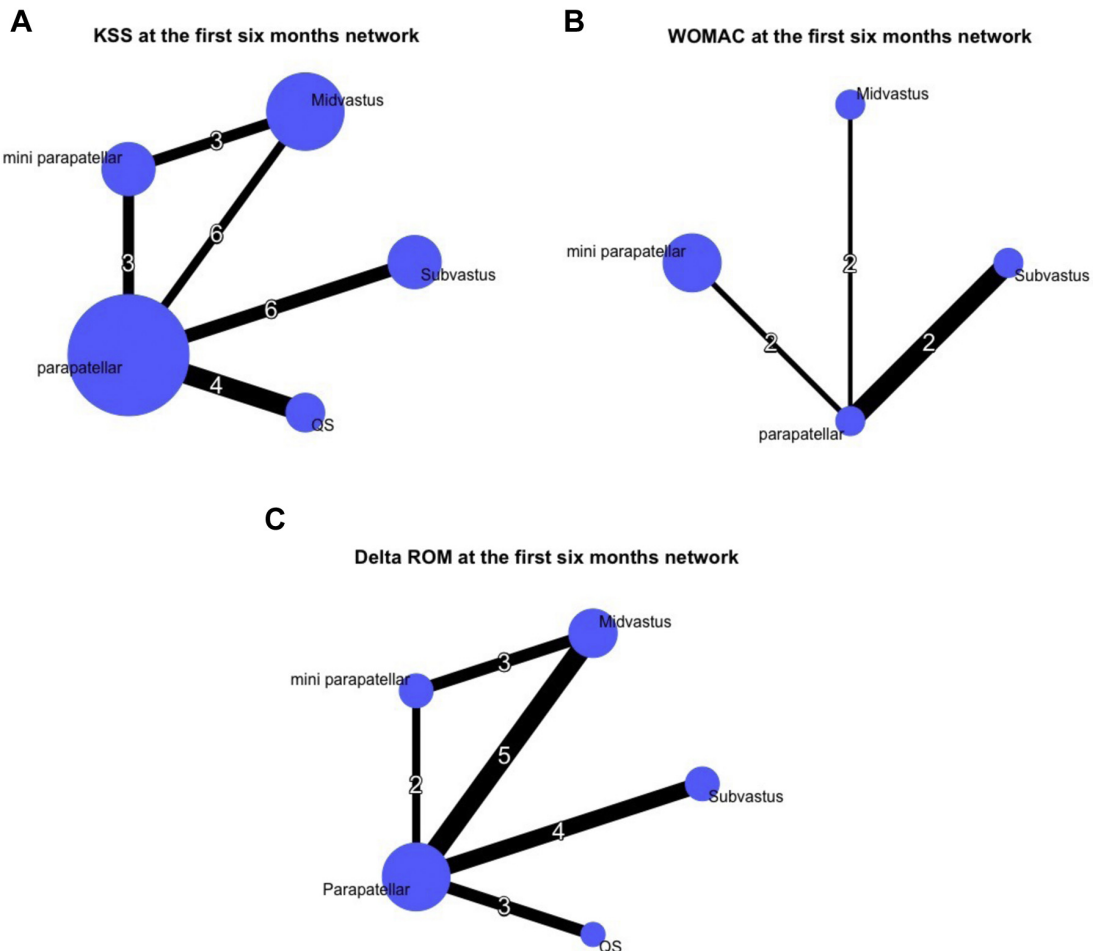


Fig. 3. A-C These network diagrams of the main outcomes show (A) KSS at the first 6 months, (B) WOMAC scores at the first 6 months, (C) ROM at the first 6 months.

ARTICLE IN PRESS

P.-A. Bouché et al. / The Journal of Arthroplasty xxx (2020) 1–8

5

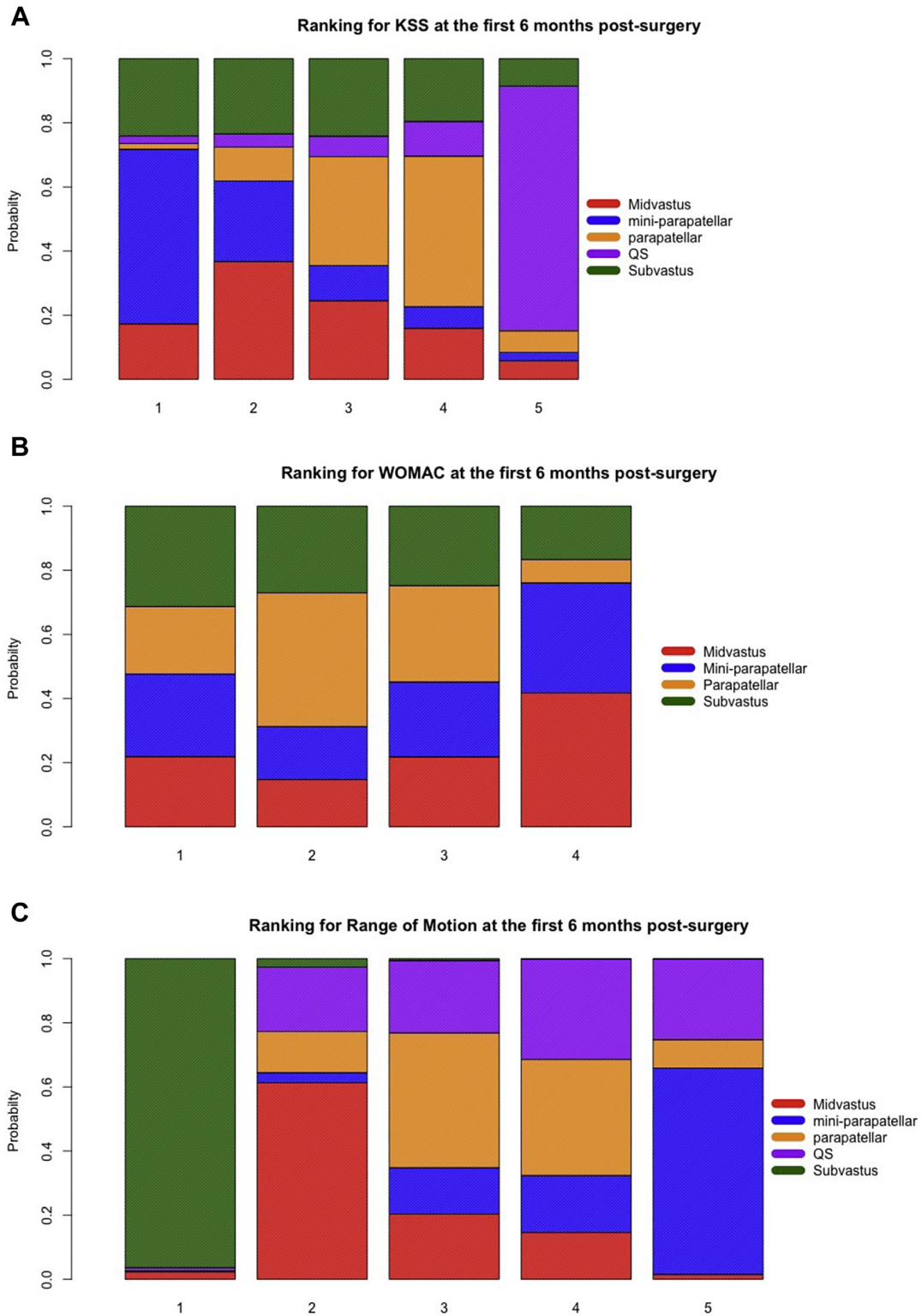


Fig. 4. A-C These images shows rank probabilities of (A) delta KSS at the first 6 months postoperatively, (B) delta WOMAC scores at the first 6 months postoperatively, and (C) delta ROM at the first 6 months postoperatively for each surgical approach. The x axis represents the ranking. The y axis represents the probability of rank: 0% to 100%. Each color represents one surgical approach: red for midvastus, blue for mini-parapatellar, orange for parapatellar, purple for QS, and green for the subvastus; QS = Quadriceps sparring.

ARTICLE IN PRESS

6

P.-A. Bouché et al. / The Journal of Arthroplasty xxx (2020) 1–8

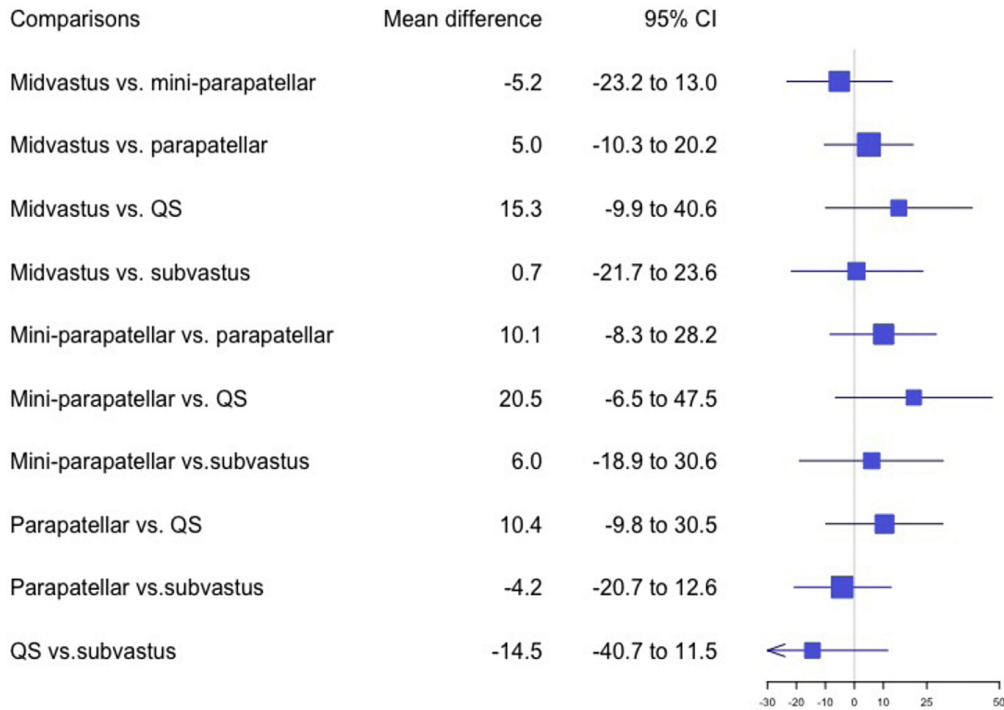


Fig. 5. Forest plots in the network meta-analysis of the mean difference of KSS between preoperatively and 6 months postoperatively. The data are presented as the relative mean differences with 95% credibility intervals, which were obtained using random effects models. The I^2 of the network meta-analysis was 0.9%.

parapatellar, and 9.2° (95% CI -16.1 to -1.8) with QS (Fig. 7). The results of pairwise comparisons confirmed that subvastus against the parapatellar approach were associated with an increased ROM at 6 months (MD 8.8 [95% CI 2.8 to 14.0]; $I^2 = 65.1\%$) (Appendix 6). Subvastus approach had the highest probability of allowing the highest ROM at 6 months (Fig. 4).

Secondary Outcomes

Regarding secondary clinical outcomes, mean KSS scores at 1 and 2 years postoperatively were higher in the mini-parapatellar group than parapatellar (MD 13.0 [95% CI 1.2 to 24.6] (Appendix 7). No difference was observed at this time regarding WOMAC at

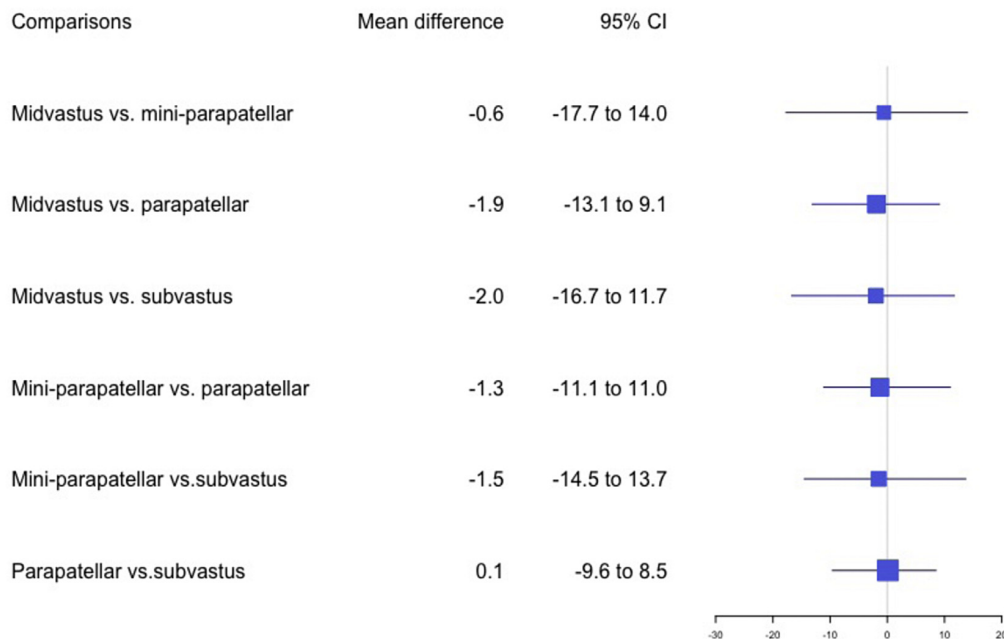


Fig. 6. Forest plots in the network meta-analysis of the mean difference of WOMAC between preoperatively and 6 months postoperatively. The data are presented as the relative mean differences with 95% credibility intervals, which were obtained using random effects models. The I^2 of the network meta-analysis was 14%.

ARTICLE IN PRESS

P.-A. Bouché et al. / The Journal of Arthroplasty xxx (2020) 1–8

7

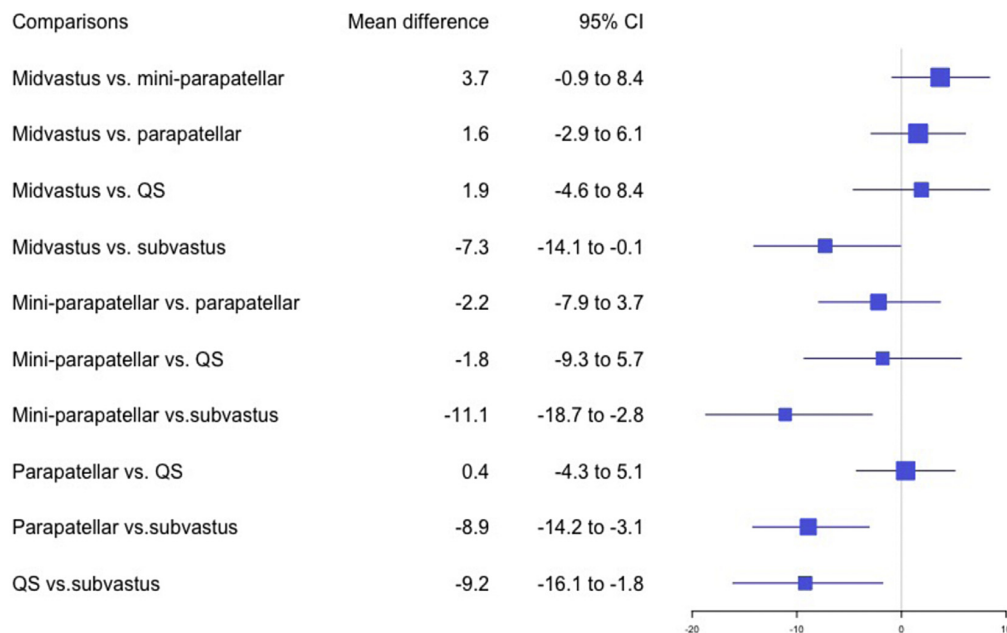


Fig. 7. Forest plots in the network meta-analysis of the mean difference of ROM between preoperatively and 6 months postoperatively. The data are presented as the relative mean differences with 95% credibility intervals, which were obtained using random effects models. The I^2 of the network meta-analysis was 12%.

1 and 2 years postoperatively (Appendix 7). Concerning the SF 36 at 6 months, 1 year and 2 years postoperatively, we had only data for physical scale and health perception. We found no difference between the different surgical approaches either SF36 scores at short or mid-term (Appendix 7).

Subvastus approach reduced the HKA $\pm 3^\circ$ outliers' rate, compared with QS (RR 0.18 [95% CI 0.0 to 0.90]) without any other difference between other approaches. Sagittal components alignment was no different between groups (Appendix 8).

Surgical procedures using parapatellar approach lasted shorter than those using midvastus (MD 9.1 [95% CI 3.2 to 15.2]), subvastus (MD 10.6 [95% CI 1.4 to 19.8]) and QS (MD 11.9 [95% CI 3.1 to 20.6]) (Appendix 9), but shown no difference with mini-parapatellar approach.

Blood loss was not different between groups.

First post-operative day mean VAS seemed to be lower with subvastus approach compared to parapatellar (MD 1.3 [95% CI 0.3 to 2.3]) (Appendix 9).

Length of stay was shorter using the mini-parapatellar approach compared to parapatellar (MD 3.0 [95% CI 1.2 to 4.8]) midvastus (MD 2.4 [95% CI 0.5 to 4.2]), subvastus (MD 2.6 [95% CI 0.3 to 4.7]) and QS approaches (MD 2.4 [95% CI 0.1 to 4.6]) (Appendix 9).

We finally observed more complications in midvastus TKAs compared to parapatellar (RR 1.6 [95% CI 1.2 to 2.3]) and subvastus interventions (RR 1.9 [95% CI 1.2 to 2.3]) (Appendix 9).

Grade Evaluation

We performed the GRADE evaluation for the three main outcomes. For KSS at the 6 months post-surgery, it indicates that the evidence is moderate for parapatellar/QS and low for other comparisons (Appendix 11). For WOMAC at the 6 months post-surgery, it suggests that the evidence is moderate for midvastus/parapatellar and low for other comparisons (Appendix 11). For ROM at the 6 months post-surgery, it suggests that the evidence is low for

mini-parapatellar/parapatellar and moderate for others comparisons (Appendix 11).

Discussion

No difference was observed between the different approaches for TKA. Despite the potential advantage of preserving the quadriceps tendon or reducing the wound length, mini invasive and standard medial parapatellar approaches lead to similar functional outcome at 6 months up to two years postoperatively.

In the current study, we were not able to demonstrate, contrary to head-to-head comparisons that early functional recovery [32] and reduction in pain [20] using minimally invasive approaches were improved compared to the classic parapatellar. Moreover, it is important to observe that those goals were not achieved without losing sight of the fact that they show potential disadvantages themselves such as added to an increased surgical time, a difficult learning curve [13–15].

A NMA was a useful approach to synthesize a large amount of data in this setting because it allowed the comparison of approaches that were not compared directly in individual trials. The different comparisons allowed by the NMA exhibited few significant results. Regarding primary outcomes, the subvastus approach was found to be the only technique showing significant positive differences in ROM at 6 months with the other approaches, but it was not associated with improvements in validated knee clinical scores. On the other hand, our comparisons showed that the mini medial parapatellar approach led to a KSS improvement and a reduction of the hospital stay compared to the standard approach. Concerning secondary outcomes, subvastus approach also reduced the early post-operative pain compared to the parapatellar approach, and mini parapatellar technique allowed a reduction of the hospital stay compared to the standard approach. Those benefits came however with an increased surgical time of approximately 10 minutes compared to the parapatellar approach.

47]. In the interest of making alignment as consistent as possible, surgeons therefore have explored a variety of approaches, including navigation, accelerometers, patient-specific instrumentation, and robotics. It is important to determine the degree to which these approaches achieve those goals without losing sight of the fact that these approaches also have potential disadvantages, including increased cost, increased surgical time, a difficult learning curve, and unique—and sometimes severe—complications [2, 5]. A network meta-analysis is a useful approach to synthesize a large amount of data in this setting because this approach allows the comparison of approaches that were not compared directly in individual trials. We found an incremental reduction in the risk of small ($\pm 3^\circ$) HKA outliers when navigation was used compared with the other approaches, but no associated improvements in validated outcomes scores. The small alignment benefit came with a large increase in surgical time. Until or unless future studies can demonstrate a benefit in terms of survivorship, knee pain, cost effectiveness, or function of navigation for these other newer approaches, based on our findings we recommend against their widespread use.

Limitations

However, this study had several limitations. Few trials evaluated more recent techniques, such as the use of an accelerometer [11, 17, 21, 34, 43] or robot [25, 26], making it difficult to assess the benefits of these two techniques in relation to their cost. Unfortunately, we were unable to collect clinical outcomes after 2 years postoperatively due to the lack of information in published studies, even though joint prostheses are expected to have a longevity of 15 years or more. Finally, the ranking results were based on point estimate probabilities, and the results do not fully reflect the uncertainty of the ranking as outlined by Trinquart et al. [45]. As a consequence, we must consider ranking results as trends rather than definitive conclusions.

We found that many trials did not evaluate clinical outcomes; only 24 trials reported clinical outcomes. Only three trials simultaneously evaluated our three primary outcomes, illustrating the high level of heterogeneity regarding the outcomes planned and reported. In addition, the reporting of clinical outcomes was notably heterogeneous across trials. The WOMAC was reported with different numeric scales, whereas the KSS was more homogeneously used and reported in the included RCTs [27]. Such heterogeneity in evaluated outcomes and their definitions makes it difficult to compare results between trials and combinations by meta-analysis. This highlights the need to develop recommendations and core sets of outcomes as promoted by the Core Outcome Measures in Effectiveness Trials (COMET) initiative [10] in the field of knee prosthetic replacement.

Small Improvement in HKA Alignment

We found that about 10% more patients could achieve an HKA of within 3° of neutral when navigation was used compared with when standard guides or patient-specific instruments were used. We caution readers against making too much of even this incremental difference. The evidence is scant and contradictory that such small alignment differences are likely to improve implant survivorship or function [9, 20, 31] and even larger differences—such as $\pm 5^\circ$ —may not matter as much as once was thought [36]. We would have preferred to set the cutoff to 5° or even analyze it as a continuous variable, but the data were not available in the studies we evaluated. Reducing the likelihood of HKA outliers in a small minority of patients undergoing TKA by widespread adoption of navigation would result in substantial increases in costs and operative time [2], at least initially.

No Improvements in Outcomes Scores

We found no evidence that these technologies improved KSS or WOMAC scores at either 6 months, 1 year, or 2 years after TKA. For clinical outcomes, similar results were found in other studies regarding KSS comparing navigation, patient-specific instrumentation, and standard cutting guides [28, 37]. Petursson et al. [37] found that navigation had a better WOMAC after 2 years. We did not find similar results at short or medium terms.

No Increase in Complications but Longer Surgical Time

Navigation is often criticized because of increased operative time and the risk of iatrogenic fracture due to the pin insertions. In this network meta-analysis, navigation increased the mean operative time by 20 minutes compared with standard cutting guides and patient-specific instrumentation. However, postoperative complications were similar among all cutting guides. We were probably underpowered to detect rare but serious complications. Thus, in our view, the benefit/risk balance between a more precise tool, such as navigation, allowing more reproducible cuts but without evidence of an impact on function and an increase of surgical time does not favor the navigation approach.

Conclusions

No new approach in this study demonstrated superiority over standard cutting guides on functional outcome, and

the alignment differences affected few patients, were small in magnitude, and were of questionable clinical importance. Thus, unless these new approaches either convincingly demonstrate superior survivorship or convincingly demonstrate superior long-term functional outcomes, there is no reason to use these approaches from either the patient's point of view or the hospital's point of view since they add cost, have a learning curve (during which some patients may be harmed), and have the risks associated with novelty.

Acknowledgments We thank Professor Sylvie Chevret MD, PhD, director of the biostatistics department in the Saint Louis Hospital, for her advice. Furthermore, we thank every corresponding author who made the effort to provide us with the requested supplementary data.

References

- Banerjee S, Cherian JJ, Elmallah RK, Jauregui JJ, Pierce TP, Mont MA. Robotic-assisted knee arthroplasty. *Expert Rev Med Devices*. 2015;12:727–735.
- Bauwens K, Matthes G, Wich M, Gebhard F, Hanson B, Ekkernkamp A, Stengel D. Navigated total knee replacement. A meta-analysis. *J Bone Joint Surg Am*. 2007;89:261–269.
- Berger RA, Crosssett LS, Jacobs JJ, Rubash HE. Malrotation causing patellofemoral complications after total knee arthroplasty. *Clin Orthop Relat Res*. 1998; 356:144–153.
- Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. *J Comput Graph Stat*. 1998;7:434–355.
- Clark GW, Wood DJ. Robotics in arthroplasty: Where are we today? *Bone & Joint* 360. 2015;4:2–7.
- Deep K, Shankar S, Mahendra A. Computer assisted navigation in total knee and hip arthroplasty. *SICOT-J*. 2015;3:50.
- Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med*. 2010;29:932–944.
- Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Mak*. 2013;33:641–656.
- Fu Y, Wang M, Liu Y, Fu Q. Alignment outcomes in navigated total knee arthroplasty: a meta-analysis. *Knee Surg Sports Traumatol Arthrosc*. 2012;20:075–1082.
- Gargon E, Williamson PR, Altman DG, Blazeby JM, Tunis S, Clarke M. The COMET Initiative database: progress and activities update (2015). *Trials*. 2017; 18:54.
- Gharaibeh MA, Solayar GN, Harris IA, Chen DB, MacDessi SJ. Accelerometer-based, portable navigation (KneeAlign) vs conventional instrumentation for total knee arthroplasty: a prospective randomized comparative trial. *J Arthroplasty*. 2017;32:777–782.
- Haglin JM, Eltorai AEM, Gil JA, Marcaccio SE, Botero-Hincapie J, Daniels AH. Patient-specific orthopaedic implants. *Orthop Surg*. 2016;8:417–424.
- Higgins JPT, Altman DG, Gotzsche PC, Jüni P, Moher D, Oxman AD, Savović J, Schulz KF, Weeks L, Sterne JAC, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;18:343:d5928.
- Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2015;5:13
- Huang EH, Copp SN, Bugbee WD. Accuracy of a handheld accelerometer-based navigation system for femoral and tibial resection in total knee arthroplasty. *J Arthroplasty*. 2015;30:1906–1910.
- Huijbregts HJTAM, Khan RJK, Sorensen E, Fick DP, Haebich S. Patient-specific instrumentation does not improve radiographic alignment or clinical outcomes after total knee arthroplasty. *Acta Orthop*. 2016;87:386–394.
- Ikawa T, Takemura S, Kim M, Takaoka K, Minoda Y, Kadoya Y. Usefulness of an accelerometer-based portable navigation system in total knee arthroplasty. *Bone Joint J*. 2017;99:1047–1052.
- Jacofsky DJ, Allen M. Robotics in arthroplasty: A Comprehensive Review. *J Arthroplasty*. 2016;31:2353–2363.
- Jeffery RS, Morris RW, Denham RA. Coronal alignment after total knee replacement. *J Bone Joint Surg Br*. 1991;73:709–714.
- Kim YH, Kim JS, Yoon SH. Alignment and orientation of the components in total knee replacement with and without navigation support: a prospective, randomised study. *J Bone Joint Surg*. 2007;89:471–476.
- Kinney MC, Cidambi KR, Severns DL, Gonzales FB. Comparison of the iAssist handheld guidance system to conventional instruments for mechanical axis restoration in total knee arthroplasty. *J Arthroplasty*. 2018;33:61–66.
- Kuzyk PRT, Higgins GA, Tunggal JAW, Sellan ME, Waddell JP, Schemitsch EH. Computer navigation vs extramedullary guide for sagittal alignment of tibial components: radiographic study and meta-analysis. *J Arthroplasty*. 2012;27:630–637.
- Laurent P. Patient-specific instruments in orthopaedics. In: Ritacco LE, Milano FE, Chao E, eds. *Computer-assisted Musculoskeletal Surgery*. Philadelphia, PA: Springer; 2016: 163–179.
- Leopold SS. Pencil and paper research? Network meta-analysis and other study designs that do not enroll patients. *Clin Orthop Relat Res*. 2015;473:2163–2165.
- Liow MHL, Goh GSH, Wong MK, Chin PL, Tay DKJ, Yeo SJ. Robotic-assisted total knee arthroplasty may lead to improvement in quality-of-life measures: a 2-year follow-up of a prospective randomized trial. *Knee Surg Sports Traumatol Arthrosc Off*. 2017;29:2942–2951.
- Liow MHL, Xia Z, Wong MK, Tay KJ, Yeo SJ, Chin PL. Robot-assisted total knee arthroplasty accurately restores the joint line and mechanical axis: A prospective randomised study. *J Arthroplasty*. 2014;29:2373–2377.
- Lovelock TM, Broughton NS, Williams CM. The popularity of outcome measures for hip and knee arthroplasties. *J Arthroplasty*. 2018;33:273–276.
- Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. 2004;23:3105–3124.
- Mannan A, Akinyooye D, Hossain F. A Meta-analysis of functional outcomes in patient-specific instrumented knee arthroplasty. *J Knee Surg*. 2017;30:668–674.
- Mannan A, Smith TO. Favourable rotational alignment outcomes in PSI knee arthroplasty: A Level 1 systematic review and meta-analysis. *The Knee*. 2016;23:186–190.
- Mason JB, Fehring TK, Estok R, Banel D, Fahrback K. Meta-analysis of alignment outcomes in computer-assisted total knee arthroplasty surgery. *J Arthroplasty*. 2007;22:1097–1106.
- Mattei L, Pellegrino P, Calò M, Bistolfi A, Castoldi F. Patient specific instrumentation in total knee arthroplasty: a state of the art. *Ann Transl Med*. 2016;4:126–126.
- Moskal J, Capps S, Mann J, Scannelli J. Navigated versus conventional total knee arthroplasty. *J Knee Surg*. 2013;27: 235–248.

34. Nam D, Cody EA, Nguyen JT, Figgie MP, Mayman DJ. Extramedullary guides versus portable, accelerometer-based navigation for tibial alignment in total knee arthroplasty: A randomized, controlled trial: Winner of the 2013 HAP PAUL award. *J Arthroplasty*. 2014;29:288–294.
35. Parratte S, Blanc G, Boussemart T, Ollivier M, Corroller TL, Argenson JN. Rotation in total knee arthroplasty: no difference between patient-specific and conventional instrumentation. *Knee Surg Sports Traumatol Arthrosc*. 2013;21:2213–2219.
36. Parratte S, Pagnano MW, Trousdale RT, Berry DJ. Effect of postoperative mechanical axis alignment on the fifteen-year survival of modern, cemented total knee replacements. *J Bone Joint Surg Am*. 2010;92:2143–2149.
37. Petursson G, Fenstad AM, Gothesen O, Dyrhovden GS, Hallan G, Rohrl SM, Aamodt A, Furnes O. Computer-assisted compared with conventional total knee replacement: a multicenter parallel-group randomized controlled trial. *J Bone Joint Surg Am*. 2018;100:1265–1274.
38. Rebal BA, Babatunde OM, Lee JH, Geller JA, Patrick DAJ, Macaulay W. Imageless computer navigation in total knee arthroplasty provides superior short term functional outcomes: a meta-analysis. *J Arthroplasty*. 2014;29:938–944.
39. Ritter MA, Faris PM, Keating EM, Meding JB. Postoperative alignment of total knee replacement. Its effect on survival. *Clin Orthop Relat Res*. 1994;299:153–156.
40. Salanti G, Giovane CD, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the Quality of Evidence from a Network Meta-Analysis. *PLoS One*. 2014;9:e99682.
41. Sikorski JM. Alignment in total knee replacement. *J Bone Joint Surg Br*. 2009;90:1121–1127.
42. Stulberg SD, Loan P, Sarin V. Computer-assisted navigation in total knee replacement: results of an initial experience in thirty-five patients. *J Bone Joint Surg Am*. 2002;84(A Supple 2):90–98.
43. Thiengwittayaporn S, Fusaku Y, Kangkano N, Jarupongprapa C, Charoenphandhu N. Hand-held navigation may improve accuracy in minimally invasive total knee arthroplasty: a prospective randomized controlled trial. *Int Orthop*. 2016;40:51–57.
44. Thienpont E, Schwab PE, Fennema P. Efficacy of patient-specific instruments in total knee arthroplasty: a systematic review and meta-analysis. *J Bone Joint Surg Am*. 2017;99:521–530.
45. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in treatment rankings: reanalysis of network meta-analyses of randomized trials. *Ann Intern Med*. 2016;164:666–673.
46. Van Valkenhoef and Joel Kuiper. gemtc: Network Meta-Analysis Using Bayesian Methods. *R package version 0.8-2*. <https://CRAN.R-project.org/package=gemtc>. Accessed March 20, 2020.
47. Vanlommel L, Vanlommel J, Claes S, Bellemans J. Slight undercorrection following total knee arthroplasty results in superior clinical outcomes in varus knees. *Knee Surg Sports Traumatol Arthrosc*. 2013;21:2325–2330.
48. Weir CJ, Butcher I, Assi V, Lewis SC, Murray GD, Langhorne P, Brady MC. Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review. *BMC Med Res Methodol*. 2018;18:25.
49. Victor J, Dujardin J, Vandenneucker H, Arnout N, Bellemans J. Patient-specific guides do not improve accuracy in total knee arthroplasty: A prospective randomized controlled trial knee. *Clin Orthop Relat Res*. 2014;472:263–271.

Chapitre 6

Évaluation des guides de coupe dans les prothèses totales de genoux

Nous nous sommes ensuite intéressés aux différents guides de coupe dans les prothèses totales de genoux. Les chirurgiens s'aident de ces guides de coupe afin de positionner correctement les implants dans les trois plans de l'espace. Ils permettent de guider la scie oscillante lors de la réalisation des coupes osseuses.

Actuellement, sur le marché, il existe cinq grandes catégories de guide de coupe : les guides mécaniques, la navigation, les guides de coupes sur mesure ou « PSI » (Patient Specific Instrumentation), les accéléromètres et la robotique (Figure 4).

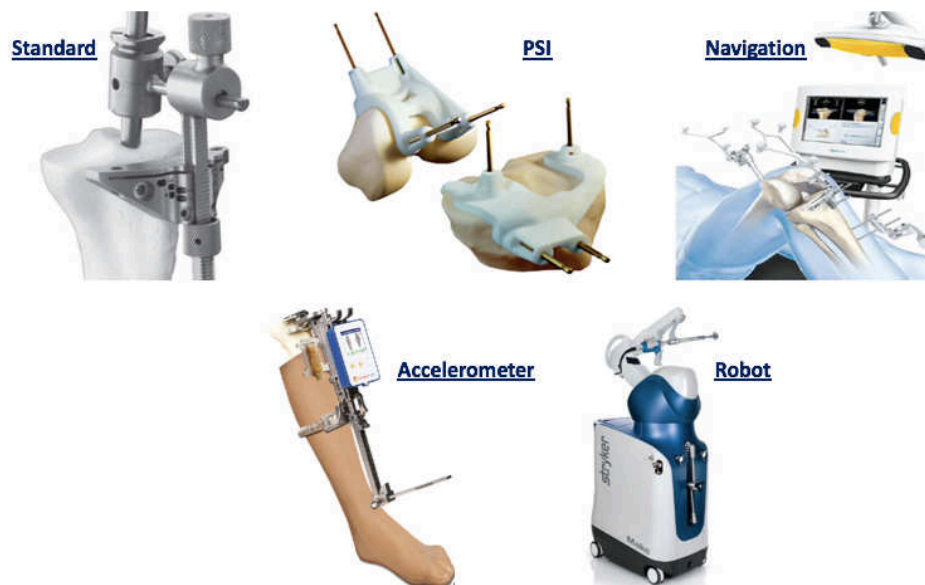


Figure 4. Représentation schématique des différents guides de coupes analysés dans l'article. De gauche à droite, guide de coupe standard, PSI, la navigation, l'accéléromètre et la robotique.

Le guide de coupe historique est la technique conventionnelle ou mécanique. Cette technique consiste à utiliser des repères anatomiques et les axes anatomiques du fémur et du tibia pour déterminer l'épaisseur des coupes osseuses. Cette coupe permettra le positionnement des futurs implants. Deux principaux types de guides de coupe standard existent pour le tibia, pour déterminer son axe anatomique: le système d'alignement intra et extra médullaire (Zeng et al., 2015). Le système intra médullaire utilise une longue tige métallique insérée dans le canal médullaire du tibia. Le système extra-médullaire utilise une tige parallèle à la crête tibiale dans le plan sagittal et du milieu du genou au milieu de la cheville dans l'axe du 2ème métatarsien dans le plan frontal. Il n'y a pas de différence de précision d'alignement entre les deux systèmes (Haglin et al., 2016; Mattei et al., 2016). Les guides conventionnels sont les plus répandus dans le monde et les moins coûteux. Ils sont faciles d'accès et leur utilisation ne nécessite pas d'aides extérieures. Leur principal inconvénient est que le chirurgien n'a finalement que très peu de recul sur ce qu'il est en train de faire. Il n'y a pas de contrôle peropératoire fiable.

La deuxième catégorie de guide de coupe apparue est la navigation. C'est une chirurgie assistée par ordinateur qui nécessite ou non la réalisation d'une imagerie préopératoire. La navigation fut développée dans les années 1990 (Deep et al., 2017). Elle permet, à l'aide de capteurs fémoraux et tibiaux positionnés en percutanée, aux chirurgiens d'évaluer en temps réel, les coupes, l'axe mécanique du membre opéré et la position des implants. Certaines navigations permettent aussi de contrôler la tension ligamentaire. La navigation permet d'augmenter la précision des coupes et d'avoir un contrôle peropératoire pour le chirurgien. Le matériel et sa maintenance rendent les coûts plus importants comparés au guide conventionnel. Certains auteurs rapportent aussi une augmentation du temps opératoire et de complications propres aux fiches de la navigation (fractures).

Le troisième guide de coupe existant est le PSI (Haglin et al., 2016). Il s'agit de guides de coupe spécifiques au patient fait sur mesure. Ils sont fabriqués à partir d'une modélisation tridimensionnelle du genou du patient, obtenue à l'aide d'une imagerie réalisée en préopératoire. Cet examen peut-être un scanner ou une IRM (Imagerie par Résonance Magnétique). Le chirurgien planifie ses coupes en amont de l'intervention. Cette technique chirurgicale permet en peropératoire un gain de temps et parfois diminue les coûts avec l'utilisation d'un ancillaire à usage unique. Cependant, sa diffusion reste médiocre car il faut avoir accès facilement à ses différents examens d'imagerie et leur conception dépende d'un laboratoire.

Le quatrième guide de coupe est l'accéléromètre. C'est un guide de coupe qui a été conçu pour remédier à certains inconvénients de la navigation. Il permet de déterminer uniquement l'alignement frontal dans les PTG sans utiliser la technologie de morphing osseux. Cela pourrait être assimilé à un système de navigation simplifié.

Enfin le dernier guide de coupe qui se développe et se diffuse ces dernières années est la robotique (Banerjee et al., 2015). Malgré l'intérêt récent pour la robotique en orthopédie, les premiers robots ont été développés au début des années 2000 avec l'apparition de l'ACROBOT®. Du fait, d'un coût important, de la nécessité d'un ingénieur en salle d'intervention et d'un temps procédure, la robotique a été laissée de côté. C'est environ 10 ans plus tard, que de nouveaux robots sont apparus sur le marché : Navio® en 2012 et Mako® en 2013. À l'aide d'une imagerie préopératoire ou le « bone morphing » peropératoire, les coupes osseuses sont réalisées par une fraise guidée par un bras robotisé. Le fraisage s'arrête une fois que la zone ciblée est atteinte, selon le planning du chirurgien. Il existe trois grands types de robots pour les arthroplasties de genou : les systèmes passifs, semi-actifs ou actifs. L'avantage de la robotique est la possibilité de contrôler les coupes dans les trois plans de l'espace ainsi que la tension ligamentaire. Toutefois, son coût, l'encombrement spatial et la nécessité d'un ingénieur en salle sont des inconvénients non négligeables. À l'heure actuelle, les bénéfices réels de la chirurgie robotique n'ont pas encore été prouvés scientifiquement.

Ces guides de coupe, dans les prothèses totales de genoux, n'ont jamais été comparés les uns aux autres. Nous avons utilisé une méta-analyse en réseau afin de les comparer et de voir si un guide de coupe apporte un avantage par rapport aux autres.

Le développement de nouvelles technologies et de nouvelles techniques dans le domaine médical et notamment prothétique est toujours attrayant autant pour le praticien que pour le patient. Malgré la pression parfois commerciale et de la mode, il est toujours important d'évaluer la balance bénéfice/risque des innovations et leur apport réel pour le chirurgien et surtout pour le patient.

De nombreux essais thérapeutiques ou méta-analyses comparant deux à deux ces différents guides de coupe ou voies d'abord dans les PTG ont été menés (Alvand et al., 2018; Blyth et al., 2017; Ollivier et al., 2016; Song et al., 2016). Ils retrouvent parfois des résultats contradictoires ou manquent parfois de puissance (du fait du faible volume des séries) (Thienpont E. et al.,

2017; Huijbregts et al., 2016; Mannan A. and Smith T.O., 2016; Parratte et al., 2013; Kim et al., 2007; Fu et al., 2012; Kuzyk et al., 2012). De plus, à notre connaissance aucune étude n'a réussi à comparer simultanément ces différentes interventions l'une à l'autre. En effet, il est difficile de mettre en place un essai randomisé comparant plusieurs techniques différentes.

Dans ce contexte, la réalisation d'une méta-analyse en réseau nous a semblé être la méthodologie la plus appropriée afin de répondre à ces différentes questions (Leopold, 2015). En effet, cette approche méthodologique permet de comparer plusieurs traitements qui n'ont pas été comparés directement l'un à l'autre dans des précédents essais. Autrement dit, si l'approche A a été comparée à l'approche B dans un certain nombre d'essais randomisés, et l'approche B a été comparée à l'approche C dans d'autres, une méta-analyse en réseau permet de comparer l'approche A à l'approche C et d'avoir des résultats sur cette comparaison indirectement. C'est une manière puissante de répondre à des questions cliniques importantes en réutilisant des données déjà publiées dans la littérature.

L'objectif de ce travail était de comparer ces différentes techniques chirurgicales à l'aide d'une méta-analyse en réseau.

La référence de l'article présenté est :

Bouché P-A, Corsia S, Dechartres A, Resche-Rigon M, Nizard R. Are There Differences in Accuracy or Outcomes Scores Among Navigated, Robotic, Patient-specific Instruments or Standard Cutting Guides in TKA? A Network Meta-analysis. Clin Orthop Relat Res. 2020;478(9):2105-2116. doi:10.1097/CORR.0000000000001324

Systematic Review or Meta-analysis

Are There Differences in Accuracy or Outcomes Scores Among Navigated, Robotic, Patient-specific Instruments or Standard Cutting Guides in TKA? A Network Meta-analysis

Pierre-Alban Bouché BS, Simon Corsia BS, Agnès Dechartres MD, PhD, Matthieu Resche-Rigon MD, PhD, Rémy Nizard MD, PhD

Received: 21 August 2019 / Accepted: 6 May 2020 / Published online: 8 June 2020
Copyright © 2020 by the Association of Bone and Joint Surgeons

Abstract

Background Several kinds of cutting guides, including patient-specific instrumentation, navigation, standard cutting guides, accelerometer-based navigation, and robotic guidance, are available to restore a planned alignment during TKA. No previous study has simultaneously

compared all of these devices; a network meta-analysis is an especially appealing method because it allows comparisons across approaches that were not compared head-to-head in individual randomized controlled trials.

Questions/purposes We performed a network meta-analysis to determine whether novel approaches to achieving implant alignment, such as patient-specific instrumentation, navigation, accelerometer-based navigation, and robotic guidance, provide any advantage over standard cutting guides in terms of: (1) hip-knee-ankle (HKA) alignment outliers greater than $\pm 3^\circ$, (2) outcome scores (1989 - Knee Society Score and WOMAC score) measured 6 months after surgery, or (3) femoral and tibial implant malalignment (greater than $\pm 3^\circ$), taken separately, in the frontal and sagittal plane, as well as other secondary outcomes including validated outcome scores 1 and 2 years after surgery.

Methods In our network meta-analysis, we included randomized controlled trials comparing the different cutting guides by using at least one of the previously specified criteria, without limitation on language or date of publication. We searched electronic databases, major orthopaedic journals, proceedings of major orthopaedic meetings, [ClinicalTrials.gov](https://www.clinicaltrials.gov/), and the World Health Organization's International Clinical Trials Registry Platform until October 1, 2018. This led to the inclusion of 90 randomized controlled trials involving 9389 patients (mean age 68.8 years) with 10,336 TKAs. Two reviewers independently selected trials and extracted data. The primary outcomes were the proportion patients with malalignment of the HKA angle (defined as $HKA > 3^\circ$ from neutral) and the Knee Society Score and WOMAC scores at 6 months postoperatively. We combined direct

Each author certifies that neither he or she, nor any member of his or her immediate family, has funding or commercial associations (consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

Each author certifies that his or her institution waived approval for the reporting of this investigation and that all investigations were conducted in conformity with ethical principles of research.

This work was performed at Hôpital Saint-Louis, INSERM, Paris, France.

P.-A. Bouché, M. Resche-Rigon, Service de Biostatistique et Information Médicale, Hôpital Saint-Louis, APHP, Université Paris Diderot, ECSTRA Team, UMR U1153, INSERM, Paris, France

P.-A. Bouché, S. Corsia, R. Nizard, Service de chirurgie orthopédique et traumatologique, Hôpital Lariboisière, Paris, France

A. Dechartres, Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Département Biostatistique Santé Publique et Information Médicale, Hôpitaux Universitaires Pitié Salpêtrière – Charles Foix, AP-HP, Paris, France

M. Resche-Rigon (✉), INSERM U1153, Hôpital Saint-Louis, Bâtiment Lailler 3ème étage, 1, avenue Claude Vellefaux, 75010 Paris, France, Email: matthieu.resche-rigon@univ-paris-diderot.fr

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

and indirect comparisons using a Bayesian network meta-analysis framework to assess and compare the effect of different cutting guides on outcomes. Bayesian estimates are based on the posterior distribution of an endpoint and are called credible intervals. Usually the 95% credible interval, corresponding to a posterior probability of 0.95 that the endpoint lies in the interval, is computed. Unlike the frequentist approach, the Bayesian approach does not allow the calculation of the *p* value.

Results The proportion of HKA outliers was lower with navigation than with patient-specific instrumentation (risk ratio 0.46 [95% credible interval (CI) 0.34 to 0.63]) and standard cutting guides (risk ratio 0.45 [95% CI 0.37 to 0.53]); however, this corresponded to an actual difference of only 12% of patients for navigation versus 21% of patients for patient-specific instrumentation, and 12% of patients for navigation versus 25% for standard cutting guides. We found no differences for other comparisons between different cutting guides, including robotics and the accelerometer. We found no differences in the Knee Society Score or WOMAC score between the different cutting guides at 6 months. Regarding secondary outcomes, navigation reduced the risk of frontal and sagittal malalignments for femoral and tibial components compared with the standard cutting guides, but none of the other cutting guides showed superiority for the other secondary outcomes.

Conclusions Navigation resulted in approximately 10% fewer patients having HKA outliers of more than 3°, without any corresponding improvement in validated outcomes scores. It is unknown whether this incremental reduction in the proportion of patients who have alignment outside a window that itself has been called into question will justify the increased costs and surgical time associated with the approach. We believe that until or unless these new approaches either (1) convincingly demonstrate superior survivorship, or (2) convincingly demonstrate superior outcomes, surgeons and hospitals should not use these approaches since they add cost, have a learning curve (during which some patients may be harmed), and have the risks associated with uncertainty of novel surgical approaches.

Level of Evidence Level I, therapeutic study.

Introduction

Achieving accurate alignment during TKA is important because of concerns for premature loosening [19]. In addition, function may be impaired in the setting of marked malalignment [39]. Nevertheless, the importance of precise alignment remains controversial. Some historical series found an association between a neutral alignment and aseptic loosening [3, 19, 39, 41]. Conversely, more

recent long-term series did not find this association [36, 47].

Five types of cutting guides are currently available. Standard cutting guides are the oldest and the most commonly used option. These guides rely on manual alignment relative to the intramedullary canal or landmarks palpated at the time of surgery (on the tibia) and angles determined in preoperative planning (on the femur). Introduced in 1997, navigation allows surgeons to have real-time information on alignment during surgery, using a system that includes trackers attached to bone followed by the use of an infrared camera and software to interpret the data [6, 42]. An accelerometer is a simplified navigation system because it allows the surgeon to determine the knee's frontal alignment without implantation of trackers in the patient's limb [15]. Developed during the 2010s, patient-specific instrumentation uses custom-fit cutting guides manufactured from a three-dimensional (3-D) model of the patient's knee determined with MRI or CT images obtained before surgery [12, 23, 32]. Finally, there was a resurgence of interest in modern robotic guidance in the last decade. It is the most recently developed technique, in which bony cuts are performed by a robotic arm that stops once it reaches the targeted area, according to the surgeon's preoperative planning [1, 18].

However, new approaches like patient-specific instrumentation, navigation, standard cutting guides, accelerometer-based navigation, and robotic guidance are expensive [5]. They also add surgical time [2], can have substantial learning curves, and have been associated with unique (and sometimes serious) complications [2]. In addition, several randomized trials have suggested that they may not improve alignment [49] (or not improve it enough to justify the added time or cost associated with their use). Long-term studies, when they have been available, have found no survivorship advantages to these techniques, including clinical function. However, these approaches have intuitive appeal, and some other studies have indeed demonstrated alignment advantages [21, 2]. The challenge when synthesizing an evidence base consisting of dozens of randomized trials is that not all of the techniques of interest have been compared head-to-head in individual trials. Moreover, many head-to-head randomized controlled trials (RCTs) or two-by-two meta-analyses have been conducted to compare these different cutting guides, but they have shown contradictory results [9, 16, 20, 22, 29, 30, 33, 35, 38, 44]. No study that we know of has simultaneously compared the different available techniques. In this setting, network meta-analysis is potentially useful [24].

This approach allows the comparison of approaches that were not evaluated directly in individual trials. That is, if approach A was compared with approach B in several randomized trials, and approach B was compared with

approach C in others, a network meta-analysis allows the comparison of approach A to approach C; as such, it is a powerful means to answer important clinical questions.

We therefore performed a network meta-analysis to determine whether novel approaches to achieving implant alignment, such as patient-specific instrumentation, navigation, accelerometer-based navigation, and robotic guidance, provide any advantage over standard cutting guides in terms of: (1) hip-knee-ankle (HKA) alignment outliers greater than $\pm 3^\circ$ (a more reasonable cutoff of 5° would be preferable to a cutoff of 3° but the 5° cutoff was too rarely reported), (2) outcome scores (1989 Knee Society Score and WOMAC score) measured 6 months after surgery, or (3) femoral and tibial implant malalignment (greater than $\pm 3^\circ$), taken separately in the frontal and sagittal plane, as well as other secondary outcomes including validated outcome scores 1 and 2 years after surgery.

Materials and Methods

This systematic review and network meta-analysis was registered in PROSPERO (Number CRD42018091913) and is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

Search Strategy

We searched electronic databases (Medline, CENTRAL and Embase) starting at the inception date up to October 1, 2018 using dedicated search equations adapted to each database (see Appendix 1; Supplemental Digital Content 1, <http://links.lww.com/CORR/A362>). We manually searched three major orthopaedic journals (the *Journal of Bone and Joint Surgery American Volume*, the *Journal of Arthroplasty*, and *Clinical Orthopaedics and Related Research*[®]) and conference proceedings of orthopaedic meetings (the American Academy of Orthopaedic Surgeons, the American Association of Hip and Knee Surgeons, and the Société Française de Chirurgie Orthopédique et Traumatologie) from January 1, 2005 to October 1, 2018. We also searched ClinicalTrials.gov and the World Health Organization's International Clinical Trials Registry Platform. Finally, we systematically screened reference lists of systematic reviews and meta-analyses for any additional references.

Study Eligibility and Selection Process

All RCTs that compared at least two types of cutting guides and reported at least one of our predefined outcomes were

included. No restriction on language or date of publication was applied. Exclusion criteria were TKA for an indication other than osteoarthritis, TKA after knee osteotomy or unicompartmental knee arthroplasty on the same knee, and TKA revision.

Two main types of outcomes were assessed in this review: the proportion of HKA outliers, defined as an HKA angle outside the 177° to 183° range within the first 6 months postoperatively and functional outcomes at 6 months postoperatively, evaluated with either the 1989 Knee Society Score (KSS) or the WOMAC score (see Appendix 2; Supplemental Digital Content 2, <http://links.lww.com/CORR/A363>).

The secondary outcomes were malalignment of tibial and femoral implants in the frontal, sagittal, and axial planes, defined as an angle more or less than 3° from the intended value within the first 6 months postoperatively (see Appendix 2; Supplemental Digital Content 2, <http://links.lww.com/CORR/A363>); functional outcomes (1989 KSS and WOMAC score) at 1 and 2 years postoperatively; the mean operative time, in minutes, from skin incision to closure; and the incidence of postoperative complications.

Two reviewers (PAB, SC) independently screened titles, summaries, and full texts whenever necessary to assess the eligibility of each study. Disagreements were resolved by a third reviewer (MRR) and by contacting the corresponding author when doubt persisted.

Selection and General Characteristics

Our search initially retrieved 1364 records. Finally, 90 RCTs with 9389 patients and 10,336 knees were included (Fig. 1). The most frequent comparison was navigation versus standard cutting guides in 58% (52 of 90) of the trials. Standard cutting guides were compared with every other type of cutting guide. Three studies directly compared patient-specific instrumentation and navigation (Fig. 2). One trial had three arms (navigation, patient-specific instrumentation, and standard cutting guides) (see Appendix 3; Supplemental Digital Content 3, <http://links.lww.com/CORR/A364>). The risk of bias of each study remains heterogeneous notably concerning the “sequence generation” and “allocation concealment” (see Appendix 4; Supplemental Digital Content 4, <http://links.lww.com/CORR/A365>).

Data Extraction and Risk of Bias Assessment

The same reviewers (PAB, SC) independently extracted the data and evaluated the risk of bias using the Cochrane risk-of-bias tool [13] in a dedicated data extraction form

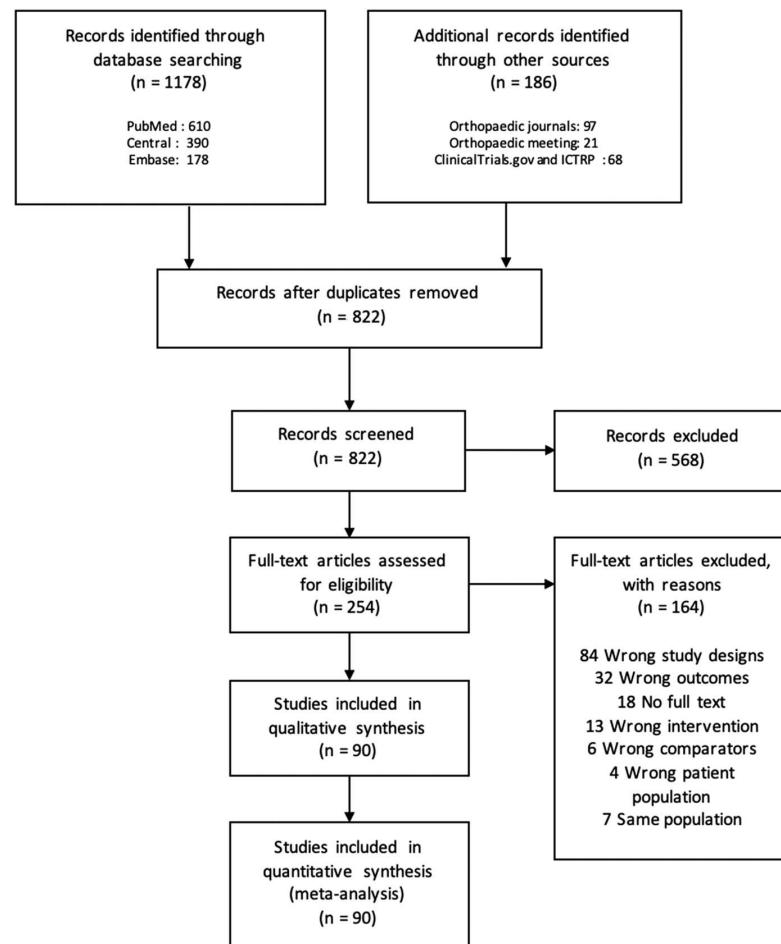


Fig. 1 This PRISMA flow chart shows the selection process of the network meta-analysis comparing the efficacy of different cutting guides in TKA; ICTRP = International Clinical Trials Registry Platform.

that was tested on 10 random records. When there was disagreement, a third reviewer (MRR) was involved to reach consensus.

We collected the following data: general study information: title, date of publication, authors, country, contact details of the corresponding author, and funding sources; characteristics of the study: design, inclusion and exclusion criteria, and follow-up duration; population characteristics: mean age, proportion of men, mean or median values for size, weight, and BMI; cutting guide used: standard cutting guides, navigation, patient-specific instrumentation, accelerometer-based navigation, or robotic guidance; prosthesis characteristics: manufacturer, product name, type of tibial bearing, type of fixation, and

conservation of the posterior cruciate ligament with a posterior-stabilized system; surgical characteristics: unilateral or bilateral simultaneous arthroplasty, use of a tourniquet, surgical approach, or patellar resurfacing; and results for outcomes of interest: frequencies of binary outcomes, mean and SD for quantitative variables, and the number of patients included in each group. Regarding continuous outcomes, when the mean or SD was not available, we estimated them from the median and the first and third quartiles and/or ranges using approximation formulas [14, 48] (see Appendix 5; Supplemental Digital Content 5, <http://links.lww.com/CORR/A366>). The corresponding authors were contacted if there were missing data.

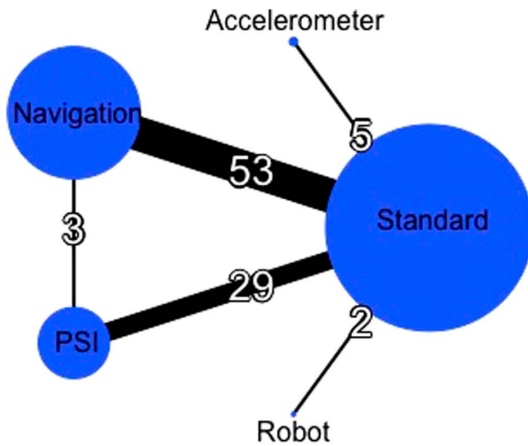


Fig. 2 In this global network diagram, the nodes (circles) represent different cutting blocks, and their sizes were proportional to the sample size of each respective intervention. The edges (lines) indicate direct comparisons, and their thicknesses were proportional to the number of studies available.

Patient Involvement

Patients were not involved in any aspect of the study design, conduct, or development of the research question or outcome measures. This study is a systematic review and network meta-analysis of existing published research; therefore, there was no active patient recruitment for data collection.

Data Synthesis and Analysis

The unit of analysis was the knee. Binary outcomes were compared between cutting guides using risk ratios, and continuous outcomes were compared using mean differences with 95% credible intervals (CIs). For direct comparisons, we conducted a conventional meta-analysis to synthesize the results, using random-effects models and fixed-effects models as sensitivity analyses. We combined direct and indirect comparisons via a network meta-analysis using the hierarchical model of Lu and Ades [28] with a Bayesian approach. In a Bayesian framework, estimates are based on the posterior distribution of the endpoint and are called credible intervals (Cis). Usually, the 95% CI corresponding to the interval has a posterior probability of 95% of the endpoint that lies within it. Unlike the frequentist approach, the Bayesian approach does not allow the calculation of the p value. Conclusions are drawn from whether or not 1 (which corresponds to a no-difference relative risk) or 0 (which corresponds to no mean differences) belongs to the credible interval. We used Markov Chain Monte Carlo to implement the model. Binary non-informative prior distribution was used for dichotomous outcomes and normal non-informative prior distribution was used for continuous outcomes [46]. The model’s convergence was assessed using Brooks-Gelman-Rubin plots [4]. We evaluated the fit via residual deviance [8]. We checked the assumption of transitivity by considering similarity in the distribution of age, sex, and BMI across the different pairwise comparisons. A network diagram was used for each outcome to present direct

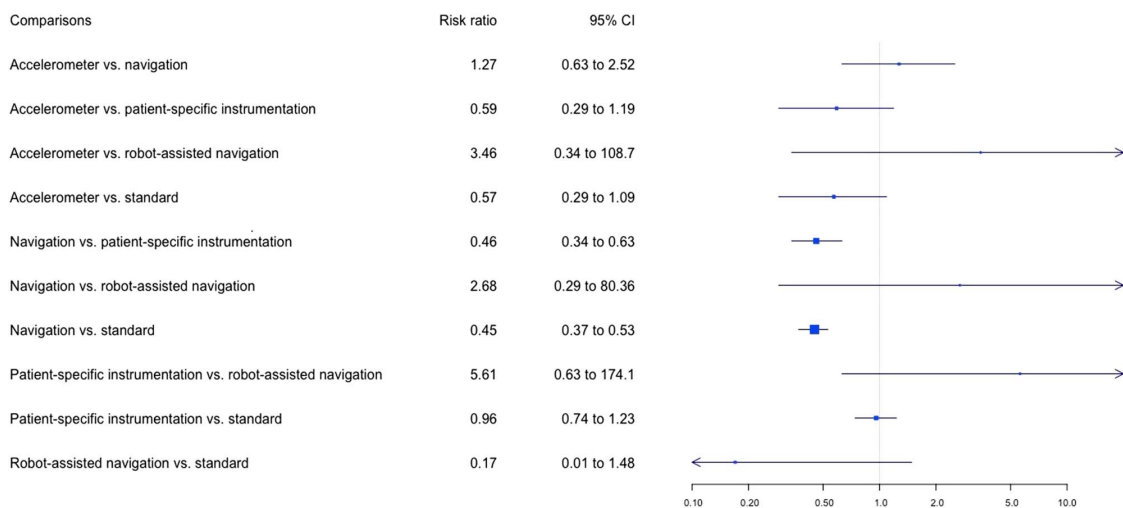


Fig. 3 Forest plots in the network meta-analysis of the relative risks of HKA misalignments between the different cutting guides are shown. The data are presented as risk ratios with 95% credibility intervals, which were obtained using random effects models. The I² of the network meta-analysis was 19%.

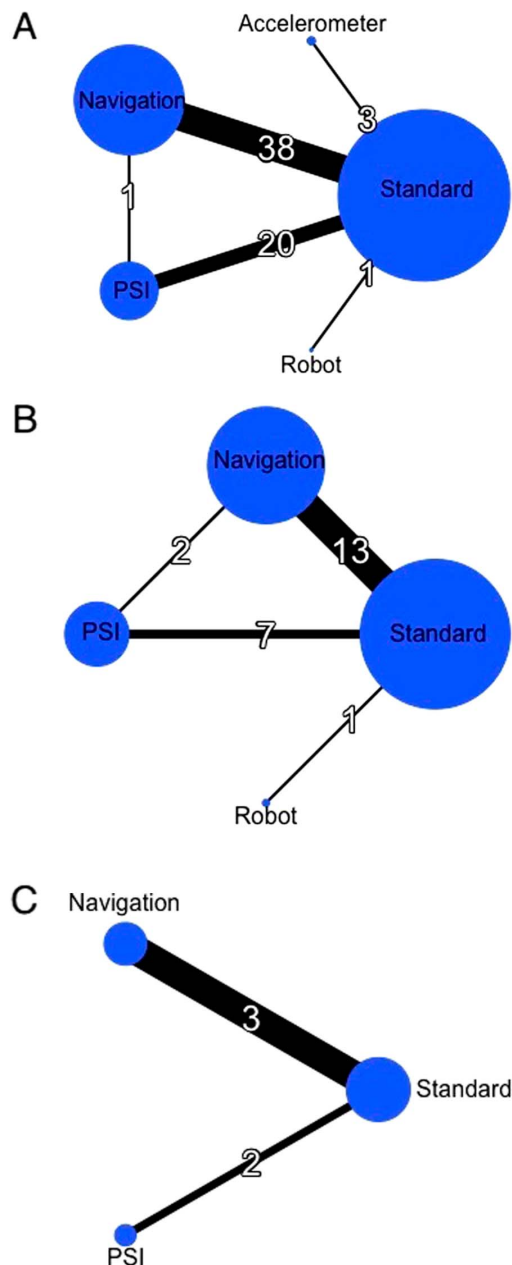


Fig. 4 A-C These network diagrams of the main outcomes show (A) the HKA malalignment, (B) KSS at the first 6 months, and (C) WOMAC scores at the first 6 months.

comparisons between interventions. We obtained a treatment hierarchy using the surface under the cumulative ranking curve and mean ranks. To evaluate the assumption of inconsistency, we used a node-splitting approach [7],

and we compared the fit between the consistency and inconsistency models. We statistically assessed the presence of heterogeneity within each network comparison using the I^2 statistic. Subgroup analyses were planned for the main outcomes according to the surgical approach and risk of bias. We conducted sensitivity analyses for the main results using network meta-regression analyses adjusted successively for the mean age, mean BMI, percentage of women, and number of TKAs per study. A comparison-adjusted funnel plot of older treatment versus newer treatment was used to assess small study effects within each pairwise comparison. Cutting guides were ranked in order from oldest to newest: standard cutting guides, navigation, patient-specific instrumentation, accelerometer-based navigation, and robotic guidance. We used the GEMTC R-package (version 0.8-2) for analyses and netmeta R-package (version 0.9-8) for network diagrams in R version 3.5.2 software (R foundation for statistical Computing, Vienna, Austria). The quality of evidence for the main outcomes was evaluated using the GRADE framework [40].

Results

HKA Angle Outliers Greater Than $\pm 3^\circ$

The proportion of HKA angle outliers greater than $\pm 3^\circ$ was lower in the navigation group than in the patient-specific instrumentation group (12% versus 21%, risk ratio 0.46 [95% CI 0.34 to 0.63]) and in the standard cutting guide group (12% versus 25%, risk ratio 0.45 [95% CI 0.37 to 0.53]), with $I^2 = 19\%$ (Fig. 3). Sixty-eight percent (61 of 90) of studies reported malalignment of the HKA angle. TKA using standard cutting guides was the most used control procedure (Fig. 4). The results of pairwise comparisons confirmed that navigation against the standard cutting guide reduced the risk of malalignment of the HKA angle (RR 0.43 [95% CI 0.36 to 0.53]; $I^2 = 51.2\%$) (see Appendix 6; Supplemental Digital Content 6, <http://links.lww.com/CORR/A367>). However, this corresponded to a relatively small proportion of the patients being influenced by the difference in accuracy. Robot-assisted surgery had the highest probability of having the lowest proportion of HKA angle outliers greater than $\pm 3^\circ$ but without any significant difference (Fig. 5).

Functional Outcomes

We found no differences between the different cutting guides either on KSS or WOMAC at 6 months.

Concerning the KSS at 6 months, the results of the pairwise comparison confirmed those results and did not

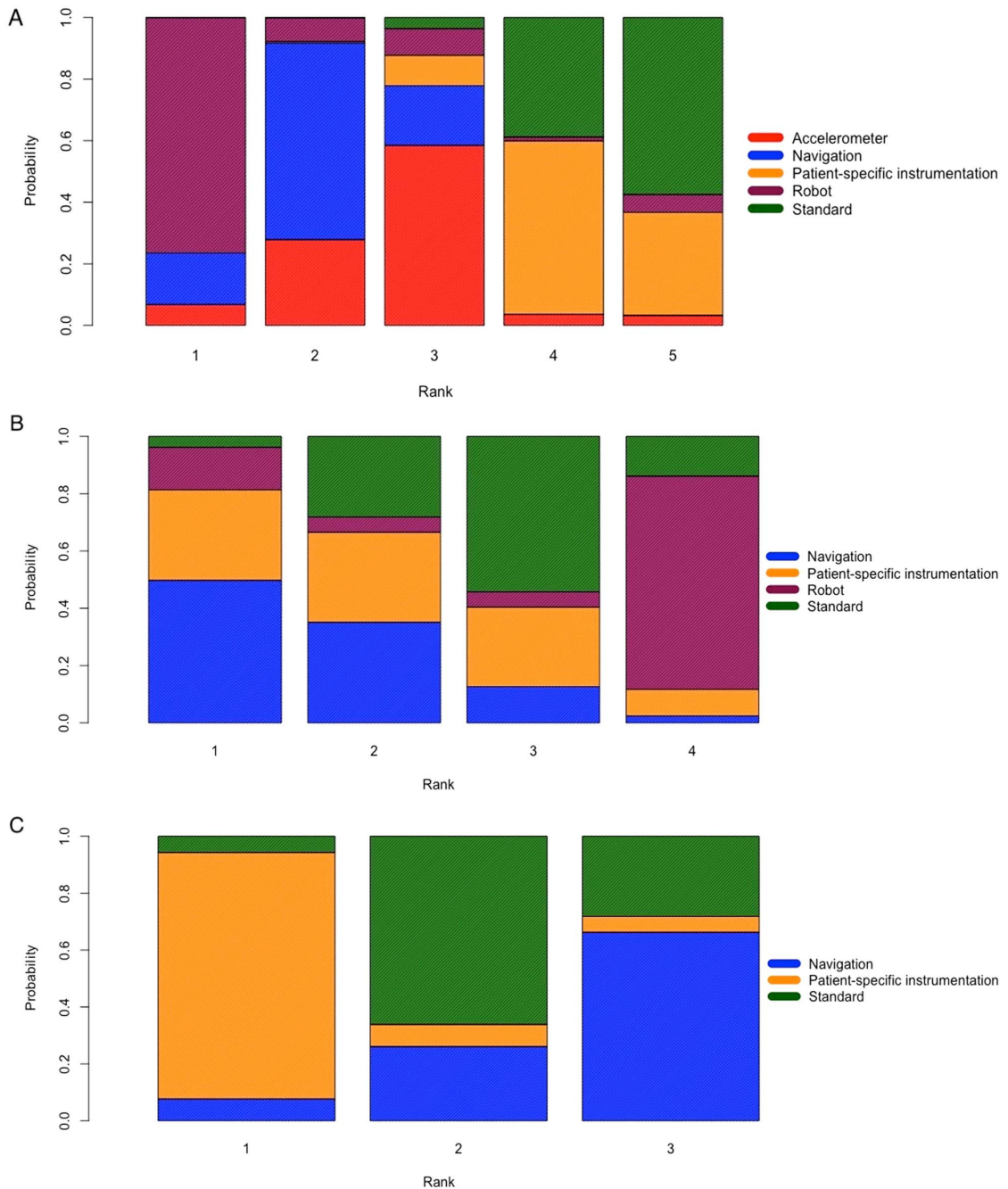


Fig. 5 A-C These images show rank probabilities of (A) the HKA malalignment angle, (B) delta KSS at the first 6 months postoperatively, and (C) delta WOMAC scores at the first 6 months postoperatively for each cutting guide. The x axis represents the ranking. The y axis represents the probability of rank: 0% to 100%. Each color represents one cutting guide: red for accelerometer-based navigation, blue for navigation, yellow for patient-specific instrumentation, purple for robot navigation, and green for the standard cutting guide; PSI = patient-specific instrumentation.

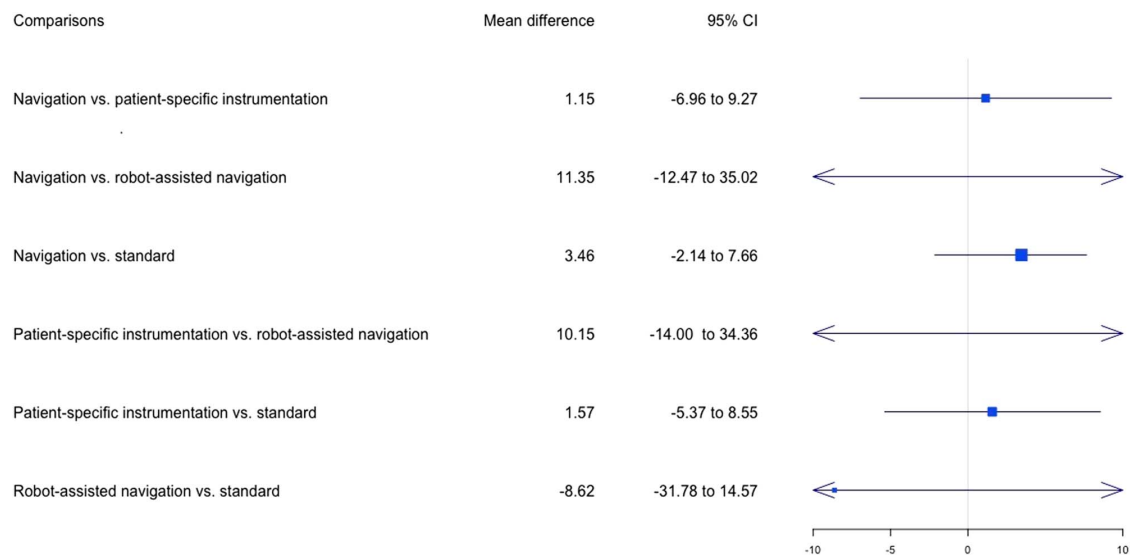


Fig. 6 Forest plots in the network meta-analysis of the mean difference of KSS between preoperatively and 6 months postoperatively are shown. The data are presented as the relative mean differences with 95% credibility intervals, which were obtained using random effects models. The I^2 of the network meta-analysis was 3%.

indicate any differences in functional outcomes (see Appendix 6; Supplemental Digital Content 6, <http://links.lww.com/CORR/A367>). Twenty-one studies (23%) reported the KSS at 6 months postoperatively (Fig. 4). Navigation had the highest probability of leading to the highest KSS mean values (Fig. 5), although no differences were found between the different cutting guides (Fig. 6).

Concerning the WOMAC at 6 months, the results of the pairwise comparison did not indicate any differences (see Appendix 6; Supplemental Digital Content 6, <http://links.lww.com/CORR/A367>). WOMAC scores at 6 months postoperatively were available in five studies (5%) (Fig. 4). Patient-specific instrumentation had the highest probability of leading to the highest WOMAC mean values (Fig. 5), although no differences were found between the different cutting guides (Fig. 7).

Secondary Outcomes

In general, all tested alternatives provided some improvement in overall alignment. Concerning frontal femoral outliers $\pm 3^\circ$, navigation reduced their proportion compared with standard cutting guides (5% versus 14%; RR 0.46 [95% CI 0.31 to 0.69]). The proportion of outliers for frontal femoral component alignment was lower in the accelerometer group than in the patient-specific instrumentation group (1% versus 11%; RR 0.23 [95% CI

0.07 to 0.69]) and in the standard cutting guide group (1% versus 14%; RR 0.18 [95% CI 0.06 to 0.49]). Concerning frontal tibial outliers $\pm 3^\circ$, navigation reduced their proportion compared with standard cutting guides (5% versus 10%; RR 0.55 [95% CI 0.36 to 0.81]) and compared with the patient-specific instrumentation group (5% versus 10%; RR 0.40 [95% CI 0.21 to 0.74]).

Regarding femoral sagittal outliers $\pm 3^\circ$, navigation reduced their proportion compared with standard cutting guides (14% versus 30%; RR 0.54 [95% CI 0.38 to 0.74]) and compared with patient-specific instrumentation (14% versus 35%; RR 0.49 [95% CI 0.30 to 0.76]). As for tibial sagittal outliers $\pm 3^\circ$, navigation reduced their proportion compared with standard cutting guides (15% versus 24%; RR 0.56 [95% CI 0.37 to 0.84]) and compared with patient-specific instrumentation (15% versus 30%; RR 0.34 [95% CI 0.18 to 0.59]). Accelerometer and standard cutting guides reduced the proportion of outliers against patient-specific instrumentation (2% versus 30%; RR 0.35 [95% CI 0.11 to 0.99] and 24% versus 30%; RR 0.68 [95% CI 0.56 to 0.89]). We found no differences between groups regarding the rotation of femoral and tibial implants (see Appendix 7; Supplemental Digital Content 7, <http://links.lww.com/CORR/A368>). For robotic guidance in sagittal, frontal and axial alignment, we did not find any direct comparison in the evidence.

As for secondary clinical outcomes, we observed no differences in the KSS and WOMAC scores at 1 and 2

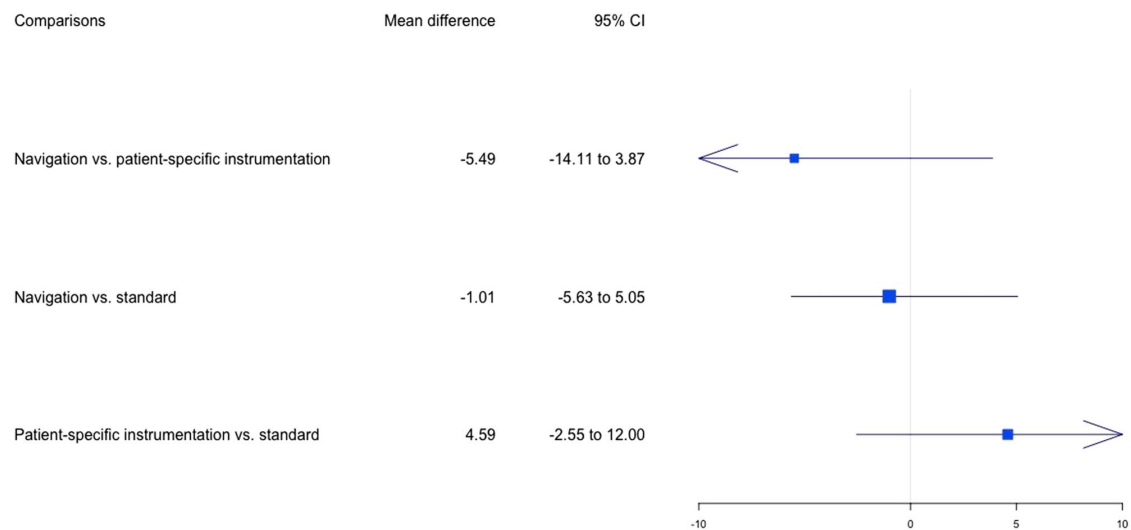


Fig. 7 Forest plots in the network meta-analysis of the mean difference of WOMAC scores between preoperatively and 6 months postoperatively are shown. The data are presented as the relative mean differences with 95% credibility intervals, which were obtained using random effects models. The I^2 of the network meta-analysis was 1%.

years postoperatively between the different cutting guides (see Appendix 8; Supplemental Digital Content 8, <http://links.lww.com/CORR/A369>).

Surgical procedures using navigation were longer than those using patient-specific instrumentation (95 minutes versus 70 minutes, mean difference 20.19 [95% CI 14.08 to 26.28] and standard cutting guides (95 minutes versus 78 minutes, mean difference 17.03 [95% CI 13.03 to 21.06]) (see Appendix 8; Supplemental Digital Content 8, <http://links.lww.com/CORR/A369>). We did not find any difference concerning accelerometer and robotic guidance in surgical time. Direct comparisons did not suggest any differences in postoperative complications (see Appendix 9; Supplemental Digital Content 9, <http://links.lww.com/CORR/A370>).

Grade Evaluation

We found no differences regarding distribution of age, sex, number of participants and BMI in each comparison of the network. An adequate random sequence generation and an adequate treatment allocation concealment were specified in 62% (56 of 90) and 50% trials (45 of 90) respectively. Outcome assessors and patients were blinded in 54% (49 of 90) and in 20% studies (18 of 90) respectively. Overall, for the items incomplete outcome data regarding radiological, clinical, and perioperative outcomes, risk was estimated low. In 60% (54 of 90) of studies, the lead final analysis was intention to treat. Finally, heterogeneity for different

outcomes was low: 19% for HKA outliers, 3% for KSS at 6 months, and 1% for WOMAC. Overall, studies were well conducted and data for our extraction was available.

We performed the GRADE evaluation for the two main outcomes. For HKA outliers $\pm 3^\circ$, it suggests that the evidence is moderate for navigation/ patient-specific instrumentation, navigation/standard, patient-specific instrumentation /standard comparison and accelerometer, and very low for robot/standard comparison (see Appendix 10; Supplemental Digital Content 10, <http://links.lww.com/CORR/A371>).

For KSS at the 6 months post-surgery, it suggests that the evidence is moderate for

navigation/ patient-specific instrumentation, low for navigation/standard and patient-specific instrumentation /standard comparison, and very low for robot/standard comparison (see Appendix 10; Supplemental Digital Content 10, <http://links.lww.com/CORR/A371>). For WOMAC at the 6 months post-surgery, it suggests that the evidence is high for navigation/standard and moderate for patient-specific instrumentation/standard comparison (see Appendix 10; Supplemental Digital Content 10, <http://links.lww.com/CORR/A371>).

Discussion

Although some studies have suggested that small differences in alignment matter in terms of TKA implant survivorship [3, 19] and knee function [39], others have not [36,

47]. In the interest of making alignment as consistent as possible, surgeons therefore have explored a variety of approaches, including navigation, accelerometers, patient-specific instrumentation, and robotics. It is important to determine the degree to which these approaches achieve those goals without losing sight of the fact that these approaches also have potential disadvantages, including increased cost, increased surgical time, a difficult learning curve, and unique—and sometimes severe—complications [2, 5]. A network meta-analysis is a useful approach to synthesize a large amount of data in this setting because this approach allows the comparison of approaches that were not compared directly in individual trials. We found an incremental reduction in the risk of small ($\pm 3^\circ$) HKA outliers when navigation was used compared with the other approaches, but no associated improvements in validated outcomes scores. The small alignment benefit came with a large increase in surgical time. Until or unless future studies can demonstrate a benefit in terms of survivorship, knee pain, cost effectiveness, or function of navigation for these other newer approaches, based on our findings we recommend against their widespread use.

Limitations

However, this study had several limitations. Few trials evaluated more recent techniques, such as the use of an accelerometer [11, 17, 21, 34, 43] or robot [25, 26], making it difficult to assess the benefits of these two techniques in relation to their cost. Unfortunately, we were unable to collect clinical outcomes after 2 years postoperatively due to the lack of information in published studies, even though joint prostheses are expected to have a longevity of 15 years or more. Finally, the ranking results were based on point estimate probabilities, and the results do not fully reflect the uncertainty of the ranking as outlined by Trinquart et al. [45]. As a consequence, we must consider ranking results as trends rather than definitive conclusions.

We found that many trials did not evaluate clinical outcomes; only 24 trials reported clinical outcomes. Only three trials simultaneously evaluated our three primary outcomes, illustrating the high level of heterogeneity regarding the outcomes planned and reported. In addition, the reporting of clinical outcomes was notably heterogeneous across trials. The WOMAC was reported with different numeric scales, whereas the KSS was more homogeneously used and reported in the included RCTs [27]. Such heterogeneity in evaluated outcomes and their definitions makes it difficult to compare results between trials and combinations by meta-analysis. This highlights the need to develop recommendations and core sets of outcomes as promoted by the Core Outcome Measures in Effectiveness Trials (COMET) initiative [10] in the field of knee prosthetic replacement.

Small Improvement in HKA Alignment

We found that about 10% more patients could achieve an HKA of within 3° of neutral when navigation was used compared with when standard guides or patient-specific instruments were used. We caution readers against making too much of even this incremental difference. The evidence is scant and contradictory that such small alignment differences are likely to improve implant survivorship or function [9, 20, 31] and even larger differences—such as $\pm 5^\circ$ —may not matter as much as once was thought [36]. We would have preferred to set the cutoff to 5° or even analyze it as a continuous variable, but the data were not available in the studies we evaluated. Reducing the likelihood of HKA outliers in a small minority of patients undergoing TKA by widespread adoption of navigation would result in substantial increases in costs and operative time [2], at least initially.

No Improvements in Outcomes Scores

We found no evidence that these technologies improved KSS or WOMAC scores at either 6 months, 1 year, or 2 years after TKA. For clinical outcomes, similar results were found in other studies regarding KSS comparing navigation, patient-specific instrumentation, and standard cutting guides [28, 37]. Petursson et al. [37] found that navigation had a better WOMAC after 2 years. We did not find similar results at short or medium terms.

No Increase in Complications but Longer Surgical Time

Navigation is often criticized because of increased operative time and the risk of iatrogenic fracture due to the pin insertions. In this network meta-analysis, navigation increased the mean operative time by 20 minutes compared with standard cutting guides and patient-specific instrumentation. However, postoperative complications were similar among all cutting guides. We were probably underpowered to detect rare but serious complications. Thus, in our view, the benefit/risk balance between a more precise tool, such as navigation, allowing more reproducible cuts but without evidence of an impact on function and an increase of surgical time does not favor the navigation approach.

Conclusions

No new approach in this study demonstrated superiority over standard cutting guides on functional outcome, and

the alignment differences affected few patients, were small in magnitude, and were of questionable clinical importance. Thus, unless these new approaches either convincingly demonstrate superior survivorship or convincingly demonstrate superior long-term functional outcomes, there is no reason to use these approaches from either the patient's point of view or the hospital's point of view since they add cost, have a learning curve (during which some patients may be harmed), and have the risks associated with novelty.

Acknowledgments We thank Professor Sylvie Chevret MD, PhD, director of the biostatistics department in the Saint Louis Hospital, for her advice. Furthermore, we thank every corresponding author who made the effort to provide us with the requested supplementary data.

References

- Banerjee S, Cherian JJ, Elmallah RK, Jauregui JJ, Pierce TP, Mont MA. Robotic-assisted knee arthroplasty. *Expert Rev Med Devices*. 2015;12:727–735.
- Bauwens K, Matthes G, Wich M, Gebhard F, Hanson B, Ekkernkamp A, Stengel D. Navigated total knee replacement. A meta-analysis. *J Bone Joint Surg Am*. 2007;89:261–269.
- Berger RA, Crossett LS, Jacobs JJ, Rubash HE. Malrotation causing patellofemoral complications after total knee arthroplasty. *Clin Orthop Relat Res*. 1998; 356:144–153.
- Brooks SP, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. *J Comput Graph Stat*. 1998;7:434–355.
- Clark GW, Wood DJ. Robotics in arthroplasty: Where are we today? *Bone & Joint* 360. 2015;4:2–7.
- Deep K, Shankar S, Mahendra A. Computer assisted navigation in total knee and hip arthroplasty. *SICOT-J*. 2015;3:50.
- Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med*. 2010;29:932–944.
- Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Mak*. 2013;33:641–656.
- Fu Y, Wang M, Liu Y, Fu Q. Alignment outcomes in navigated total knee arthroplasty: a meta-analysis. *Knee Surg Sports Traumatol Arthrosc*. 2012;20:075–1082.
- Gargon E, Williamson PR, Altman DG, Blazeby JM, Tunis S, Clarke M. The COMET Initiative database: progress and activities update (2015). *Trials*. 2017; 18:54.
- Gharabeh MA, Solayar GN, Harris IA, Chen DB, MacDessi SJ. Accelerometer-based, portable navigation (KneeAlign) vs conventional instrumentation for total knee arthroplasty: a prospective randomized comparative trial. *J Arthroplasty*. 2017;32:777–782.
- Haglin JM, Eltorai AEM, Gil JA, Marcaccio SE, Botero-Hincapie J, Daniels AH. Patient-specific orthopaedic implants. *Orthop Surg*. 2016;8:417–424.
- Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savović J, Schulz KF, Weeks L, Sterne JAC, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;18:343:d5928.
- Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2015;5:13
- Huang EH, Copp SN, Bugbee WD. Accuracy of a handheld accelerometer-based navigation system for femoral and tibial resection in total knee arthroplasty. *J Arthroplasty*. 2015;30:1906–1910.
- Huijbregts HJTAM, Khan RJK, Sorensen E, Fick DP, Haebich S. Patient-specific instrumentation does not improve radiographic alignment or clinical outcomes after total knee arthroplasty. *Acta Orthop*. 2016;87:386–394.
- Ikawa T, Takemura S, Kim M, Takaoka K, Minoda Y, Kadoya Y. Usefulness of an accelerometer-based portable navigation system in total knee arthroplasty. *Bone Joint J*. 2017;99:1047–1052.
- Jacofsky DJ, Allen M. Robotics in arthroplasty: A Comprehensive Review. *J Arthroplasty*. 2016;31:2353–2363.
- Jeffery RS, Morris RW, Denham RA. Coronal alignment after total knee replacement. *J Bone Joint Surg Br*. 1991;73:709–714.
- Kim YH, Kim JS, Yoon SH. Alignment and orientation of the components in total knee replacement with and without navigation support: a prospective, randomised study. *J Bone Joint Surg*. 2007;89:471–476.
- Kinney MC, Cidambi KR, Severns DL, Gonzales FB. Comparison of the iAssist handheld guidance system to conventional instruments for mechanical axis restoration in total knee arthroplasty. *J Arthroplasty*. 2018;33:61–66.
- Kuzyk PRT, Higgins GA, Tungal JAW, Sellan ME, Waddell JP, Schemitsch EH. Computer navigation vs extramedullary guide for sagittal alignment of tibial components: radiographic study and meta-analysis. *J Arthroplasty*. 2012;27:630–637.
- Laurent P. Patient-specific instruments in orthopaedics. In: Ritacco LE, Milano FE, Chao E, eds. *Computer-assisted Musculoskeletal Surgery*. Philadelphia, PA: Springer; 2016: 163–179.
- Leopold SS. Pencil and paper research? Network meta-analysis and other study designs that do not enroll patients. *Clin Orthop Relat Res*. 2015;473:2163–2165.
- Liow MHL, Goh GSH, Wong MK, Chin PL, Tay DKJ, Yeo SJ. Robotic-assisted total knee arthroplasty may lead to improvement in quality-of-life measures: a 2-year follow-up of a prospective randomized trial. *Knee Surg Sports Traumatol Arthrosc Off*. 2017;29:2942–2951.
- Liow MHL, Xia Z, Wong MK, Tay KJ, Yeo SJ, Chin PL. Robot-assisted total knee arthroplasty accurately restores the joint line and mechanical axis: A prospective randomised study. *J Arthroplasty*. 2014;29:2373–2377.
- Lovelock TM, Broughton NS, Williams CM. The popularity of outcome measures for hip and knee arthroplasties. *J Arthroplasty*. 2018;33:273–276.
- Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. 2004;23:3105–3124.
- Mannan A, Akinyooye D, Hossain F. A Meta-analysis of functional outcomes in patient-specific instrumented knee arthroplasty. *J Knee Surg*. 2017;30:668–674.
- Mannan A, Smith TO. Favourable rotational alignment outcomes in PSI knee arthroplasty: A Level 1 systematic review and meta-analysis. *The Knee*. 2016;23:186–190.
- Mason JB, Fehring TK, Estok R, Banel D, Fahrback K. Meta-analysis of alignment outcomes in computer-assisted total knee arthroplasty surgery. *J Arthroplasty*. 2007;22:1097–1106.
- Mattei L, Pellegrino P, Calò M, Bistolfi A, Castoldi F. Patient specific instrumentation in total knee arthroplasty: a state of the art. *Ann Transl Med*. 2016;4:126–126.
- Moskal J, Capps S, Mann J, Scanelli J. Navigated versus conventional total knee arthroplasty. *J Knee Surg*. 2013;27: 235–248.

34. Nam D, Cody EA, Nguyen JT, Figgie MP, Mayman DJ. Extramedullary guides versus portable, accelerometer-based navigation for tibial alignment in total knee arthroplasty: A randomized, controlled trial: Winner of the 2013 HAP PAUL award. *J Arthroplasty*. 2014;29:288–294.
35. Parratte S, Blanc G, Boussemart T, Ollivier M, Corroller TL, Argenson JN. Rotation in total knee arthroplasty: no difference between patient-specific and conventional instrumentation. *Knee Surg Sports Traumatol Arthrosc*. 2013;21:2213–2219.
36. Parratte S, Pagnano MW, Trousdale RT, Berry DJ. Effect of postoperative mechanical axis alignment on the fifteen-year survival of modern, cemented total knee replacements. *J Bone Joint Surg Am*. 2010;92:2143–2149.
37. Petursson G, Fenstad AM, Gothesen O, Dyrhovden GS, Hallan G, Rohrl SM, Aamodt A, Furnes O. Computer-assisted compared with conventional total knee replacement: a multicenter parallel-group randomized controlled trial. *J Bone Joint Surg Am*. 2018;100:1265–1274.
38. Rebal BA, Babatunde OM, Lee JH, Geller JA, Patrick DAJ, Macaulay W. Imageless computer navigation in total knee arthroplasty provides superior short term functional outcomes: a meta-analysis. *J Arthroplasty*. 2014;29:938–944.
39. Ritter MA, Faris PM, Keating EM, Meding JB. Postoperative alignment of total knee replacement. Its effect on survival. *Clin Orthop Relat Res*. 1994;299:153–156.
40. Salanti G, Giovane CD, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the Quality of Evidence from a Network Meta-Analysis. *PLoS One*. 2014;9:e99682.
41. Sikorski JM. Alignment in total knee replacement. *J Bone Joint Surg Br*. 2009;90:1121–1127.
42. Stulberg SD, Loan P, Sarin V. Computer-assisted navigation in total knee replacement: results of an initial experience in thirty-five patients. *J Bone Joint Surg Am*. 2002;84(A Suppl 2):90–98.
43. Thiengwittayaporn S, Fusaku Y, Kangkano N, Jarupongprapa C, Charoenphandhu N. Hand-held navigation may improve accuracy in minimally invasive total knee arthroplasty: a prospective randomized controlled trial. *Int Orthop*. 2016;40:51–57.
44. Thienpont E, Schwab PE, Fennema P. Efficacy of patient-specific instruments in total knee arthroplasty: a systematic review and meta-analysis. *J Bone Joint Surg Am*. 2017;99:521–530.
45. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in treatment rankings: reanalysis of network meta-analyses of randomized trials. *Ann Intern Med*. 2016;164:666–673.
46. Van Valkenhoef and Joel Kuiper. gemtc: Network Meta-Analysis Using Bayesian Methods. *R package version 0.8-2*. <https://CRAN.R-project.org/package=gemtc>. Accessed March 20, 2020.
47. Vanlommel L, Vanlommel J, Claes S, Bellemans J. Slight undercorrection following total knee arthroplasty results in superior clinical outcomes in varus knees. *Knee Surg Sports Traumatol Arthrosc*. 2013;21:2325–2330.
48. Weir CJ, Butcher I, Assi V, Lewis SC, Murray GD, Langhorne P, Brady MC. Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review. *BMC Med Res Methodol*. 2018;18:25.
49. Victor J, Dujardin J, Vandenneucker H, Arnout N, Bellemans J. Patient-specific guides do not improve accuracy in total knee arthroplasty: A prospective randomized controlled trial knee. *Clin Orthop Relat Res*. 2014;472:263–271.

Conclusion

Ce travail s'est intéressé dans un premier temps à faire une méta-analyse en réseau afin d'essayer de répondre de manière claire à deux questions d'actualités sur les prothèses totale de genoux : la voie d'abord dans les prothèses totales de genoux peut-elle jouer un rôle dans la réussite de cette intervention ? Un guide de coupe est-il supérieur à un autre dans la réalisation des PTG ?

Concernant les voies d'abords médiale dans les PTG, l'hypothèse principale était qu'elle pouvait avoir une incidence sur l'enveloppe ligamentaire, le positionnement des guides de coupes et sur la récupération fonctionnelle (Goble and Justin, 2004; Vail, 2004) . Nous avons retrouvé que la voie subvastus améliorait les mobilités dans les six premiers mois postopératoires sans différence sur les scores fonctionnels. La voie parapatellaire interne diminuait le temps opératoire comparée aux autres voies d'abord chirurgicales. La durée d'hospitalisation était diminuée par la voie mini-parapatellaire interne comparée aux autres techniques. Aucune voie d'abord chirurgicale médiale pour les prothèses totales de genoux ne semblent clairement supérieure à une autre. Au vu des résultats de l'étude et d'un nombre non négligeable d'essais randomisés traitant de ce sujet, nous pensons avoir répondu à cette problématique. Notre conclusion est que toutes les voies d'abords chirurgicales médiales du genou pour réaliser une PTG sont équivalentes et le choix revient au chirurgien en fonction de ses habitudes, de ses préférences et de sa maîtrise technique. Il serait intéressant de comparer à expérience égale toutes ses voies d'abords, notamment sur le temps chirurgical, le taux de complications et la récupération fonctionnelle en postopératoire.

Concernant les guides de coupes sur mesure dans les PTG, l'hypothèse principale était qu'ils peuvent jouer un rôle dans la précision des coupes osseuses et par conséquent améliorer la fonction en postopératoire. En effet, un guide de coupe qui permet à un chirurgien d'être plus précis et reproductible pourrait permettre de s'approcher de l'objectif fixé par le chirurgien. Nous avons comparé 5 guides de coupes l'un à l'autre. Nous n'avons que très peu d'essais randomisés qui comparaient l'accéléromètre et la robotique. La navigation

permettait de diminuer le taux de valeurs aberrantes dans le plan frontal et dans le plan sagittal comparée aux guides de coupes standard mais augmentait d'environ 20 minutes le temps opératoires. Au vu de ces résultats, nous ne pouvons pas recommander d'utiliser un guide de coupe plutôt qu'un autre car les différences observées sur les critères radiologiques n'ont pas eu de répercussions sur les scores cliniques. Il est intéressant, de voir dans quelques années les premiers résultats de la robotique. Il nous semble que la robotique apportera une précision que les autres guides de coupes ne permettent pas. Le manque de résultat plus concret peut s'expliquer. En effet, il existe peu d'essais randomisés dans la littérature qui collectent des critères de jugement radiologique. Au maximum seulement 61 études sur 90 relevaient un de ces critères (68 %). De plus, la plupart des mesures des critères radiologiques, dans les études sélectionnées, ont été effectués sur des radiographies. La fiabilité des mesures radiologiques est variable et moins reproductible que celles effectuées sur une imagerie en coupe (De Valk et al., 2016; Riddle, 2012; Specogna et al., 2007). Toutefois, une étude récente comparant les résultats du robot à un ancillaire mécanique, avec un recul minimum de 10 ans, ne retrouvait pas de différence (Kim et al., 2020). Des données supplémentaires sont donc nécessaires afin de prouver réellement l'intérêt ou non de la robotique dans la chirurgie prothétique.

Concernant les critères de jugement clinique, pour les deux méta-analyses, nous avons choisi de faire la différence des scores en préopératoire par rapport au score en postopératoire à différent moment du suivi. Nous avons opté pour cette option afin de diminuer l'hétérogénéité entre les différents centres, entre les différentes techniques et d'organisation des systèmes de santé. Aucune différence n'a pu être mise en évidence au regard des scores cliniques. Plusieurs explications sont possibles. Premièrement, il existe une multitude de score clinique et aucune recommandation sur quel score utiliser dans une pathologie donnée. Nous avons collecté trois scores différents qui au final n'apportent que très peu d'information différente.

De plus, les scores ont été collectés en postopératoire à des moments différents. Il n'existe pas une homogénéité sur le suivi des patients. Il paraît cependant difficile pour une équipe chirurgicale de revoir régulièrement et fréquemment tous les patients opérés sans surcharger ses consultations et en continuant de consulter de nouveaux patients (Piriou et al., 2006).

La multiplicité des scores cliniques et des moments de recueils, fait qu'il existe une énorme perte d'informations. Le score qui comporte le plus d'informations est le total KSS dans les 6 premiers mois postopératoire. Ce score est collecté dans 22 études (37%) pour la méta-analyse sur les voies d'abord chirurgicales et dans 21 études (23%) pour celle sur les guides de coupes.

Cela met en évidence la nécessité obligatoire d'homogénéiser et d'élaborer des recommandations sur les critères de jugement à collecter dans les études en orthopédie et spécialement dans la chirurgie prothétique comme le recommande le groupe Core Outcome Measures in Effectiveness Trials (COMET) (Gargon et al., 2017).

Deuxièmement, la fonction est une évaluation subjective du patient sur son état et plusieurs paramètres rentrent en ligne de compte. Résumer la future fonction d'une articulation prothétique à l'utilisation ou non d'un outil technologique nous semble réducteur. Comme nous l'avons vu, précédemment, les articles inclus dans l'étude sont issus de plusieurs continents et d'une multitude de pays différents. L'environnement préopératoire, peropératoire et postopératoire est forcément différent (Lyman et al., 2018; Putman et al., 2021). Afin d'explorer l'hétérogénéité entre les études, nous avons fait des analyses en sous-groupe et des méta-régressions pour les critères de jugement principaux dans les deux méta-analyses. Nous n'avons observé aucune différence. La validité externe est par conséquent bonne.

Enfin, il faudrait savoir à partir de quelle différence observée entre le préopératoire et le postopératoire, pour un score donné, celle-ci devient pertinente pour le patient. Pour cela, il existe des méthodes statistiques intéressantes et qui se développent de plus en plus en orthopédie que sont les MCID (minimal clinically important difference) (Copay et al., 2007). Les MCID sont utilisés en recherche clinique comme seuil de décision pour tester l'efficacité d'un nouveau traitement par rapport aux pratiques actuelles (Maredupaka et al., 2020).

En prenant en compte toutes ces considérations, il est actuellement impossible, à mon sens, de savoir si un guide de coupe ou une voie d'abord permet d'améliorer ou non la fonction des patients après une pose de prothèse totale de genou.

Nous avons choisi de rapporter l'effet des différentes interventions en utilisant des modèles à effets aléatoires. C'est un choix que nous avons fait a priori de toutes analyses. En effet, dans cette méta-analyse en réseau, nous nous attendions au moins à des différences dans les populations étudiées (comme l'âge des patients) ou sur les interventions reçues (type de prothèse, procédure pré et postopératoire, système de santé, ...). En effet, ignorer l'hétérogénéité peut conduire à un résultat de synthèse trop précis (c'est-à-dire que l'intervalle de crédibilité est trop étroit) et peut faire conclure à tort à une différence (Riley et al., 2011).

Nous avons choisi un modèle bayésien car cela, permet de renforcer l'inférence concernant l'efficacité relative de deux traitements, en incluant à la fois des comparaisons « directes » et « indirectes ». Les statistiques bayésiennes permettent aussi de mesurer l'effet du traitement sous forme de probabilités et donc d'obtenir un classement (Lu and Ades, 2004). Nous avons fait le choix de n'inclure que les essais randomisés dans la méta-analyse en réseau pour limiter le nombre d'études, l'hétérogénéité et l'incohérence. Une distribution a priori non informative binaire a été utilisée pour les résultats dichotomiques et une distribution a priori non informative normale pour le modèle continu.

Nous n'avons pas réussi à obtenir la moyenne et l'écart type pour tous les résultats continus malgré le fait que nous avons contacté les auteurs pour essayer de recueillir un maximum de données manquantes. Nous avons dû, par conséquent, les estimer. Pour l'estimation, nous avons utilisé les formules décrites dans les annexes des différentes études. La majorité des formules utilisées font l'hypothèse d'une distribution gaussienne.

Concernant les congrès américains, nous n'avons pas trouvé d'informations sur les communications de l'AAOS après 2015 et pour les communications de l'AAKS après 2012. C'est un potentiel risque de biais pour identifier la littérature grise. Néanmoins, ce risque est limité car nous avons recherché des essais randomisés non publiés dans ClinicalTrials.gov et dans la plateforme d'enregistrement international des essais cliniques (ICTRP) de l'Organisation mondiale de la santé.

Les résultats des classements sont basés sur des probabilités. En conséquence, nous ne pouvons pas les considérer comme des résultats fixes et définitifs mais plutôt comme des tendances. Cependant, pour tous les résultats, les classements étaient bien établis entre toutes les interventions. Donc, nous pensons que nous pouvons prendre en compte ces classements pour faire le meilleur choix.

Par ailleurs, nous n'avons pas évalué le coût de chaque guide de coupe, ce qui est une caractéristique pouvant influencer le choix de l'instrumentation d'une PTG. Une autre limite potentielle est que le nombre d'études incluses dans chaque réseau n'était pas identique.

Enfin, une donnée que nous n'avons pas pu analyser dans ces méta-analyses est la courbe d'apprentissages de chaque technique.

Les méta-analyses en réseau comme toute méthodologie ont des limites. La principale limite est représentée par le biais de dissémination, c'est-à-dire le biais lié à l'identification des études. En effet les essais avec résultats statistiquement significatifs sont plus souvent publiés, plus rapidement publiés et plus souvent cités que les essais négatifs. La méta-analyse a donc plus de chance d'être positive. C'est pour ça qu'il faut s'assurer que le recherche d'article soit exhaustive. Un autre biais possible des méta-analyses est le fait qu'il s'agit d'un processus rétrospectif. Les données existent déjà et les choix réalisés (critères de sélection, de jugement, analyse statistique) peuvent conditionner les résultats. Il faut s'assurer qu'un protocole a été fait et a été enregistré sur PROSPERO. Enfin, la dernière limite est qu'il s'agit d'un processus subjectif pour la sélection des études et l'extraction des données. Il est nécessaire pour limiter ce biais que ces étapes soient effectuées par deux intervenants indépendants.

Au total, les méta-analyses en réseau en orthopédie est une méthodologie fiable et utile afin d'obtenir de nouvelles informations et de répondre à des questions précises. Il faut cependant être critique par rapport aux conclusions et savoir être modéré.

Partie 3

Évaluation de l'apprentissage

Chapitre 7

Méthode de modélisation des courbes d'apprentissage en orthopédie

Après avoir statué sur deux aspects qui peuvent potentiellement avoir un impact sur le résultat à court, moyen et long terme des prothèses totales de genoux, nous nous sommes intéressés à l'apprentissage dans le domaine médical et à son évaluation. En effet, de plus en plus de nouvelles techniques et de technologies font leur apparition en médecine et en notamment en chirurgie orthopédique. La question de l'apprentissage devient alors essentielle. Par exemple, si une nouvelle technologie apporte un bénéfice en termes de précision, mais que pour savoir l'utiliser correctement et être reproductible, il faut réaliser une centaine d'interventions, alors son bénéfice est limité.

L'acquisition de compétences théoriques ou pratiques représente un défi tant pour l'enseignant, que pour l'étudiant. Les courbes d'apprentissage permettent d'observer l'acquisition d'une tâche au fil du temps. Elles ont été développées et largement utilisées pendant la Seconde Guerre mondiale pour la production d'armes (Hall, 2002). Elles ont ensuite été diffusées dans différents domaines, notamment dans l'industrie et plus tard dans la santé (Ramsay et al., 2002; Wright, 1936). L'un des principaux avantages des courbes d'apprentissage est l'association d'une représentation numérique et graphique du processus d'acquisition. La médecine s'est intéressée, en particulier dans les années 1980, aux courbes d'apprentissage afin de déterminer l'efficacité d'une procédure et le temps nécessaire pour acquérir les nouvelles techniques par les étudiants (Ramsay et al., 2002; Valsamis et al., 2018a).

Plusieurs techniques ont été développées pour les décrire en médecine (Valsamis et al., 2018b). L'une des plus grandes difficultés des courbes d'apprentissage est de modéliser l'apprentissage d'une compétence et ce qu'elle varie potentiellement selon les domaines. Afin de pouvoir utiliser au mieux les courbes d'apprentissage en orthopédie nous avons effectué une revue de la

littérature pour recenser les différentes méthodologies existantes et définir la plus adaptée à notre spécialité. Nous avons aussi recherché si des tests statistiques existaient permettant de comparer deux courbes d'apprentissage entre elles.

BMC Medical Education

Learning curves in orthopedic and trauma surgery: a systematic review.

--Manuscript Draft--

Manuscript Number:	MEED-D-21-00415
Full Title:	Learning curves in orthopedic and trauma surgery: a systematic review.
Article Type:	Research article
Section/Category:	Assessment and evaluation of admissions, knowledge, skills and attitudes
Funding Information:	
Abstract:	<p>Purpose: The processes of teaching, learning, and evaluating technical skills represent a challenge for surgeons. In the past, companionship has allowed the transmission of knowledge but without evaluating its performance. Techniques in development were performed most often with limited or without any validation or control. But, recently, a need for a change in that domain emerged. In order to handle this important issue, the number of studies interested in learning curves in medicine, particularly in orthopedic and trauma surgery, is increasing. The aim of this study was to assess if a technique provides a better approach to evaluate the learning curve in orthopedic and trauma surgery and if methods have been developed to compare the learning curves between them. Methods: A systematic literature search of Medline was carried out. From each article, data regarding the design of the study, the specialty, the primary objective, the primary endpoint, the series sizes, the methods applied to model the learning curves, the use of graphs and the use of tests were retrieved. We used Principal Component Analysis (PCA), to enhance characteristics shared by different studies and illustrate some pattern in those kind of studies Results: 119 studies were found relevant. Most trials were single center, retrospective and public funding. We found 6 different methods applied to evaluate learning curve: CUSUM (cumulative summation) methods, LC-CUSUM (Learning Curve cumulative summation), comparisons of experimented vs beginners' participants, comparisons of first procedures vs others, general linear regression models and survival models. Using PCA, we found 3 different clusters. We didn't find any application of a method which compares directly two learning curves. Conclusion: In orthopedic and trauma literature, 6 different methods are used to model and evaluate learning curves. The LC-CUSUM test seems to be the best technique to model learning curve because it's the simplest, most flexible and robust technique. No method was found in order to compare the learning curves between them.</p>
Corresponding Author:	pierre alban bouche Universite Paris Diderot FRANCE
Corresponding Author E-Mail:	pierrealban309@gmail.com
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universite Paris Diderot
Corresponding Author's Secondary Institution:	
First Author:	pierre alban bouche, MD
First Author Secondary Information:	
Order of Authors:	pierre alban bouche, MD Matthieu Resche-Rigon, MD,PHD Rémy Nizard, MD,PHD
Order of Authors Secondary Information:	
Opposed Reviewers:	
Additional Information:	

Question	Response
Has this manuscript been submitted before to this journal or another journal in the BMC series</ a>?	No

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

Learning curves in orthopedic and trauma surgery: a systematic review.

Pierre-Alban Bouché^{1,2} (resident), Matthieu Resche-Rigon^{1,3} (MD PHD) *, Rémy Nizard²
(MD PHD) *

¹ Université de Paris, ECSTRAA Team, UMR U1153, INSERM, Paris, France

² Service de chirurgie orthopédique et traumatologique, Hôpital Lariboisière, APHP, Paris,
France

³ Service de Biostatistique et Information Médicale, Hôpital Saint-Louis, APHP, Paris, France

*Co-last author

Running title: Learning curves in orthopedic and trauma surgery

33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Corresponding author:

M. Bouché ✉

INSERM U1153

Hôpital Saint-Louis

Bâtiment Lailler 3ème étage

1, avenue Claude Vellefaux

75010 Paris, France

Email: pierrealban309@gmail.com

1	1	Learning curves in orthopedic and trauma surgery: a systematic
2		
3	2	review.
4		
5		
6	3	
7		
8	4	
9		
10		
11	5	
12		
13	6	
14		
15		
16	7	
17		
18	8	
19		
20		
21	9	
22		
23	10	
24		
25		
26	11	
27		
28	12	
29		
30	13	<u>Key words:</u> Orthopaedic, trauma, education, learning curve, models
31		
32		
33	14	
34	15	
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		

16 Abstract:

17

18 Background: The processes of teaching, learning, and evaluating technical skills represent a
19 challenge for surgeons. In the past, companionship has allowed the transmission of
20 knowledge but without evaluating its performance. Techniques in development were
21 performed most often with limited or without any validation or control. But, recently, a need
22 for a change in that domain emerged. In order to handle this important issue, the number of
23 studies interested in learning curves in medicine, particularly in orthopedic and trauma
24 surgery, is increasing. The aim of this study was to assess if a technique provides a better
25 approach to evaluate the learning curve in orthopedic and trauma surgery and if methods have
26 been developed to compare the learning curves between them.

27

28 Methods: A systematic literature search of Medline was carried out. From each article, data
29 regarding the design of the study, the specialty, the primary objective, the primary endpoint,
30 the series sizes, the methods applied to model the learning curves, the use of graphs and the
31 use of tests were retrieved. We used Principal Component Analysis (PCA), to enhance
32 characteristics shared by different studies and illustrate some pattern in those kind of studies

33

34 Results: 119 studies were found relevant. Most trials were single center, retrospective and
35 public funding. We found 6 different methods applied to evaluate learning curve: CUSUM
36 (cumulative summation) methods, LC-CUSUM (Learning Curve cumulative summation),
37 comparisons of experimented vs beginners' participants, comparisons of first procedures vs
38 others, general linear regression models and survival models. Using PCA, we found 3
39 different clusters. We didn't find any application of a method which compares directly two
40 learning curves.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

41

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

42 Conclusion: In orthopedic and trauma literature, 6 different methods are used to model and
43 evaluate learning curves. The LC-CUSUM test seems to be the best technique to model
44 learning curve because it's the simplest, most flexible and robust technique. No method was
45 found in order to compare the learning curves between them.
46 Level of evidence: Level II, systematic review

47 Introduction:

1
2
3 48 Acquisition of technical skills represents a challenge for both the teacher and the learner. The
4
5 49 learning curves represent the acquisition of a task over time. They were developed and largely
6
7
8 50 used during the Second World War for the production of weapons [1]. They were then
9
10 51 disseminated in various field, notably in the industry and later in health [2, 3]. One of the
11
12 52 main advantages of learning curves is the association of numerical and graphical
13
14
15 53 representations of the learning process. Thus, a direct application would be to compare them
16
17 54 with each other in order to evaluate the potential of learning different techniques. Medicine
18
19 55 became interested, particularly in the 1980s, in learning curves in order to determine the
20
21 56 effectiveness of procedure and time to acquire by students the new techniques [2, 4]. Several
22
23 57 techniques to describe them have been developed in medicine [5]. One of the biggest
24
25 58 difficulties in learning curves is to model the learning of the skill. A skill that varies to
26
27 59 different domains. Learning curves have mainly been developed and used in surgery because
28
29 60 patients need to know if their surgeon is competent for a given procedure. Indeed,
30
31 61 acceptability by the patients and the authorities of being treated by a surgeon unable to prove
32
33 62 his abilities to perform a conventional or developing technique dramatically decreased.
34
35 63 Additionally, before autonomizing a junior surgeon in a team, an objective measure of skill
36
37 64 acquisition is highly desirable. Therefore, the traditional way of this acquisition process
38
39 65 mostly done without objective measurements is questionable. Recently, developments of
40
41 66 evaluation of technical skills learning have been observed. In recent years, the development of
42
43 67 minimally invasive techniques and the arrival of new technologies, notably robotics, have
44
45 68 seen the use of learning curves as an evaluation criterion increased [6]. Indeed, by searching
46
47 69 on Pubmed for "learning curve", in 2000, 204 articles were published against 2718 in 2019.
48
49 70 We focused on the learning curve for orthopedic or trauma surgical procedures to observe if a
50
51 71 technique provide a better way to evaluate the learning curve in this area. To reach this aim,
52
53
54
55
56
57
58
59
60
61
62
63
64
65

72 we performed a systematic review on the different technics used in order to model the
73 learning curves for orthopedic or trauma surgical procedures. To our knowledge, none method
74 to compare different learning curves has been described, presumably due to the critical
75 difficulties to model them, and thus to develop comparison test [5]. So, the second aim of this
76 study was to observe methods available to compare the learning curves between them.
77 Finally, we described the different studies found using Principal Component Analysis (PCA)
78 [7], to enhance characteristics shared by different studies and illustrate some pattern in those
79 kinds of studies. Our hypothesis was that several methods existed to model learning curve but
80 without guidelines and that none method was described to compare different learning curves.

81

82

83 **SEARCH SRATEGY AND CRITERIA**

84

85 *Eligibility criteria*

86 Studies were included if they reported on learning curve of orthopaedic and trauma surgical
87 procedure and if the article was a clinical research study. No restriction on language or date of
88 publication was applied.

89

90 *Information sources*

91 For this review, studies were identified by searching Medline electronic database via PubMed.
92 Initial search was run on 18/08/2020. Finally, reference lists of systematic reviews and meta-
93 analyses were systematically screened for any additional references.

94

95 *Search*

96 We used the following search terms:

97 (((learning curve[MeSH Terms]) OR (learning curves[MeSH Terms])) OR (learning

98 curve[Title])) OR (learning curves[Title])) AND (((orthopedic[MeSH Terms]) OR

99 (orthopedic procedure[MeSH Terms])) OR (traumatology[MeSH Terms]))

100

101 *Study selection*

102 One author made a first selection of eligible studies based on title and abstracts. All eligible

103 studies were then extracted, and further selection was performed based on the full text by the

104 same author.

106 *Data collection process*

107 A data extraction sheet was developed by all authors. It was pilot tested on five studies and

108 refined. All selected studies were then reviewed independently by one author. Because this

109 was a study on how reporting is performed, authors were not contacted to provide missing

110 information.

111

112 *Data items*

113 For the review, we collected the following data from each article: General study information

114 (date of publication, country, authors, title, funding sources), specialty, type of intervention,

115 design (retrospective or prospective) of the study, primary objective of the study, primary

116 endpoint of the study, series size, method used to model the learning curves and test applied if

117 learning curves were compared. We collected also the type of the learning curves graphically

118 represented.

119

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65120 *Summary measures*

121 Binary outcomes are reported as number and percentage. Continuous outcomes are presented
122 with median and min and max values. Cluster analysis is a method aiming at identifying
123 groups sharing some pattern of characteristics. To achieve this goal, PCA reduce the
124 dimensions of observed continuous characteristics and allows us to highlight the main
125 differences in the different trials [7]. In this paper, study characteristics were quantitative and
126 qualitative variables, so we used Factor Analysis of Mixed Data (FAMD) which extends PCA
127 to different type of variables [8]. We used the *FactoMineR* R-package [9] for FAMD and
128 *factoextra* R-package [10] for cluster analyses in R version 3.5.2 software (R foundation for
129 statistical Computing, Vienna, Austria).

130

131 *Patient Involvement*

132 Patients were not involved in any aspect of the study design, conduct, or development of the
133 research question or outcome measures. This study is a systematic review of existing
134 published research; therefore, there was no active patient recruitment for data collection.

135

136 **Results:**137 *Selection and General Characteristics*

138 Our search initially retrieved 358 records. None duplicated article was found, thus, the titles
139 and abstracts of 358 studies were reviewed. 178 reports were included in the full text articles
140 assessed for eligibility. Finally, 119 studies met our inclusion criteria and were eligible in the
141 review (fig 1). Publication years ranged from 1992 to 2019, with a recent increase in the
142 number of publications. Most trials were run in a single center and funded by public grants

143 (respectively 97.5% and 87.4%). Studies were performed in 25 different countries in 5
144 different continents. Europe and North America were the most represented continents (33.6%
145 and 32.8% respectively). The design was retrospective in 109 (89.1%) studies. Sub-specialties
146 of interest were prosthesis in 48 (40.3%) studies, spine in 30 (25.2%), arthroscopy in 24
147 (20.2%), trauma in 11 (9.2%), foot surgery in 3 (2.5%), pediatric surgery in 2 (1.7%), and
148 tumor in 1 (0.9%). The characteristics of trials are detailed in the table 1 and the detailed of
149 each trial are presented in appendix A.

150

151 *Models used for learning curve:*

152 The aim of trials was the analysis of the learning curve of a new technique in 110 (92.4%)
153 studies and for education of students (mainly residents) in 9 (7.6%). The primary objective
154 was to model the learning curve of one technique in 102 (85.7%) studies, to compare the
155 learning curves of 2 surgical techniques in 14 (11.8%) and to compare the learning curves of 3
156 surgical techniques in 3 (2.5%). The median (range) number of procedures controlled per
157 study was 100 (8–8041). The median of number of participants per study was 1 (1-251). We
158 found 6 different approaches to analyze learning curves in orthopedic or trauma surgical
159 procedures: CUSUM (cumulative summation) methods, LC-CUSUM (Learning Curve
160 cumulative summation), comparisons of experimented vs beginners' participants,
161 comparisons of first procedures vs others (i.e. following procedures), general linear regression
162 models and survival models. The most frequent methods used was “comparisons of first
163 procedures vs others”: 75 studies (63.0%) (table 2). A graph was plotted in 89 (74.8%)
164 studies. Most of the times, plots consist in scatter plots in 37 (41.6%) studies.

165

166 *Methods used for learning curve comparisons:*

167 A comparison between different learning curve were presented in 30 (25.2%) studies. In most
168 of the cases (27 trials (90.0%), standards tests were used (t-test, chi-2 test, ...) to compare
169 surgical parameters like surgical duration or the rate of complications, between procedures of
170 the two participants. In other cases, regression and survival test were used, respectively in 2
171 (6.7%) studies and 1 study (3.3%). All details are available on table 2. We didn't find any
172 method which compares directly the learning curve and allows to assess if a student acquire
173 the competence before another with a significant difference neither if a technique had a
174 significant quicker learning curve.

175

176 *Principal Component analysis and clusters:*

177 For the FAMD, we selected 8 variables: 2 quantitative variables and 6 qualitative variables.

178 The 2 quantitative variables were: the number of procedures and the number of participants.

179 The 6 qualitative variables were: the design, the aim, the main outcome of the study, the

180 methods used, if a comparison were made and if a graph was presented. We find 3 different

181 clusters among studies, represented in figure 2. The first cluster included studies aiming the

182 analysis of the learning curves for a new technique, using "First versus others" methods, not

183 making other comparison and a small number of participants per study (mean of 1.6). The

184 second cluster included studies aiming the analysis of the learning curves for students

185 education, including a larger number of participants (mean of 25.2 per study) and considering

186 a great number of procedures (mean of 933.6) . The last cluster included studies comparing

187 "Experimented vs beginners" surgeons and making comparisons. The characteristics of

188 clusters are detailed in the table 3.

189

190 **Discussion:**

191 The main result of our study is that 6 different approaches were identified in the literature.

192 This result was obtained after 119 studies were included among 358 retrieved relevant

193 articles.

194 Learning curves in medicine allow to observe when a procedure is acquired or when a person

195 in training has mastered the procedure or the surgical technique [11, 12]. It is important to

196 evaluate these curves in order to answer to patients and institutional demand to demonstrate

197 that mastering of the technique was achieved to improve quality and safety of care. If there is

198 still no duty of result, demonstration of a continuous process of evaluation, especially for new

199 techniques or younger surgeons appears mandatory for the future. This trend is confirmed by

200 a Pubmed extraction over the past 20 years, which demonstrate that the number of studies

201 stating in their title or their abstract "learning curve" has skyrocketed: from 155 in 1999 to

202 537 in 2009 and 951 in 2019, i.e. a 6.3-fold increase in 20 years.

203 To our knowledge and despite the craze for learning curves, none review has reported any

204 information on the methods used. Thus, a review of the literature seemed interesting to more

205 strictly evaluate the methods used to report learning process.

206 We deliberately had chosen to search only for the word "learning curve" in the title in

207 PubMed because this term is used in an abusive manner possibly to try to promote a study in

208 our opinion. The choice to look at studies only in orthopedics was deliberately chosen in order

209 to be homogeneous and to be able to have reliable method in this field.

210

211 The first and most used method is "first vs others". This method compares the first patients

212 operated by a surgeon with the following ones. The threshold to define "first patients" is

213 either chosen arbitrarily or to equally separate the sample of patients or procedures. For

214 example, out of 100 procedures, the first 50 are compared to the next 50 or 4 groups of 25

1 215 procedures are sequentially created and compared. This method is easy to implement,
2
3 216 reproducible and easily understood. However, it does not take into account the fact that a
4
5 217 learning curve is not linear and can be composed of different phases. This notion is very well
6
7 218 detailed in the article by Valsamis [5]. Six different scenarios of learning process were
8
9 219 described using 4 models: “plateau” model, “line-line” model, “line-plateau” model and
10
11 220 “line” model.

12
13
14
15 221 The second most frequent method used is called, by us, “Experimented vs Beginners”. This
16
17 222 method is close to the "first vs others" but rather than taking procedures over time as unit
18
19 223 under control, authors compare the procedures performed by surgeons considered to be
20
21 224 experienced to those who are beginners. Thus, this approach shares the advantages and limits
22
23 225 of the previous approach mentioned above.
24
25
26
27

28 226 Two other methods are CUSUM and LC-CUSUM. CUSUMs have been developed and
29
30 227 applied by industry since the beginning of the last century to monitor procedures and detect
31
32 228 repeated deviations from a goal in a controlled process [13]. From the defined goal, limits are
33
34 229 determined for a controlled or uncontrolled process. They have been used in medicine for
35
36 230 about thirty years to monitor stable conditions (surgical procedures, mortality in a pathology,
37
38 231 occurrence of nosocomial infection) [14]. However, since this procedure is not adapted to an
39
40 232 uncontrolled process with a high failure rate as a learning phase can be, a modified version
41
42 233 has been developed to allow the monitoring of learning curves: the LC-CUSUM [15]. This
43
44 234 method has the same properties as the CUSUM but here, the limit is crossed when the student
45
46 235 has acquired the required skills in other words when it is judged that he has entered in a
47
48 236 controlled process. The main advantage of the LC-CUSUM is that it does not require a
49
50 237 parametric model of the learning curve. Whatever the type of learning curve only the passage
51
52 238 in a controlled process is detected. Thus, it is easy to implement, adaptable to every
53
54 239 operator and able to valid specifically this operator for a new technique. Furthermore, this
55
56
57
58
59
60
61
62
63
64
65

240 technic can be performed in other specialty [16, 17] due to its flexibility and its robustness.

241 The main limit of the LC-CUSUM is that it focuses on only one dimension thus on only one
242 characteristics of the learning process.

243 The penultimate method is the use of survival analysis approach. This method has been used
244 to see the survival of implants therefore in studies concerning prostheses. The design of such
245 studies is developp to assess if there is a difference in survival between implants placed early
246 in the experiment against those implanted later. This approach is very specific to protheses or
247 medical devices and represents a minority of studies. Some studies compared survival
248 analysis and CUSUM methods to detect aseptic loosening of Total Hip Arthroplasty (THA)
249 [18, 19]. CUSUM methods allowed to detect five years and half years before basic survival
250 analysis the poor performance of this THA [19]. Thus, CUSUM should be considered as an
251 very good alternative to standard statistical methods to early detect deviance.

252 The last method is based on a regression model. Valsamis has developed a whole method
253 concerning learning curves [4]. As previously introduced, six different scenarios of learning
254 curves were described using 4 models. This model is probably one of the most promising to
255 analyse learning curves. Nevertheless, each regression model is specific to each participant
256 and makes comparisons difficult. Moreover, this model could be difficult to apply.

257 To our knowledge and in view of the results of the review, the simplest, most flexible and
258 robust technique to model learning curve is the LC-CUSUM test. This technic could be
259 applied whatever the type of acquisition of the student and the type of the interest procedure.
260 Furthermore, it's easy to disseminate and it can be used in other specialty [16, 17].

261 The second interest of this review was to detect if a method was available to compare
262 different learning curves between them. We didn't find any test comparing directly learning
263 curves or allowing to detect if a student acquired the competence before another with a

1 264 significant difference neither if a technique had a significant quicker (or improved) learning
2 265 curve.

3
4
5 266 To our knowledge and in view of the results of the review, none method exists to compare
6
7
8 267 different learning curves. Different surgical techniques constantly appear and evolve, several
9
10 268 authors will want to compare the learning curves of these new techniques against the older in
11
12 269 order to show that they will not require too long or chaotic learning process. Thus, it appears
13
14
15 270 mandatory to develop approaches to compare these curves with a valid method.
16
17

18 **271 Conclusions:**

19
20
21 272 On orthopedic and trauma literature, 6 different methods are used to describe learning curves.
22

23
24 273 From our point of view, the better technics to model learning curve in orthopedic and trauma
25

26 274 surgery is the LC-CUSUM test. None method was found in order to compare the learning
27

28
29 275 curves between them.
30
31

32 276
33

34 **277 Declarations:**

35
36 278
37
38

39 279 Abbreviations:
40

41
42 280 • CUSUM (cumulative summation)
43

44 281 • FAMD: Factor Analysis of Mixed Data
45

46 282 • LC-CUSUM: Learning Curve Cumulative summation
47
48

49 283 • PCA: Principal Component Analysis
50

51
52 284 • THA: Total Hip Arthroplasty
53

54 285 Ethics approval and consent to participate: Not applicable. This is a systematic review.
55

56 286 Consent for publication: All authors gave their consent for publication.
57
58

59 287 Availability of data and material: available on request
60

61 288
62
63
64
65

289	Competing interests: The authors declare no conflict of interests.
1	
2	290 Funding: The researchers did not receive any funding to conduct this study.
3	
4	291 Authors' contributions:
5	
6	
7	292 • Pierre-Alban Bouché : Article design, data analysis and article writing
8	
9	293 • Matthieu Resche-Rigon: Design and proofreading the article
10	
11	
12	294 • Rémy Nizard: Design and proofreading the article
13	
14	295 Acknowledgements:none
15	
16	
17	296
18	
19	297
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	

- 299 1. Hall RW. Lessons from the Learning Curve Part 2 of Out of the Box : Overcoming Our
300 Mind Set Legacies. 2002. /paper/Lessons-from-the-Learning-Curve-Part-2-of-Out-of-%3A-
301 Hall/82149c48a0a38ef6ca1a56999db3781fb7e4baa8. Accessed 24 Jul 2020.
- 302 2. Ramsay CR, Wallace SA, Garthwaite PH, Monk AF, Russell IT, Grant AM. Assessing the
303 learning curve effect in health technologies. Lessons from the nonclinical literature. *Int J*
304 *Technol Assess Health Care*. 2002;18:1–10.
- 305 3. WRIGHT TP. Factors Affecting the Cost of Airplanes. *Journal of the Aeronautical*
306 *Sciences*. 1936;3:122–8.
- 307 4. Valsamis EM, Chouari T, O’Dowd-Booth C, Rogers B, Ricketts D. Learning curves in
308 surgery: variables, analysis and applications. *Postgraduate Medical Journal*. 2018;94:525–30.
- 309 5. Valsamis EM, Golubic R, Glover TE, Husband H, Hussain A, Jenabzadeh A-R. Modeling
310 Learning in Surgical Practice. *Journal of Surgical Education*. 2018;75:78–87.
- 311 6. Soomro NA, Hashimoto DA, Porteous AJ, Ridley CJA, Marsh WJ, Ditto R, et al.
312 Systematic review of learning curves in robot- assisted surgery. *BJS Open*. 2019;4:27–44.
- 313 7. Abdi H, Williams LJ. Principal component analysis. *WIREs Computational Statistics*.
314 2010;2:433–59.
- 315 8. Pagès J. Analyse factorielle de données mixtes. *Revue de Statistique Appliquée*.
316 2004;52:93–111.
- 317 9. Husson F, Josse J, Le S, Mazet J. FactoMineR: Multivariate Exploratory Data Analysis and
318 Data Mining. 2020. <https://CRAN.R-project.org/package=FactoMineR>. Accessed 24 Sep
319 2020.
- 320 10. Kassambara A, Mundt F. factoextra: Extract and Visualize the Results of Multivariate
321 Data Analyses. 2020. <https://CRAN.R-project.org/package=factoextra>. Accessed 24 Sep
322 2020.
- 323 11. Bristol case surgeon claimed to have been on “learning curve.” *BMJ*. 1999;319:1456.
- 324 12. Green ML, Aagaard EM, Caverzagie KJ, Chick DA, Holmboe E, Kane G, et al. Charting
325 the road to competence: developmental milestones for internal medicine residency training. *J*
326 *Grad Med Educ*. 2009;1:5–20.
- 327 13. Steiner SH, Cook RJ, Farewell VT. Risk-adjusted monitoring of binary surgical outcomes.
328 *Med Decis Making*. 2001;21:163–9.
- 329 14. Biau DJ, Resche-Rigon M, Godiris-Petit G, Nizard RS, Porcher R. Quality control of
330 surgical and interventional procedures: a review of the CUSUM. *Qual Saf Health Care*.
331 2007;16:203–7.
- 332 15. Biau DJ, Porcher R. A method for monitoring a process from an out of control to an in
333 control state: Application to the learning curve. *Stat Med*. 2010;29:1900–9.

- 334 16. Dessolle L, Fréour T, Barrière P, Jean M, Ravel C, Daraï E, et al. How soon can I be
1 335 proficient in embryo transfer? Lessons from the cumulative summation test for learning curve
2 336 (LC-CUSUM). *Hum Reprod.* 2010;25:380–6.
3
- 4 337 17. Weil G, Motamed C, Biau DJ, Guye ML. Learning curves for three specific procedures by
5 338 anesthesiology residents using the learning curve cumulative sum (LC-CUSUM) test. *Korean*
6 339 *J Anesthesiol.* 2017;70:196–202.
7
8
- 9 340 18. Haroon SL, Lewsey JD, Gregg PJ, Reeves BC, van der Meulen JHP. Continuous
10 341 monitoring of the performance of hip prostheses. *The Journal of Bone and Joint Surgery*
11 342 *British volume.* 2006;88-B:716–20.
12
13
- 14 343 19. Haroon SL, Lewsey JD, Meulen JHP van der. Continuous monitoring of long-term
15 344 outcomes with application to hip prostheses. *Statistics in Medicine.* 2007;26:5081–99.
16
17 345
18
19 346
20
21
22 347
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

348

Tables:1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

349 Table 1. Characteristics of trials

Parameters	Values	N	Statistics*
		119	
Funding	Private	15	12.6 %
	Public	104	87.4 %
Continent	Asia	31	26.1 %
	Europe	40	33.6 %
	International	1	0.8 %
	North america	39	32.8 %
	Oceania	7	5.9 %
	South america	1	0.8 %
N° of center	mono	116	97.5 %
	multi	3	2.5 %
Design	Prospectif	13	10.9 %
	Retrospectif	106	89.1 %
Sub-speciality	Arthroscopy	24	20.2 %
	Foot	3	2.5 %
	Pediatric	2	1.7 %
	prosthesis	48	40.3 %
	Spine	30	25.2 %
	Trauma	11	9.2 %
	Tumor	1	0.9 %

350

351 Table 2. Characteristics of methods used for learning curve comparisons

Parameters	Values	N	Statistics*
Aim	Education	9	7.6 %
	Technique	110	92.4 %
Main outcome	Comparison of 2 techniques	14	11.8 %
	Comparison of 3 techniques	3	2.5 %
	Learning curve of one technique	102	85.7 %
N° of procedures median(min-max)		118	100 (8-8041)

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

N° of participants median(min-max)		114	1(1-251)
Methods	CUSUM	12	10.1 %
	Experimented vs beginners	14	11.8 %
	First versus others	75	63.0 %
	LC-CUSUM	5	4.2 %
	Regression	12	10.1 %
	Survival	1	0.8 %
Graph	No	30	25.2 %
	Yes	89	74.8 %
Comparison	No	89	74.8 %
	Yes	30	25.2 %
Type of tests (if there is a comparison)	Regression	2	6.7 %
	Survie	1	3.3 %
	Tests standard	27	90.0 %

352

353

Table 3. Characteristics of clusters:

Cluster	Variables	N	Statistics ⁺
1		95	
	Comparaison : « no »	91	95.5 %
	Aim : « technique »	82	86.2%
	Methods : « first versus others »	87	91.8%
	Center : « mono centric »	79	82.8%
	Number of participants (mean [sd])		1.6 [2.2]
2		9	
	Aim : « education »	8	88.9 %
	Number of participants (mean [sd])		25.2 [16.8]
	Number of procedures (mean [sd])		933.6 [237.6]
3		14	
	Comparaison : « Yes »	6	46.7%

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

	Methods : « Experimented vs beginners »	14	100.0%
	Center : « multi centric »	14	100.0%
	Graph : »no »		23.3%

354

Legends

Fig. 1 This PRISMA flow chart shows the selection process of the review of learning curves in orthopedic and trauma surgery.

Fig. 2 The different cluster found with FAMM among studies. Cluster 1 included studies

which the aim was the analysis of the learning curve for a new technique, using “First versus

others” methods and not making comparison. Cluster 2 included studies which the aim was

the analysis of the learning curve for education to student. Cluster 3 included studies which

used “Experimented vs beginners” methods and making a comparison.

360

Additional files:

- Appendix_A: Appendix of the study

- PRISMA: Prisma check list

364

365

366

367

368

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1.

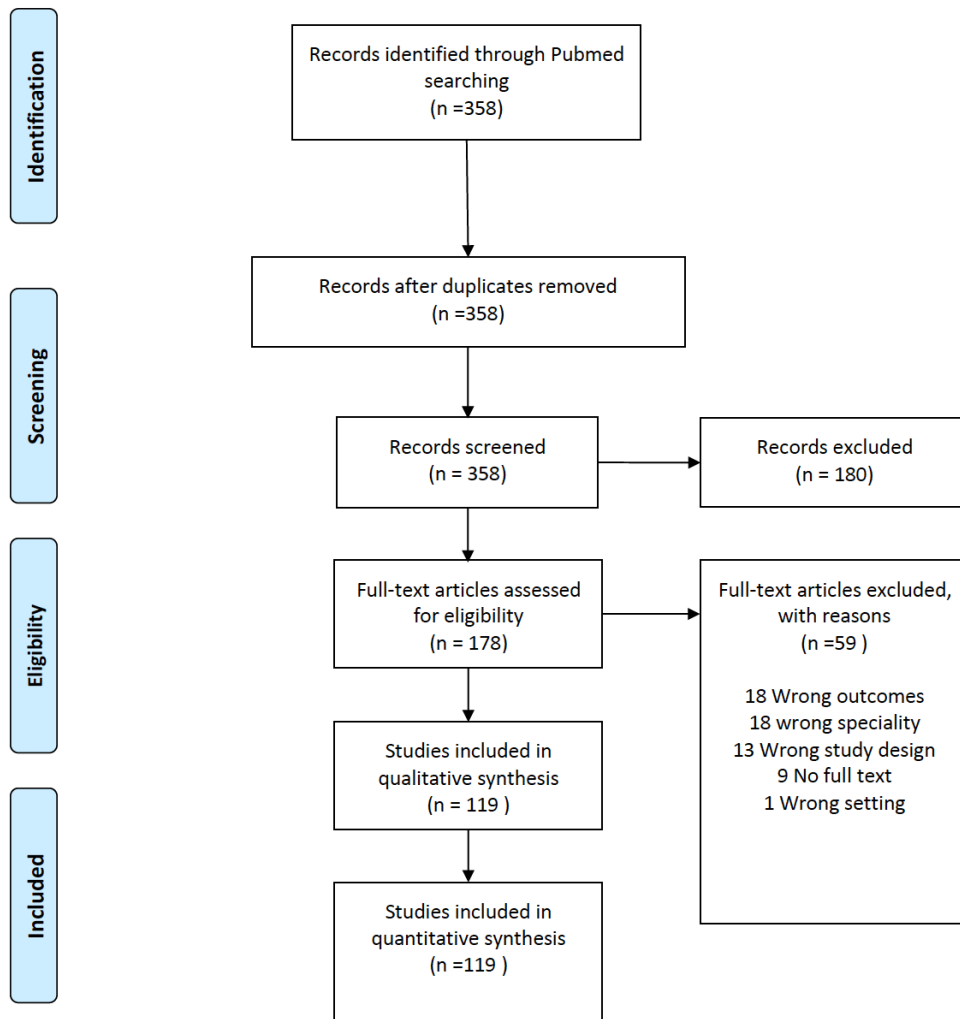
Parameters	Values	N	Statistics*
		119	
Funding	Private	15	12.6 %
	Public	104	87.4 %
Continent	Asia	31	26.1 %
	Europe	40	33.6 %
	International	1	0.8 %
	North america	39	32.8 %
	Oceania	7	5.9 %
	South america	1	0.8 %
N° of center	mono	116	97.5 %
	multi	3	2.5 %
Design	Prospectif	13	10.9 %
	Retrospectif	106	89.1 %
Sub-speciality	Arthroscopy	24	20.2 %
	Foot	3	2.5 %
	Pediatric	2	1.7 %
	prosthesis	48	40.3 %
	Spine	30	25.2 %
	Trauma	11	9.2 %
	Tumor	1	0.9 %

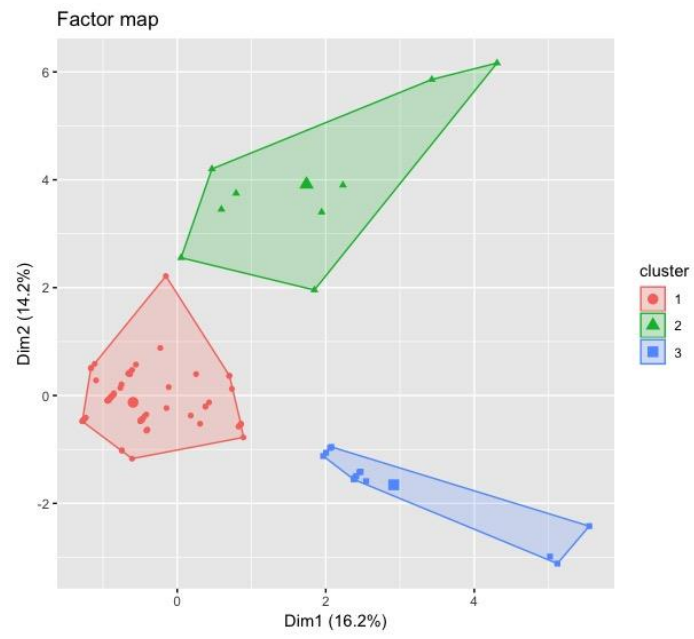
Table 2.

Parameters	Values	N	Statistics*
Aim	Education	9	7.6 %
	Technique	110	92.4 %
Main outcome	Comparison of 2 techniques	14	11.8 %
	Comparison of 3 techniques	3	2.5 %
	Learning curve of one technique	102	85.7 %
N° of procedures median(min-max)		118	100 (8-8041)
N° of participants median(min-max)		114	1(1-251)
Methods	CUSUM	12	10.1 %
	Experimented vs beginners	14	11.8 %
	First versus others	75	63.0 %
	LC-CUSUM	5	4.2 %
	Regression	12	10.1 %
	Survival	1	0.8 %
Graph	No	30	25.2 %
	Yes	89	74.8 %
Comparison	No	89	74.8 %
	Yes	30	25.2 %
Type of tests (if there is a comparison)	Regression	2	6.7 %
	Survie	1	3.3 %
	Tests standard	27	90.0 %

Table 3:

Cluster	Variables	N	Statistics*
1		95	
	Comparaison : « no »	91	95.5 %
	Aim : « technique »	82	86.2%
	Methods : « first versus others »	87	91.8%
	Center : « mono centric »	79	82.8%
	Number of participants (mean [sd])		1.6 [2.2]
2		9	
	Aim : « education »	8	88.9 %
	Number of participants (mean [sd])		25.2 [16.8]
	Number of procedures (mean [sd])		933.6 [237.6]
3		14	
	Comparaison : « Yes »	6	46.7%
	Methods : « Experimented vs beginners »	14	100.0%
	Center : « multi centric »	14	100.0%
	Graph : »no »		23.3%





Chapitre 8

Méthode de sommation cumulées

Suites à notre revue de la littérature, nous nous sommes intéressés aux méthodes de sommation cumulée et notamment au LC-CUSUM test pour modéliser l'apprentissage. Dans cette partie est présentée les principes généraux du CUSUM et du LC-CUSUM.

A. CUSUM :

Le CUSUM test (Cumulative Sum) a été développé, initialement, afin d'évaluer le contrôle qualité dans le domaine de l'industrie. Il s'agit d'une technique d'analyse séquentielle publiée en 1954 pour la première fois (Page,1954). Le CUSUM a ensuite été développé et utilisé dans le domaine de la santé. Une adaptation de la méthode statistique a été effectuée afin de pouvoir surveiller de manière prospective un processus qui devrait être stable et continu dans le temps comme le résultat d'une intervention chirurgicale ou le taux d'infections nosocomiales.

Afin d'évaluer le contrôle qualité d'une intervention, il faut définir une variable d'intérêt qui permettra de suivre les performances au cours du temps. Cette variable de suivi peut être continue (temps de réalisation de la procédure par exemple) ou binaire (taux d'échec d'une procédure). Il faut bien choisir ce paramètre afin que le CUSUM puisse émettre une alarme le plus rapidement possible lorsqu'une déviation se produit dans la procédure évaluée. Dans la suite de l'explication, nous ne considérons que le cas de variable binaire.

Le CUSUM (S_t) est une somme des scores (W_t) attribués à chaque procédure (t) avec une valeur dépendant du résultat de la procédure. Ces scores attribués à chaque procédure sont, par convention, positifs lorsqu'ils sont associés à un échec et négatifs lorsqu'ils sont associés à des réussites. Ainsi, le score CUSUM augmente avec la détérioration des performances et l'accumu-

lation des échecs.

Le CUSUM est conçu pour détecter des augmentations ou des diminutions de p (taux d'échec). Il teste séquentiellement l'hypothèse $H_0 : p = p_0$ contre $H_1 : p = p_1$ où p_0 est déterminé par la performance actuelle du processus (taux d'échec acceptable lorsque la performance est adéquate et que le processus est sous contrôle), tandis que p_1 représente une valeur alternative d'intérêt correspondant à des performances inférieures (taux d'échec inacceptable lorsque les performances deviennent inadéquates et que le processus devient incontrôlable). Le choix de p_0 et p_1 dépend de ce qui est jugé acceptable et de la performance considérée comme inadéquate selon l'expérience ou les données de la littérature.

Le CUSUM est la somme S_t de t procédures, dont la formule mathématique est :

$$S_t = \max(0, S_{t-1} + W_t)$$

Avec $t (= 1, 2, 3, \dots)$ le nombre de procédure, $S_0 = 0$ le score CUSUM au début du processus, et W_t le score attribué à la t ème procédure, en fonction du succès ou non de l'intervention.

Le CUSUM a pour objectif de détecter une dégradation des performances d'un processus. Les valeurs du CUSUM sont forcément positives. Le fait de ne pas pouvoir attribuer de valeurs négatives évite l'accumulation de scores négatifs générés par des séries de succès. Cette particularité permet au CUSUM de détecter rapidement et d'être sensible aux séries récentes de mauvaises performances. Par exemple, sans cette spécificité, le score d'un chirurgien ayant cumulé des succès pourrait être très inférieur à 0 et un changement dans sa pratique conduisant à la survenue d'échecs serait alors masqué. En effet, pour atteindre le seuil d'alerte fixé, le nombre d'échecs « nécessaires » devra être plus important en partant d'un score qu'en partant de 0.

Le choix optimal pour W_t est basé sur le rapport de log-vraisemblance. Pour une variable binaire et pour la procédure t , si en cas de succès $Y_t = 0$ et si en cas d'échec $Y_t = 1$ alors nous avons comme scores CUSUM:

$$W_t = \log [(1 - p_1) / (1 - p_0)] \quad \text{if } Y_t = 0 \text{ (succès)}$$

$$W_t = \log (p_1 / p_0) \quad \text{if } Y_t = 1 \text{ (échec)}$$

Le passage de l'hypothèse H_0 à l'hypothèse H_1 se produit lorsque le score du CUSUM a atteint un seuil limite de contrôle noté h ($S_i \geq h$). Lorsque ce seuil est atteint cela signifie que le processus n'est plus sous contrôle, le test émet alors une alarme. Le but de cette alarme est de générer un audit et de vérifier le processus afin de détecter la cause de la détérioration. Cela permet de mettre en places des mesures nécessaires afin de résoudre le problème avant de réinitialiser la surveillance. Le choix de la limite h définit la sensibilité et la spécificité du test. Plus la limite sera basse, plus tôt elle détectera des changements dans le paramètre surveillé mais plus grand sera le taux de faux positifs (signaux déclenchés bien que le processus soit sous contrôle). À l'inverse, plus il sera élevé, plus tard les changements seront détectés avec un risque plus élevé d'ignorer les vrais positifs.

En cas de suivi d'une courbe d'apprentissage d'un stagiaire, l'objectif est de détecter quand le stagiaire atteint des performances adéquates. Dans ce cas, le test CUSUM est insatisfaisant.

B. LC-CUSUM :

Biau et col. (Biau and Porcher, 2010) ont proposé une nouvelle méthode statistique, le test LC-CUSUM, qui a la capacité de signaler lorsqu'un processus atteint un état de contrôle, c'est-à-dire lorsqu'un stagiaire a atteint un niveau de performance adéquat. Par conséquent, les hypothèses sont inversées par rapport à CUSUM : Le LC-CUSUM teste après chaque procédure Y_i l'hypothèse nulle $H_0: p = p_0 + \delta$ (le processus est hors de contrôle) par rapport à l'hypothèse alternative $H_1 : p = p_0$ (le processus est sous contrôle) où p_0 est la probabilité d'échec lorsque les performances sont adéquates et δ un écart maximal acceptable par rapport aux performances adéquates. Les scores attribués à chaque procédure sont, à l'inverse du CUSUM, positifs lorsqu'ils sont associés à des réussites et négatifs lorsqu'ils sont associés à des échecs. Ensuite, le score LC-CUSUM augmentera avec l'amélioration des performances. Comme pour le choix de p_0 et p_1 pour le CUSUM, le choix de p_0 et δ pour le LC-CUSUM est quelque peu arbitraire car il n'y a pas de consensus sur la manière de sélectionner ces paramètres dans la littérature. Pour le CUSUM, $|p_1 - p_0|$ correspond à la différence de performance (performance inacceptable moins performance acceptable). Pour le LC-CUSUM, l'écart acceptable par rapport à la cible, peut varier entre 0 et cette différence $|p_1 - p_0|$. Comme proposé pour les essais d'équivalence, un raisonnable choix de cet écart δ peut être $\delta = (p_1 - p_0)/2$. Cette option est le meilleur compromis entre ce qui est jugé acceptable pour le LC-CUSUM et la performance considérée comme

inadéquate pour le CUSUM.

Lorsque la statistique du LC-CUSUM est égale ou dépasse une limite de compétence h , alors H_0 est rejetée et le t LC-CUSUM indique que le processus a atteint un état de contrôle (le stagiaire a atteint la compétence).

Comme pour le CUSUM, le LC-CUSUM ne peut être négatif ce qui empêche l'accumulation de scores au-delà de 0 qui pourraient pénaliser le stagiaire. Le score reste ainsi réactif et augmentera dès que les performances du stagiaire s'amélioreront. Il n'aura donc pas besoin de compenser ses échecs précédents cumulés.

Le LC-CUSUM est la somme S_t de t procédures, dont la formule mathématique est :

$$S_t = \max(0, S_{t-1} + W_t)$$

Avec $t (= 1, 2, 3, \dots)$ le nombre de procédure, $S_0 = 0$ le score LC-CUSUM au début du processus, et W_t le score attribué à la t ème procédure, en fonction du succès ou non de l'intervention.

Comme pour le CUSUM, le choix de W_t est basé sur le rapport de log-vraisemblance. Pour une variable binaire et pour la procédure t , si en cas de succès $Y_t = 0$ et si en cas d'échec $Y_t = 1$ alors nous avons comme score LC-CUSUM :

$$W_t = \log [(1 - p_0) / (1 - p_0 - \delta)] \text{ if } Y_t = 0 \text{ (succès)}$$

$$W_t = \log [p_0 / (p_0 + \delta)] \text{ if } Y_t = 1 \text{ (échec)}$$

Le CUSUM et le LC-CUSUM dépendent uniquement de ce qui est considéré comme un taux d'échec acceptable et inacceptable. Ils ne sont pas influencés par les caractéristiques du patient. Toutefois, dans certains contextes, ce qui est considéré comme acceptable peut varier en fonction de ces caractéristiques.

Chapitre 9

Comparaison de courbes d'apprentissage à partir du LC-CUSUM

Comme nous l'avons observé dans la revue de la littérature présentée dans la partie précédente, aucune méthodologie n'existe afin de comparer deux courbes d'apprentissages en dehors d'une analyse visuelle subjective. La dernière partie de cette thèse s'intéresse au développement d'un test statistique permettant de comparer aux moins deux courbes d'apprentissage l'une à l'autre.

Il nous semble nécessaire d'avoir cet outil statistique dans l'arsenal d'évaluation des actes chirurgicaux. De plus en plus, d'innovations en chirurgie prothétique font appel aux nouvelles technologies et l'apprentissage de ces techniques nous semble un point important afin d'évaluer la balance bénéfique/risque. En effet, si une nouvelle technique chirurgicale permet d'améliorer la précision, mais qu'en contrepartie pour la maîtriser et être reproductible dans le temps, il faut réaliser un minimum d'une centaine de procédure, alors la balance bénéfique/risque n'est peut-être pas en faveur de sa diffusion.

La première partie de cet article s'attache à présenter le principe des tests statistiques développés. La deuxième partie présente les performances des deux tests à l'aide de simulations. Enfin, la dernière partie est une application des tests sur un jeu de données.

Développement d'un test statistique permettant de comparer deux courbes d'apprentissage à partir du LC-CUSUM :

Pierre-Alban Bouché^{1,2} (resident), Rémy Nizard² (MD PHD), Matthieu Resche-Rigon^{1,3} (MD PHD)

1 Université de Paris, ECSTRAA Team, UMR U1153, INSERM, Paris, France

2 Service de chirurgie orthopédique et traumatologique, Hôpital Lariboisière, APHP, Paris, France

3 Service de Biostatistique et Information Médicale, Hôpital Saint-Louis, APHP, Paris, France

Corresponding author:

M. Bouché

INSERM U1153

Hôpital Saint-Louis

Bâtiment Lailler 3ème étage

1, avenue Claude Vellefaux

75010 Paris, France

Email: pierrealban309@gmail.com

Pierre-Alban Bouché^{1,2} (resident) : pierrealban309@gmail.com

Matthieu Resche-Rigon¹ (MD PHD) : matthieu.resche-rigon@univ-paris-diderot.fr

Rémy Nizard² (MD PHD) : remy.nizard@aphp.fr

Introduction :

La chirurgie prothétique et notamment du genou a beaucoup évolué ces dernières années. À la fin du vingtième siècle, les grands principes de pose des prothèses totales de genou ont été définis ainsi que le cahier des charges des différents implants. L'innovation actuelle se fait sur certains aspects techniques notamment avec l'apport de technologies : chirurgie assistée par ordinateur, chirurgie robotisée, utilisation de lunettes de réalité virtuelle, etc... (Banerjee et al., 2015; Deep et al., 2017; Gregory et al., 2018).

Le développement de nouvelles technologies dans le domaine médical et notamment prothétique est toujours attrayant autant pour le praticien que pour le patient. Malgré la pression parfois commerciale et des effets de mode, il est toujours important d'évaluer la balance bénéfique/risque des innovations et leur apport réel pour le chirurgien et surtout pour le patient.

Un des éléments à évaluer lors de la diffusion d'une nouvelle technique est sa courbe d'apprentissage. Les courbes d'apprentissage sont une représentation graphique de l'évolution de la performance au cours du temps. Elles permettent de voir l'évolution des performances mais aussi d'observer une constance dans l'apprentissage. La médecine s'est intéressée, en particulier dans les années 1980, aux courbes d'apprentissage afin de déterminer l'efficacité d'une procédure et le temps nécessaire pour acquérir les nouvelles techniques par les étudiants (Ramsay et al., 2002; Valsamis et al., 2018).

À la suite d'une revue de la littérature, six méthodologies différentes pour les modéliser en orthopédie ont été identifiées : comparaison des premiers patients opérés par un chirurgien avec les suivants méthodes, comparaison des actes pratiqués par des chirurgiens expérimentés à ceux réalisés par des débutants, le CUSUM test, le LC-CUSUM test, une approche utilisant les analyses de survie et la méthode basée sur les régressions linéaires. Nous avons aussi observé qu'aucun test statistique n'existe permettant de comparer deux courbes d'apprentissages.

L'objectif de cette étude était de développer ce test et de présenter des premiers résultats.

La première partie de cet article s'attache à présenter le principe de deux tests statistiques développer. La deuxième partie présente les performances de ces deux tests à l'aide de simulations. Enfin, la dernière partie est une application des tests sur un jeu de données.

Présentation des tests :

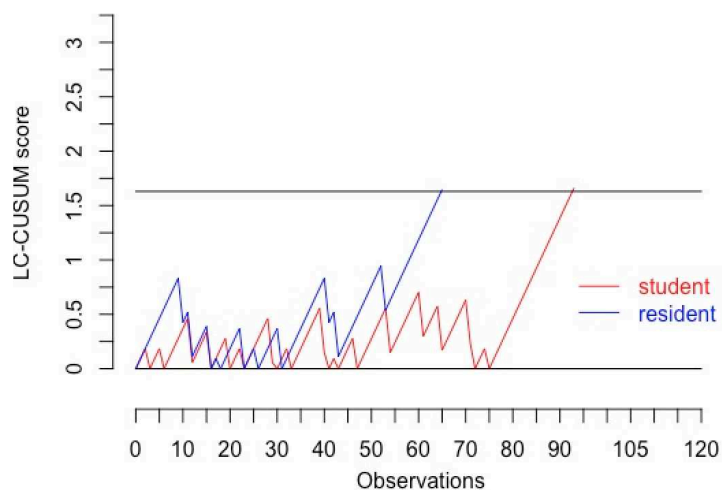
Parmi les différentes méthodologies retrouvées dans la littérature afin de modéliser les courbes d'apprentissages en orthopédie, le LC-CUSUM test semble être la méthodologie la plus simple, la plus flexible et la plus robuste. Afin de développer ce test statistique, nous nous sommes basés sur cette méthodologie.

Le test LC-CUSUM suppose que l'étudiant n'est pas compétent au début du suivi. Cette méthode est dérivée du CUSUM test (Steiner et al., 2001). Le CUSUM-test a été développé et appliqué dans l'industrie initialement. Il permet de contrôler des procédures et de détecter les écarts répétés par rapport à un objectif dans un processus contrôlé. Par exemple, pour entreprise fabriquant des vis, le CUSUM-test permettra de détecter lorsque les vis produites s'écartent du modèle standard. Il permettra, ainsi, à l'entreprise de paramétrer de nouveau la machine pour produire des vis standards.

En médecine et plus particulièrement en chirurgie une telle approche permet de monitorer un taux de complications ou la capacité à bien aligner une prothèse. Cette méthodologie est peu adaptée à un processus non contrôlé avec un taux d'échec élevé comme peut l'être une phase d'apprentissage.

Une version modifiée a été développée pour permettre le suivi des courbes d'apprentissage : le LC-CUSUM (Biau and Porcher, 2010). Cette méthode a les mêmes propriétés que le CUSUM mais ici, la limite est franchie lorsque l'étudiant a acquis les connaissances requises, c'est-à-dire lorsqu'il est jugé compétent (Fig 1). Le LC-CUSUM suppose que l'étudiant n'est pas compétent au début du suivi et signale lorsqu'il atteint le niveau de performance prédéfini acceptable.

Figure 1 : Courbes d'apprentissage de deux étudiants utilisant le LC-CUSUM. En abscisse, le nombre de procédures. En ordonnée, les valeurs du LC-CUSUM test. Ici, la limite à dépasser, pour qu'un étudiant soit compétent, était de 1,6. L'étudiant représenté par la courbe bleue devient compétent à la 60ème procédure. L'étudiant représenté par la courbe rouge devient compétent à la 92ème procédure.



Nous voyons, sur cet exemple, que l'étudiant représenté par la courbe bleue devient compétent à la 60ème procédure alors que l'étudiant représenté par la courbe rouge devient compétent à la 92ème procédure.

Pour comparer ces deux méthodes d'apprentissages ou 2 groupes de professionnels, nous avons envisagé deux possibilités. La première consistait à considérer le nombre moyen de procédures nécessaire pour être performant et de le comparer en utilisant un test de Wilcoxon (Wilcoxon, 1945) ou de Student (Zabell, 2008). La deuxième technique consistait à prendre en compte l'aire sous la courbe du LC-CUSUM avant la sortie et de comparer la moyenne des aires sous la courbe des deux techniques à l'aide d'un test de Wilcoxon ou de Student.

Simulations :

Afin de voir lequel des deux tests statistiques étaient le plus performant, nous avons réalisé des simulations en utilisant des tests de Wilcoxon et des tests de Student.

Les simulations consistaient à comparer différents groupes de praticiens effectuant des scénarios d'apprentissages différents. Chaque groupe était composé de 20 praticiens. Nous avons effectué trois séries de simulations pour faire varier le nombre de procédures par praticiens et les performances du LC-CUSUM.

Pour chaque série de simulation, dans un premier temps, nous avons comparé le premier scénario à lui-même afin de vérifier la valeur nominale du risque alpha. 7300 simulations ont été effectuées de façon à estimer ce risque avec une précision de l'ordre du pourcent. Nous avons choisi délibérément ce nombre de simulations car cela permettait d'obtenir une distribution d'une loi normal centrée. Nous avons, ensuite, simulé 200 fois cinq autres scénarios d'apprentissages pour les comparer au premier. Ces comparaisons permettaient d'évaluer la puissance de chaque test dans différentes situations.

Au début de chaque série de simulations sont présentés les paramètres des différents scénarios simulés, leur représentation graphique et les performances du LC-CUSUM test choisi. Le rang moyen de sortie, l'aire moyenne sous la courbe de chaque scénario, le risque alpha et la puissance de chaque test sont recueillis dans un tableau.

Les paramètres nécessaires d'être définis pour chaque série de simulation de scénarios pour le LC-CUSUM comprenaient :

- Le nombre moyen de procédures par praticien pour affirmer la compétence -> n
- Le taux d'échec acceptable -> p_1
- Le taux d'échec inacceptable -> p_0
- L'équivalence (écart acceptable par rapport à p_1) -> δ

Avec ces paramètres, nous avons déterminé la limite « h » et les performances correspondantes du LC-CUSUM, c'est-à-dire le taux de faux positifs (TFP) et le taux de vrais positifs (TVP). La limite h correspond au seuil du LC-CUSUM score à partir duquel le praticien est considéré comme compétent.

Le taux de faux positifs (TFP) est la probabilité que la limite soit franchie sur n procédures si la performance réelle du praticien est de p_1 d'échec (taux d'échec inacceptable). Cela correspond à l'erreur de type 1 dans les essais contrôlés randomisés.

Le taux de vrai positifs (TVP) est la probabilité que la limite h soit franchie sur n procédures si la performance réelle du praticien est de p_0 d'échec (taux d'échec acceptable). Cela correspond à la puissance dans les essais contrôlés randomisés

Le choix de la limite « h » est totalement arbitraire. Il n'est pas important que la limite h soit haute ou basse. Ce qui compte vraiment, ce sont les performances (TFP et TVP) du test.

Résultats simulations :

1. Première simulation :

a. Paramètres :

- $n < -100$
- $p_0 < -0.4$
- $p_1 < -0.1$
- $\delta < -0.075$
- $h < -1.64$

Nous avons considéré $p_1 = 10\%$ comme critère acceptable pour un pourcentage de procédures non-réussis. Nous avons considéré $p_0 = 40\%$ comme critère non-acceptable pour un taux de procédures non-réussis. Un delta de 0,075, écart acceptable par rapport aux performances adéquates, a été choisi. Sur la base de ces paramètres et en utilisant la simulation informatique (10000 simulations de séries de 100 procédures), une limite de h de 1,64 a été choisie pour le test LC-CUSUM afin de fournir les taux de découverte d'alarme suivants sur 100 procédures: taux de vrai positifs (TVP; ici la probabilité qu'une alarme soient déclenchée sur 100 procédures si la performance réelle de l'étudiant est de 10% d'échec, c'est-à-dire que la procédure est maîtrisée) de 95% et un taux de faux positifs (TFP; ici la probabilité qu'une alarme soit déclenchée sur 100 procédures si la performance de l'étudiant est de 40% d'échec, c'est-à-dire que la procédure n'est pas maîtrisée) de 1,5%.

b. Risque de première espèce :

Pour le risque alpha, nous avons comparé le premier scénario à lui-même afin de vérifier la valeur nominale du risque alpha. Le premier scénario est défini par :

- Performance de départ ($Perf_0$) = 0,4
- Performance de fin ($Perf_1$) = 0,1
- Procédure à partir de laquelle la performance commence à varier (N_{proc0}) = 10
- Procédure à partir de laquelle la performance ne varie plus (N_{proc1}) = 50
- Nombre de procédures totale effectuées (N_{br}) = 100

Tableau 1 : performance des tests pour le risque alpha. *E.T:écart type.*

Moyenne _{aire} + ET _{aire}	Moyenne _{sorties} + ET _{sorties}	Risque α_{wilcoxon} Aire	Risque α_{wilcoxon} sortie	Risque α_{student} Aire	Risque α_{student} sortie
32.807+10.218	75.546+17.500	0.045	0.045	0.045	0.045

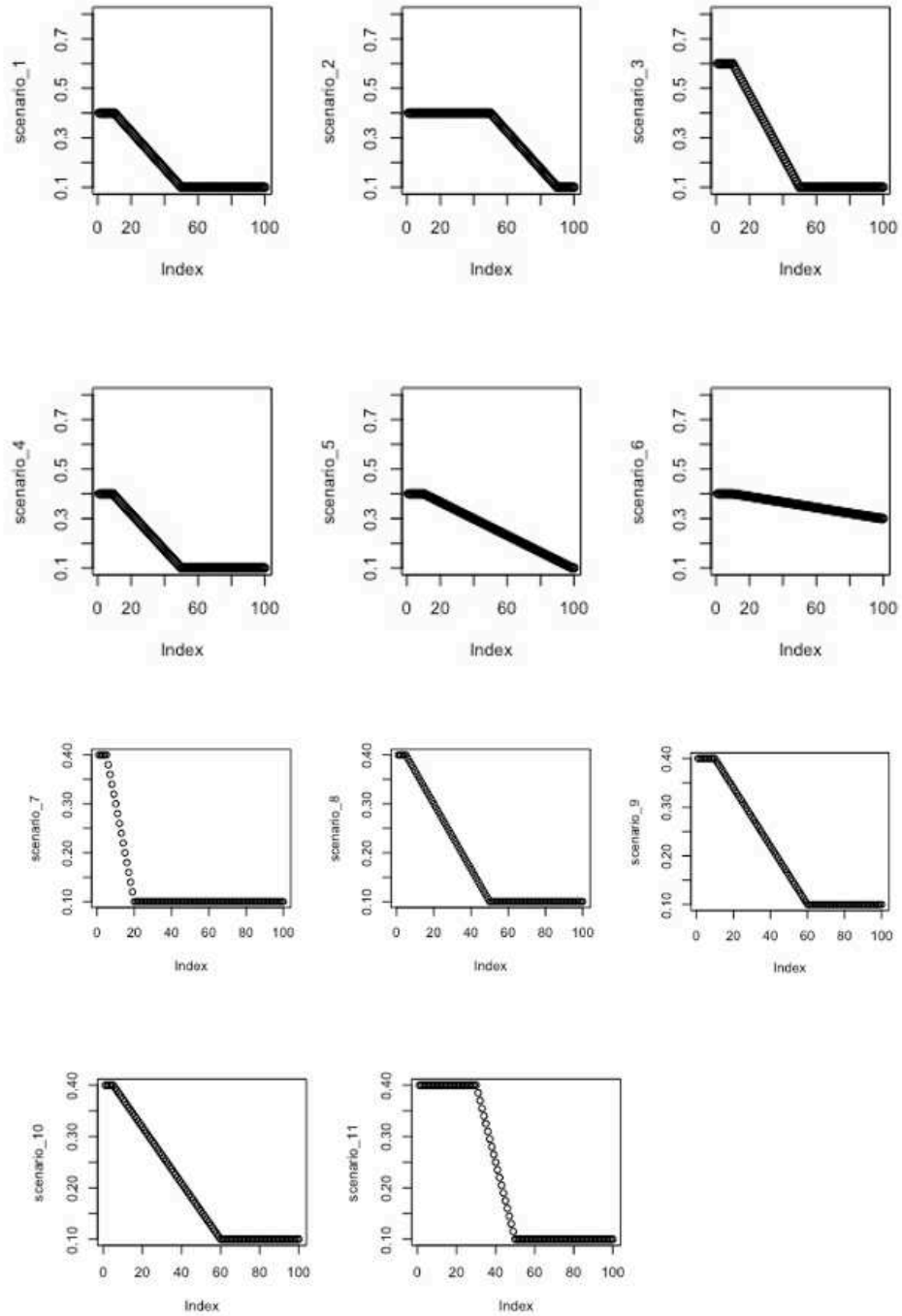
c. Différents scénarii :

Les différents scénarii diffèrent soit par leur performance de départ (Perf0), la performance de fin (Perf1), la procédure à partir de laquelle la performance commence à varier (Nproc0), la procédure à partir de laquelle la performance ne varie plus (Nproc1) et le nombre de procédures totale effectuées (Nbr). Le détail des différents scénarii est disponible dans le tableau suivant.

Tableau 2 : détails des différents scénarii

Scénario	Perf0	Perf1	N _{proc0}	N _{proc1}	Nbr
1	0,4	0,1	10	50	100
2	0,4	0,1	50	90	100
3	0,6	0,1	10	50	100
4	0,4	0,1	9	50	100
5	0,4	0,1	10	99	100
6	0,4	0,3	10	99	100
7	0,4	0,1	5	20	100
8	0,4	0,1	5	50	100
9	0,4	0,1	10	60	100
10	0,4	0,1	5	60	100
11	0,4	0,1	30	50	100

Figure 2 : représentations des différents scénarii



d. Résultats :Tableau 3 : performances des tests des différents scénarii. *E.T:écart type.*

Comparaisons	Moyenne _{aire 1}	Moyenne _{aire 2}	Moyenne _{sorties 1}	Moyenne _{sorties 2}	Puissance _{wilcoxon}	Puissance _{wilcoxon}	Puissance _{student}	Puissance _{student}
	+ ET _{aire 1}	+ ET _{aire 2}	+ ET _{sorties 1}	+ ET _{sorties 2}	Aire	Sortie	Aire	Sortie
1 Vs 2	32.807 (10.218)	24.273 (6.616)	75.546 (17.500)	98.379 (4.578)	1	1	1	1
1 Vs 3	32.807 (10.218)	30.270 (9.350)	75.546 (17.500)	78.856 (15.818)	0.58	0.54	0.61	0.545
1 Vs 4	32.807 (10.218)	32.901 (10.191)	75.546 (17.500)	75.445 (17.554)	0.07	0.04	0.075	0.04
1 Vs 5	32.807 (10.218)	28.723 (8.111)	75.546 (17.500)	96.297 (8.855)	0.98	1	0.99	1
1 Vs 6	32.807 (10.218)	16.584 (5.131)	75.546 (17.500)	99.726 (2.725)	1	1	1	1
1 Vs 7	32.807 (10.218)	30.415 (12.677)	75.546 (17.500)	55.082 (21.763)	0.91	1	0.58	1
1 Vs 8	32.807 (10.218)	33.312 (10.430)	75.546 (17.500)	74.669 (18.051)	0.075	0.09	0.085	0.105
1 Vs 9	32.807 (10.218)	33.272 (9.448)	75.546 (17.500)	81.367 (9.448)	0.14	0.95	0.065	0.94
1 Vs 10	32.807 (10.218)	33.707 (9.667)	75.546 (17.500)	80.551 (16.316)	0.21	0.89	0.165	0.865
1 Vs 11	32.807 (10.218)	31.310 (9.228)	75.546 (17.500)	79.715 (15.475)	0.255	0.67	0.30	0.725

2. Deuxième simulation :

a. Paramètres :

La principale modification de cette série est sur le taux d'échec acceptable (p_1) qui a été augmenté de 10 à 20%.

- $n < 100$
- $p_0 < 0.4$
- $p_1 < 0.2$
- $\delta < 0.075$
- $h < 1.36$

Nous avons considéré $p_1 = 20\%$ comme critère acceptable pour un pourcentage de procédures non-réussis. Nous avons considéré $p_0 = 40\%$ comme critère non-acceptable pour un taux de procédures non-réussis. Un δ de 0,075, écart acceptable par rapport aux performances adéquates, a été choisi. Sur la base de ces paramètres et en utilisant la simulation informatique (10000 simulations de séries de 100 procédures), une limite de h de 1,36 a été choisie pour le test LC-CUSUM afin de fournir les taux de découverte d'alarme suivants sur 100 procédures: taux de vrais positifs (TVP; ici la probabilité qu'une alarme soit déclenchée sur 100 procédures si la performance réelle de l'étudiant est de 10% d'échec, c'est-à-dire que la procédure est maîtrisée) de 91% et un taux de faux positifs (TFP; ici la probabilité qu'une alarme soit déclenchée sur 100 procédures si

le la performance de l'étudiant est de 40% d'échec, c'est-à-dire que la procédure n'est pas maîtrisée) de 5%.

b. Risque de première espèce :

Pour le risque alpha, nous avons comparé le premier scénario à lui-même afin de vérifier la valeur nominale du risque alpha. Le premier scénario est défini par :

- Performance de départ (Perf0) = 0,4
- Performance de fin (Perf1) = 0,1
- Procédure à partir de laquelle la performance commence à varier (Nproc0) = 10
- Procédure à partir de laquelle la performance ne varie plus (Nproc1) = 50
- Nombre de procédures totale effectuées (Nbr)= 100

Tableau 4 : performance des tests pour le risque alpha. *E.T:écart type.*

Moyenne _{aire} + ET _{aire}	Moyenne _{sorties} + ET _{sorties}	Risque $\alpha_{wilcoxon}$ Aire	Risque $\alpha_{wilcoxon}$ sortie	Risque $\alpha_{student}$ Aire	Risque $\alpha_{student}$ sor
24.761+6.884	89.708+17.240	0.04	0.035	0.045	0.05

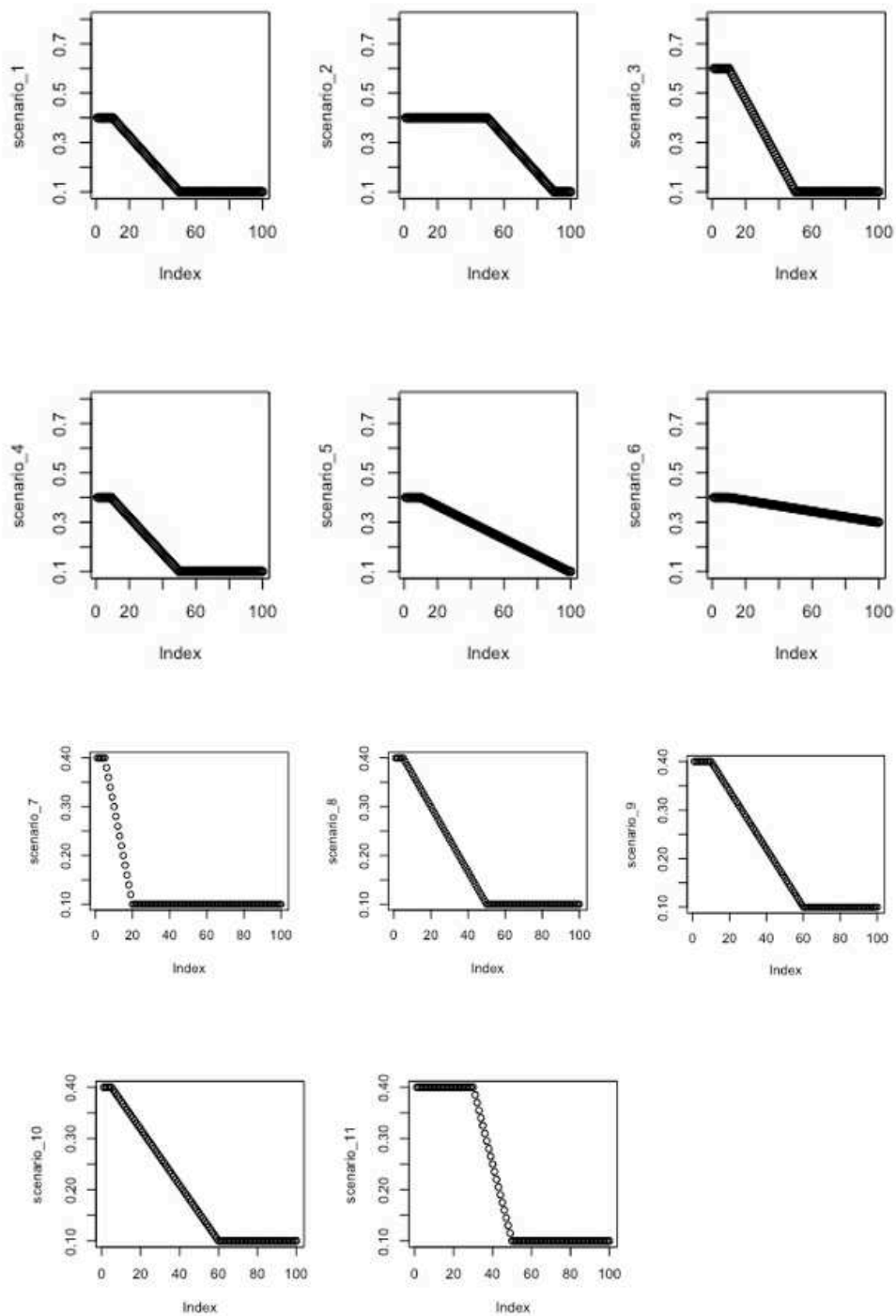
c. Différents scénarii :

Les différents scénarii diffèrent soit par leur performance de départ (Perf0), la performance de fin (Perf1), la procédure à partir de laquelle la performance commence à varier (Nproc0), la procédure à partir de laquelle la performance ne varie plus (Nproc1) et le nombre de procédures totale effectuées (Nbr). Le détail des différents scénarii est disponible dans le tableau suivant.

Tableau 5 : détails des différents scénarii

Scénario	Perf0	Perf1	N _{proc0}	N _{proc1}	Nbr
1	0,4	0,1	10	50	100
2	0,4	0,1	50	90	100
3	0,6	0,1	10	50	100
4	0,4	0,1	9	50	100
5	0,4	0,1	10	99	100
6	0,4	0,3	10	99	100
7	0,4	0,1	5	20	100
8	0,4	0,1	5	50	100
9	0,4	0,1	10	60	100
10	0,4	0,1	5	60	100
11	0,4	0,1	30	50	100

Figure 3 : représentations des différents scénarii



d. Résultats :

Tableau 6 : performances des tests des différents scénarii. *E.T:écart type.*

Comparaisons	Moyenne _{aire 1} + ET _{aire 1}	Moyenne _{aire 2} + ET _{aire 2}	Moyenne _{sorties 1} + ET _{sorties 1}	Moyenne _{sorties 2} + ET _{sorties 2}	Puissance _{wilcoxon Aire}	Puissance _{wilcoxon Sortie}	Puissance _{student Aire}	Puissance _{student Sortie}
1 Vs 2	24.761 (6.884)	18.336 (5.080)	89.708 (17.240)	98.583 (6.201)	1	1	1	1
1 Vs 3	24.761 (6.884)	22.050 (6.167)	89.708 (17.240)	91.807 (14.296)	0.98	0.17	0.99	0.31
1 Vs 4	24.761 (6.884)	24.799 (6.831)	89.708 (17.240)	89.545 (17.341)	0.08	0.04	0.08	0.045
1 Vs 5	24.761 (6.884)	20.593 (5.718)	89.708 (17.240)	97.184 (9.609)	1	1	1	1
1 Vs 6	24.761 (6.884)	16.270 (4.613)	89.708 (17.240)	98.960 (6.458)	1	1	1	1
1 Vs 7	24.761 (6.884)	26.280 (8.577)	89.708 (17.240)	80.538 (25.506)	0.415	0.92	0.515	0.99
1 Vs 8	24.761 (6.884)	25.002 (6.989)	89.708 (17.240)	89.153 (17.855)	0.085	0.045	0.065	0.065
1 Vs 9	24.761 (6.884)	24.147 (6.547)	89.708 (17.240)	91.817 (15.164)	0.12	0.245	0.125	0.325
1 Vs 10	24.761 (6.884)	24.476 (6.703)	89.708 (17.240)	91.321 (15.747)	0.075	0.165	0.075	0.195
1 Vs 11	24.761 (6.884)	23.518 (6.409)	89.708 (17.240)	91.572 (14.704)	0.44	0.16	0.46	0.25

3. Troisième simulation :

a. Paramètres :

La principale modification de cette série est sur le nombre de procédures par praticiens qui est passé de 100 à 50.

- $n < 50$
- $p_0 < 0.4$
- $p_1 < 0.1$
- $\delta < 0.075$
- $h < 1.64$

Nous avons considéré $p_1 = 10\%$ comme critère acceptable pour un pourcentage de procédures non-réussis. Nous avons considéré $p_0 = 40\%$ comme critère non-acceptable pour un taux de procédures non-réussis. Un delta de 0,075, écart acceptable par rapport aux performances adéquates, a été choisi. Sur la base de ces paramètres et en utilisant la simulation informatique (10000 simulations de séries de 100 procédures), une limite de h de 1,64 a été choisie pour le test LC-CUSUM afin de fournir les taux de découverte d'alarme suivants sur 50 procédures:

taux de vrai positifs (TVP; ici la probabilité qu'une alarme soient déclenchée sur 50 procédures si la performance réelle de l'étudiant est de 10% d'échec, c'est-à-dire que la procédure est maîtrisée) de 95% et un taux de faux positifs (TFP; ici la probabilité qu'une alarme soit déclenchée sur 100 procédures si la performance de l'étudiant est de 40% d'échec, c'est-à-dire que la procédure n'est pas maîtrisée) de 1.5%.

b. Risque de première espèce :

Pour le risque alpha, nous avons comparé le premier scénario à lui-même afin de vérifier la valeur nominale du risque alpha. Le premier scénario est défini par :

- Performance de départ (Perf0) = 0,4
- Performance de fin (Perf1) = 0,1
- Procédure à partir de laquelle la performance commence à varier (Nproc0) = 10
- Procédure à partir de laquelle la performance ne varie plus (Nproc1) = 25
- Nombre de procédures totale effectuées (Nbr) = 50

Tableau 7 : performance des tests pour le risque alpha. *E.T:écart type.*

Moyenne _{aire} + ET _{aire}	Moyenne _{sorties} + ET _{sorties}	Risque $\alpha_{wilcoxon}$ Aire	Risque $\alpha_{wilcoxon}$ sortie	Risque $\alpha_{student}$ Aire	Risque $\alpha_{student}$ sortie
19.712+5.642	46.108+5.882	0.03	0.03	0.04	0.025

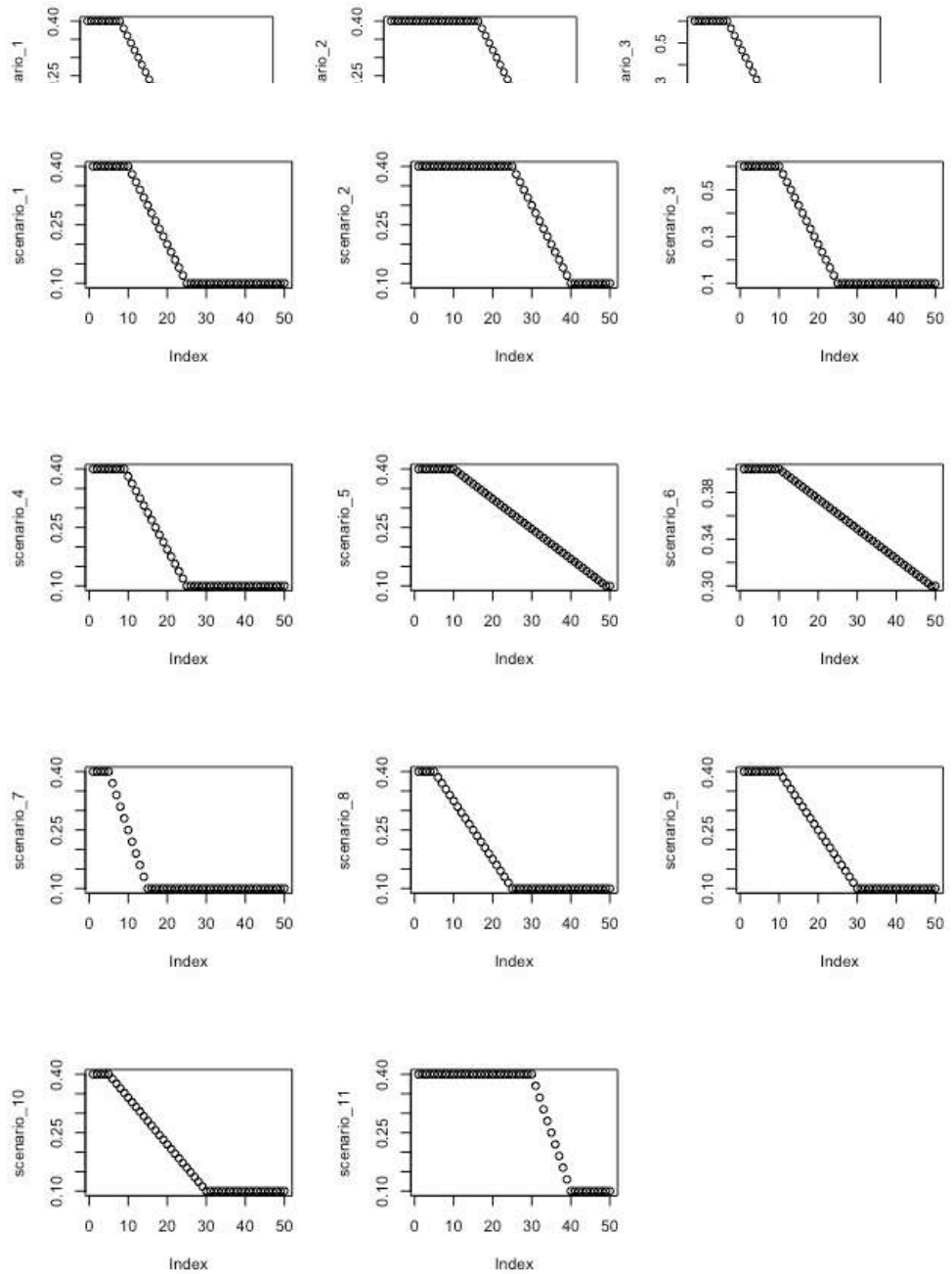
c. Différents scénarii :

Les différents scénarii diffèrent soit par leur performance de départ (Perf0), la performance de fin (Perf1), la procédure à partir de laquelle la performance commence à varier (Nproc0), la procédure à partir de laquelle la performance ne varie plus (Nproc1) et le nombre de procédures totale effectuées (Nbr). Le détail des différents scénarii est disponible dans le tableau suivant.

Tableau 8 : détails des différents scénarii

Scénario	Perf0	Perf1	N _{proc0}	N _{proc1}	Nbr
1	0,4	0,1	10	25	50
2	0,4	0,1	25	40	50
3	0,6	0,1	10	25	50
4	0,4	0,1	9	25	50
5	0,4	0,1	10	49	50
6	0,4	0,3	10	49	50
7	0,4	0,1	5	15	50
8	0,4	0,1	5	25	50
9	0,4	0,1	10	30	50
10	0,4	0,1	5	30	50
11	0,4	0,1	30	40	50

Figure 4 : représentations des différents scénarii



d. Résultats :

Tableau 9 : performances des tests des différents scénarii. *E.T:écart type.*

Comparaisons	Moyenne _{aire 1} + ET _{aire 1}	Moyenne _{aire 2} + ET _{aire 2}	Moyenne _{sorties 1} + ET _{sorties 1}	Moyenne _{sorties 2} + ET _{sorties 2}	Puissance _{wilcoxon} Aire	Puissance _{wilcoxon} Sortie	Puissance _{student} Aire	Puissance _{student} Sortie
1 Vs 2	19.712 (5.642)	12.690 (4.636)	46.108 (5.882)	49.761 (1.168)	1	1	1	1
1 Vs 3	19.712 (5.642)	17.939 (5.314)	46.108 (5.882)	47.001 (4.850)	0.88	0.21	0.89	0.34
1 Vs 4	19.712 (5.642)	19.832 (5.670)	46.108 (5.882)	45.981 (6.044)	0.05	0.08	0.045	0.075
1 Vs 5	19.712 (5.642)	12.912 (5.167)	46.108 (5.882)	49.660 (1.768)	1	1	1	1
1 Vs 6	19.712 (5.642)	7.856 (3.302)	46.108 (5.882)	49.967 (0.436)	1	1	1	1
1 Vs 7	19.712 (5.642)	21.872 (6.251)	46.108 (5.882)	42.067 (8.944)	0.835	1	0.99	1
1 Vs 8	19.712 (5.642)	20.376 (5.782)	46.108 (5.882)	45.480 (6.549)	0.205	0.16	0.225	0.195
1 Vs 9	19.712 (5.642)	18.514 (5.539)	46.108 (5.882)	47.398 (4.707)	0.44	0.53	0.535	0.66
1 Vs 10	19.712 (5.642)	19.321 (5.750)	46.108 (5.882)	46.904 (5.329)	0.09	0.18	0.10	0.24
1 Vs 11	19.712 (5.642)	11.658 (4.270)	46.108 (5.882)	49.898 (0.701)	1	1	1	1

4. Résultats :

L'ensemble des risques de premières espèces retrouvés étaient proche de 5%.

a. Test comparant l'aire sous la courbe VS test comparant le rang de sortie :

Pour le test de Wilcoxon, la puissance est similaire entre le test comparant l'aire sous la courbe et celui comparant le rang de sortie dans 30% des cas (9/30). Dans c'est neuf cas, la puissance était de 100%. La puissance du test comparant l'aire sous la courbe était supérieure à celle du test comparant le rang de sortie dans 27% (8/30). La puissance du test comparant l'aire sous la courbe était inférieure à celle du test comparant le rang de sortie dans 43% (13/30). Au total, la puissance du test de Student est supérieure ou égale à celle du test de Wilcoxon dans 73% (22/30).

Pour le test de student, la puissance est similaire entre le test comparant l'aire sous la courbe et celui comparant le rang de sortie dans 34% des cas (10/30). Dans c'est dix cas, la puissance était de 100% dans 9/10. La puissance du test comparant l'aire sous la courbe était supérieure à celle du test comparant le rang de sortie dans 23% (7/30). La puissance du test comparant l'aire

sous la courbe était inférieure à celle du test comparant le rang de sortie dans 43% (13/30). Au total, la puissance du test comparant le rang de sortie est supérieure ou égale à celle du test comparant l'aire sous la courbe dans 77% (23/30).

b. Test de Student VS test de Wilcoxon :

Pour l'aire sous la courbe, la puissance est similaire entre le test de Student et le test de Wilcoxon dans 36% des cas (11/30). Dans c'est onze cas, la puissance était de 100% dans 9/11, soit 82%. La puissance du test de Student était supérieure à du test de Wilcoxon dans 47% (14/30). La puissance du test de Student était inférieure à celle du test de Wilcoxon dans 17% (5/30). Au total, la puissance du test de Student est supérieure ou égale à celle du test de Wilcoxon dans 83% (25/30).

Pour le rang de sortie, la puissance est similaire entre le test de Student et le test de Wilcoxon dans 43% des cas (13/30). Dans c'est treize cas, la puissance était de 100% dans 12/13, soit 92%. La puissance du test de Student était supérieure à du test de Wilcoxon dans 47% (14/30). La puissance du test de Student était inférieure à celle du test de Wilcoxon dans 10% (3/30). Au total, la puissance du test de Student est supérieure ou égale à celle du test de Wilcoxon dans 90% (27/30).

C. Interprétation :

Au vu des résultats des simulations, le test comparant le rang moyen de sortie des courbes de LC-CUSUM est plus performant que celui comparant l'aire moyenne sous la courbe et le test de Student plus que le test de Wilcoxon. Le test de Student comparant le rang moyen est le test qui sera utilisé dans note application sur un jeu de données réel.

Application :

Après avoir validé les performances du test statistique utilisant le rang moyen de sortie et le test de Student, nous devons voir s'il est applicable sur un jeu de données cliniques.

Le jeu de donnée est composé de réponses à des dossiers progressifs de tous les étudiants de l'uni-

versité de Paris ayant passés les Épreuves Classantes Nationales (ECNi) en 2020. Après avoir reçu l'accord scientifique et éthique de l'Université Numérique en Santé et Sport (UNESS), nous avons obtenu les données des entraînements aux dossiers progressifs en cardiologie et pneumologie sur SIDES de ces étudiants en DFASM 2 et 3. Nous n'avons considéré que les étudiants ayant répondu à un minimum de 100 questions pour chaque spécialité. Nous avons recueilli pour chaque étudiant : le classement final aux ECNi, la faculté d'origine (Paris 5 et Paris 7) et le taux de bonnes réponses pour chaque spécialité aux dossiers progressifs.

L'objectif était de comparer les méthodes d'apprentissages de deux groupes d'étudiants à l'aide du test statistique développé. Pour cela, nous avons comparé les courbes d'apprentissage des étudiants de l'université de Paris Descartes avec ceux de l'université Paris Diderot. Nous avons aussi comparé les courbes d'apprentissages des étudiants arrivés dans classés les 1000 premiers et ceux classés après.

D'après une étude d' E. rivièrè (Rivière et al., 2019), pour être classé dans les 1000 premiers il faut environ 70% de bonnes réponses (7640,62/10800) pour être classés dans les 1000 premiers.

Pour déterminer que la compétence a été atteinte, le LC-CUSUM a été appliqué aux différents étudiants. Le LC-CUSUM suppose que l'étudiant n'est pas compétent au début du suivi et signale lorsqu'il atteint le niveau de performance prédéfini acceptable.

Par conséquent, l'hypothèse nulle (H_0) est que le processus est hors de contrôle (la performance de l'étudiant s'écarte de la performance adéquate, p_0 , d'au moins une valeur delta), et l'hypothèse H_1 est que le processus est sous contrôle (la performance de l'étudiant est égale à la performance jugée adéquate). L'étudiant n'est pas considéré compétent tant que le score LC-CUSUM reste inférieur à la limite h . Lorsque le score LC-CUSUM dépasse cette limite, l'étudiant est considéré comme compétent. Nous avons considéré $p_0 = 30\%$ comme critère acceptable pour un pourcentage de questions non-réussites. Nous avons considéré $p_1 = 50\%$ comme un taux de questions non-réussites. Une valeur de delta de 0,05 a été choisie comme écart acceptable. Sur la base de ces paramètres et en utilisant la simulation informatique (10000 simulations de séries de 100 procédures), une limite de h de 1,15 a été choisie pour le test LC-CUSUM afin de fournir les taux de découverte d'alarme suivants sur 100 procédures : taux de vrai positifs (TVP; ici la probabilité qu'une alarme soit déclenchée sur 100 procédures si la performance réelle de l'étudiant est de 30% d'échec, c'est-à-dire que l'étudiant est compétent) de 94% et un taux de faux positifs (TFP; ici la probabilité qu'une alarme soit déclenchée sur 100 procédures si la performance de l'étudiant est

de 50% d'échec, c'est-à-dire que l'étudiant n'est pas compétent) de 7%.

En 2020, 854 étudiants de l'université de Paris ont passé l'ECNi en 2020.

Sept cent soixante-cinq étudiants ont répondu à au moins une question de dossier progressif de pneumologie pour préparer les ECNi 2020, soit 89,6% (765/854). La moyenne de questions de dossiers progressifs en pneumologie faites par étudiants était de 215,2 [1,0-3228,0]. Parmi ces étudiants, 482 (63,0%) étudiants avaient au moins fait 100 questions de dossiers progressifs de pneumologie. Le LC-CUSUM a permis de déterminer que 321 étudiants étaient compétents, soit 66,6%.

Sur les 482 étudiants, 290 étudiants venaient de Paris 5 (60%) et 192 de Paris 7 (40%). Le rang moyen de sortie des étudiants de Paris 5 était de 128 (min-max :11-1208) et de ceux de Paris 7 était de 114 (min-max :11-1238). Aucune différence entre les deux facultés n'a été observée, p.valeur=0.28.

Sur ces 482 étudiants, 72 étudiants étaient classés dans les 1000 (15%) et 410 après (85%). Le rang moyen de sortie des étudiants classés dans les 1000 était de 223 (min-max :12-1208) et des autres de 105 (min-max :11-1238). Une différence significative a été observée, p.valeur<0.01.

Sur 854 étudiants, 780 étudiants ont répondu à au moins une question de DP de cardiologie pour préparer l'ECNi 2020, soit 91,3%. La moyenne de questions de dossiers progressifs en cardiologie faites par étudiants était de 258,6 [1,0-4924,0]. Parmi ces étudiants, 494 (63,3%) étudiants avaient au moins fait 100 questions de dossiers progressifs en cardiologie. Le LC-CUSUM a permis de déterminer que 359 étudiants étaient compétents, soit 72,7%.

Sur les 494 étudiants, 271 étudiants venaient de Paris 5 (55%) et 223 de Paris 7 (45%). Le rang moyen de sortie des étudiants de Paris 5 était de 136 (min-max :11-1891) et de ceux de Paris 7 était de 141 (min-max :11-1036). Aucune différence entre les deux facultés n'a été observée, p.valeur=0.72.

Sur ces 494 étudiants, 71 étudiants étaient classés dans les 1000 (14%) et 423 après (86%). Le rang moyen de sortie des étudiants classés dans les 1000 était de 264 (min-max :15-1891) et des autres de 117 (min-max :11-1012). Une différence significative a été observée, p.valeur<0.01.

Discussions :

Après fait une revue de la littérature sur les courbes d'apprentissages, aucun test statistique comparant deux méthodes d'apprentissage n'a été retrouvé.

Ces dernières années, le développement des techniques mini-invasives et l'arrivée de nouvelles technologies, notamment la robotique, ont vu se développer l'utilisation des courbes d'apprentissage comme critère d'évaluation. En effet, en recherchant sur Pubmed « learning curve », en 2000, 204 articles ont été publiés contre 2718 en 2019. Le nombre d'articles traitant des courbes d'apprentissages a été multiplié par 10.

Partant de ces différentes constatations, le développement d'un test statistique permettant de faire des comparaisons de courbes d'apprentissages nous paraissait indispensable.

Pour débiter, nous avons dû choisir quelle méthodologie de modélisation des courbes d'apprentissage utiliser. Nous avons opté pour le LC-CUSUM. Cette méthode semble être la plus simple, la plus flexible et la plus robuste pour modéliser une courbe d'apprentissage. Cette méthodologie peut être appliquée quel que soit le type d'acquisition et le type de technique d'intérêt. De plus, elle est facile à diffuser et peut être utilisée dans plusieurs spécialités (Dessolle et al., 2010; Weil et al., 2017).

Nous avons, ensuite imaginé, deux manières de pouvoir comparer les courbes d'apprentissages : soit en comparant l'aire sous la courbe ou soit en comparant le rang moyen de sortie. D'après les résultats des simulations, le test comparant le rang moyen de sortie associé à l'utilisation du test de Student nous semble le puissant. Néanmoins, l'ensemble des comparaisons des différents scénarios ne vont pas toute en faveur de ce test. En effet, la puissance du test comparant l'aire sous la courbe était supérieure à celle du test comparant le rang de sortie dans 27% cas et dans 23%, respectivement en utilisant le test de Wilcoxon et de Student. Une série de simulation plus importante semble être indispensable afin de valider ces observations.

Nous l'avons ensuite appliqué sur un jeu de données réelles. Cette application pratique nous a permis de constater que le test était performant pour détecter des différences lorsqu'elles existaient. Toutefois, l'interprétation des résultats semble incohérente. En effet, d'après notre analyse, les étudiants arrivés après les 1000 premières places auraient des courbes d'apprentissages plus courtes que les étudiants arrivés dans les 1000 premiers pour les dossiers progressifs de cardiologie et de pneumologie. La première explication est que le classement aux ECNi n'est pas forcément corrélé aux performances des étudiants en cardiologie et en pneumologie. Deuxièmement, dans cette situation, un étudiant qui a une courbe d'apprentissage plus courte qu'un autre n'est pas le garant de

réussite aux ECNi. Enfin, nous n'avons pas évalué, si une fois que l'étudiant est jugé compétent par le LC-CUSUM, ses performances restent stables dans le temps.

Toutefois, sur les deux autres comparaisons effectuées, le test ne retrouve pas de différence significative. Cette analyse semble cohérente car les étudiants proviennent de la même faculté de médecine (université de Paris) mais les enseignements ne se passent pas dans les mêmes locaux. Cette application pratique aura permis de constater que le test développer semble être performant. Afin de le valider de manière définitive, une étude avec des effectifs plus importants et une méthodologie plus rigoureuse doit être effectuée.

Le LC-CUSUM ne s'intéresse, pour le moment, à un seul paramètre. Par exemple, est-ce que la coupe osseuse a bien été effectuée comme la planification : oui ou non ? La réussite d'une intervention chirurgicale ne dépend pas que d'une seule variable. Il nous semble nécessaire de développer un LC-CUSUM test multidimensionnel afin d'obtenir un outil d'évaluation de courbe d'apprentissage en chirurgie orthopédique fiable et complet.

Conclusion :

En se basant sur la modélisation des courbes d'apprentissages par le LC CUSUM, nous avons développé un test statistique performant permettant de comparer deux courbes d'apprentissage. Ce test utilise le rang moyen de sortie associé à un test de Student. Toutefois sa validation à travers une étude clinique puissante est nécessaire.

Bibliographies :

Banerjee, S., Cherian, J.J., Elmallah, R.K., Jauregui, J.J., Pierce, T.P., Mont, M.A., 2015. Robotic-assisted knee arthroplasty. *Expert Review of Medical Devices* 12, 727–735. <https://doi.org/10.1586/17434440.2015.1086264>

Biau, D.J., Porcher, R., 2010. A method for monitoring a process from an out of control to an in control state: Application to the learning curve. *Stat Med* 29, 1900–1909. <https://doi.org/10.1002/sim.3947>

Deep, K., Shankar, S., Mahendra, A., 2017. Computer assisted navigation in total knee and hip arthroplasty. *SICOT-J* 3, 50. <https://doi.org/10.1051/sicotj/2017034>

Delacour, H., Servonnet, A., Perrot, A., Vigezzi, J., Ramirez, J., 2005. La courbe ROC (receiver operating characteristic) : principes et principales applications en biologie clinique. *undefined*.

Dessolle, L., Fréour, T., Barrière, P., Jean, M., Ravel, C., Darai, E., Biau, D.J., 2010. How soon can I be

proficient in embryo transfer? Lessons from the cumulative summation test for learning curve (LC-CUSUM). *Hum Reprod* 25, 380–386. <https://doi.org/10.1093/humrep/dep391>

Gregory, T.M., Gregory, J., Sledge, J., Allard, R., Mir, O., 2018. Surgery guided by mixed reality: presentation of a proof of concept. *Acta Orthop* 89, 480–483. <https://doi.org/10.1080/17453674.2018.1506974>

Ramsay, C.R., Wallace, S.A., Garthwaite, P.H., Monk, A.F., Russell, I.T., Grant, A.M., 2002. Assessing the learning curve effect in health technologies. Lessons from the nonclinical literature. *Int J Technol Assess Health Care* 18, 1–10.

Rivière, E., Quinton, A., Dehail, P., 2019. Analyse de la discrimination des notes aux ECN informatisées de 2016. *La Revue de Médecine Interne* 40, 286–290. <https://doi.org/10.1016/j.revmed.2018.10.386>

Steiner, S.H., Cook, R.J., Farewell, V.T., 2001. Risk-adjusted monitoring of binary surgical outcomes. *Med Decis Making* 21, 163–169. <https://doi.org/10.1177/0272989X0102100301>

Valsamis, E.M., Chouari, T., O'Dowd-Booth, C., Rogers, B., Ricketts, D., 2018. Learning curves in surgery: variables, analysis and applications. *Postgraduate Medical Journal* 94, 525–530. <https://doi.org/10.1136/postgradmedj-2018-135880>

Weil, G., Motamed, C., Biau, D.J., Guye, M.L., 2017. Learning curves for three specific procedures by anesthesiology residents using the learning curve cumulative sum (LC-CUSUM) test. *Korean J Anesthesiol* 70, 196–202. <https://doi.org/10.4097/kjae.2017.70.2.196>

Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. *Biometrics Bulletin* 1, 80–83. <https://doi.org/10.2307/3001968>

Zabell, S.L., 2008. On Student's 1908 Article "The Probable Error of a Mean." *Journal of the American Statistical Association* 103, 1–7. <https://doi.org/10.1198/016214508000000030>

Conclusion

Dans la seconde partie de ce travail, nous nous sommes intéressés aux courbes d'apprentissages en orthopédie. À la fin des deux méta-analyses en réseau effectuées, nous nous sommes posés la question de l'apprentissage de ces nouvelles techniques. En effet, ce paramètre peut jouer un rôle dans l'interprétation des résultats et dans le choix à effectuer par le chirurgien.

Notre première démarche a été de faire une revue de la littérature internationale afin de voir comment cette variable était traitée et analysée. Nous avons fait, délibérément, le choix de nous limiter à la chirurgie orthopédique et traumatologique afin d'être homogène et de pouvoir disposer d'une méthode fiable dans cette spécialité. En effet, l'évaluation d'une procédure chirurgicale en orthopédie est différente d'une évaluation de l'apprentissage d'une procédure en radiologie interventionnelle par exemple.

Notre première remarque, au début de cette revue de la littérature, était que cette thématique, en recherche clinique, faisait l'objet de plus en plus d'études. À l'aide d'une extraction Pubmed sur les 20 dernières années, le nombre d'études mentionnant dans leur titre ou leur résumé « learning curve » a explosé : de 155 en 1999 à 537 en 2009 et 951 en 2019. Soit une augmentation de 6,3 fois en 20 ans du nombre d'études traitant de ce sujet.

Notre deuxième remarque était qu'il n'existait pas de recommandations particulières sur la méthodologie à utiliser pour traiter les courbes d'apprentissages en recherche clinique contrairement à d'autres types d'études. En effet, il existe des recommandations pour les essais cliniques randomisés (CONSORT (Schulz et al., 2010)), les études observationnelles (STROBE (von Elm et al., 2014)) et les méta-analyses (PRISMA (Page et al., 2021)).

Notre revue de la littérature a permis d'identifier 6 méthodologies différentes afin d'aborder la question des courbes d'apprentissages.

La première méthode et la plus rencontrée dans la littérature est celle qui consiste à comparer les premiers patients opérés par un chirurgien avec les suivants. Le seuil pour définir les « premiers patients » est généralement choisi arbitrairement. Cette méthode est séduisante car elle est facile

à mettre en œuvre, reproductible et facilement compréhensible. Cependant, il ne prend pas en compte le fait qu'une courbe d'apprentissage n'est pas linéaire et peut être composée de différentes phases (Val-samis et al., 2018) .

Enfin notre dernière constatation est que notre recherche n'a pas permis d'identifier un test statistique permettant de comparer directement des méthodes d'apprentissages. Aucune méthodologie ne permet de détecter si un élève a acquis une compétence avant une autre avec une différence significative, ni de détecter si une technique avait une courbe d'apprentissage significativement plus rapide.

Dans la dernière partie de ce travail, nous avons donc essayé de mettre en place un test statistique permettant de comparer deux courbes d'apprentissages en nous basant sur le LC-CUSUM test.

Nous avons comparé, à l'aide de simulation, deux techniques de test. Une prenant en compte le nombre de procédure nécessaire pour être performant et de comparer la moyenne des rangs de sortie en utilisant soit un test de Wilcoxon (Wilcoxon, 1945) ou un test de Student (Zabell, 2008). La deuxième technique consistait à prendre en compte l'air sous la courbe du LC-CUSUM test d'une manière similaire aux courbes ROC (Delacour et al., 2005) et de comparer la moyenne des aires sous la courbe des deux techniques à l'aide soit d'un test de Wilcoxon soit d'un test de Student. Aux vues des simulations, le test prenant en compte le rang de sortie et utilisant le test de Student était le plus performant. Nous avons ensuite utilisé des données d'entraînements aux ECNi d'étudiants en médecine pour voir si le test était applicable. Quatre comparaisons ont été effectuées : Deux entre étudiants de deux facultés de médecine différente et deux entre étudiants arrivés classés dans les 1000 premiers et les autres. Le test statistique a permis de mettre en évidence une différence significative pour comparaison et une absence de différence dans les deux autres. Ce test statistique est un premier pas dans ce domaine et propose une première approche pour comparer des courbes d'apprentissage. Toutefois, afin de le valider définitivement, une étude de plus grande envergure est nécessaire pour limiter le biais et augmenter la puissance de celle-ci.

Perspectives

Comme énoncé précédemment, 20% des patients après une PTG ne sont pas entièrement satisfaits. Lors de la réalisation d'une prothèse totale de genou, nous avons vu dans la première partie de ce mémoire, que le chirurgien avait un choix important dans le type d'implants à utiliser et dans les techniques. Dans ce travail, nous nous sommes intéressés aux voies d'abord dans les PTG et aux guides de coupes comme possible moyen d'amélioration.

Concernant les voies d'abord, la voie subvastus permet une amélioration des mobilités dans les premiers mois mais sans retentissements sur les scores fonctionnels à moyen terme. La littérature sur ce sujet semble assez fournie et nous pensons avoir donné une conclusion à cette question : toutes les voies d'abord chirurgicales médiales du genou pour réaliser une PTG sont équivalentes et le choix revient au chirurgien en fonction de ses habitudes, de ses préférences et sa maîtrise technique.

Concernant les guides de coupes, nous ne pouvons pas recommander d'utiliser un guide de coupe plutôt qu'un autre. Malgré le fait que la navigation permet de réduire le taux de valeurs aberrantes des composants prothétiques fémoral et tibial dans les plans frontal et sagittal, elle augmente le temps opératoire et aucune différence n'a été trouvée sur le plan axial. Concernant les résultats cliniques à court et moyen termes, sur la base du Knee Society Score, une tendance en faveur de la navigation existe. Cependant, nous n'avons pas de résultats sur les résultats cliniques à long terme, sur les complications et sur le coût des différents guides de coupe. L'accéléromètre et la chirurgie assistée par robot semblent augmenter la précision des résections osseuses, mais il n'y a actuellement pas suffisamment d'essais randomisés contrôlés évaluant ces procédures par rapport aux plus courantes pour les résultats radiologiques ou cliniques pour nous permettre de conclure sur ces dispositifs chirurgicaux récents. Plusieurs essais sont « en cours » pour évaluer l'accéléromètre et la robotique.

Contrairement aux voies d'abord où la littérature internationale a fait le tour, nous pensons qu'une mise à jour de la méta-analyse en réseau sur les guides de coupes dans les PTG est nécessaire dans les prochaines années. En effet, depuis la publication de cet article, plusieurs études comparant la robotique dans les PTG ont été publiées (Held et al., 2021; Kayani et al., 2021; Smith et al., 2021; Sun et al., 2021) et d'autres le seront prochainement.

Il reste d'autres sujets sur lesquelles la littérature internationale n'a pas encore tranché afin d'améliorer le résultat des PTG. Ils sont à explorer et à étudier : est-ce qu'un type de plateau tibial prothétique est plus performant qu'un autre ? Quel objectif d'alignement ? alignement systématisé ou spécifiques ? Qu'est-ce qu'une PTG bien équilibrée ? Conservation du ligament croisé postérieur ? Resurfaçage de rotule ? Quel type de fixation utilisé, cimenté, sans ciment ou hybride ?

Pour répondre à ces différentes questions, nous avons vu que les méta-analyses en réseau était une bonne approche méthodologique et adaptée. Néanmoins, les résultats de ces deux méta-analyses en réseau ont mis en évidence la nécessité obligatoire d'homogénéiser et d'élaborer des recommandations sur les critères de jugement à collecter dans les études dans le domaine de l'orthopédie et spécialement dans la chirurgie prothétique comme promu par le groupe Core Outcome Measures in Effectiveness Trials (COMET) (Gargon et al., 2017).

Concernant les courbes d'apprentissages en orthopédie, notre revue de la littérature a permis d'identifier 6 différentes méthodologies. Nous recommandons l'utilisation du LC-CUSUM test. En effet, cette méthodologie nous paraît être la plus simple, la plus flexible, la plus robuste pour modéliser une courbe d'apprentissage et elle peut être utilisée dans plusieurs d'autres spécialités.

Enfin, nous avons observé qu'aucun test statistique validé n'existait dans la littérature permettant des méthodes d'apprentissages entre elles. Nous avons développé un test statistique se basant sur le calcul de l'aire sous la courbe du LC-CUSUM test. La nécessité de le tester dans une application clinique plus importante sera intéressante afin de le valider de manière définitive. S'il s'avère être performant, notre objectif sera de le diffuser et de le rendre accessible. Pour cela, nous envisageons de développer un package R afin de pouvoir le diffuser et de simplifier son utilisation.

Parallèlement, à la validation du test, nous travaillons sur le développement d'un LC-CUSUM test multidimensionnel. En effet, la principale limite du LC-CUSUM test est qu'il se concentre sur une seule dimension. En 2015, Mishra et All (Mishra et al., n.d.) ont passé en revue les différentes méthodes d'application du CUSUM test de manière multidimensionnelle. L'une de ces méthodes consiste à élaborer une norme (ou distance) sur l'espace des dimensions. Différentes distances ont été proposées telles que la distance euclidienne et la distance de Mahalanobis. C'est cette approche que nous allons essayer d'étendre au LC-CUSUM test. Avoir un LC-CUSUM

test multidimensionnel permettra de prendre en compte plusieurs paramètres d'une technique et d'avoir une courbe d'apprentissage plus proche de la réalité. L'objectif final sera de créer une application numérique utilisable par tous les chirurgiens, notamment les internes en orthopédie. Ils pourront ainsi s'autoévaluer et voir de manière objective leur progression.

Bibliographies:

- Abdi, H., Williams, L.J., 2010. Principal component analysis. *WIREs Computational Statistics* 2, 433–459. <https://doi.org/10.1002/wics.101>
- Alan, R., Tria, A., 2010. Quadriceps-Sparing Total Knee Arthroplasty Using the Posterior Stabilized TKA Design. *The Journal of Knee Surgery* 19, 71–76. <https://doi.org/10.1055/s-0030-1248082>
- Alcelik, I., Sukeik, M., Pollock, R., Misra, A., Shah, P., Armstrong, P.,
- Alvand, A., Khan, T., Jenkins, C., Rees, J.L., Jackson, W.F., Dodd, C.A.F.,
- ASNM, 2015. Surveillance des dispositifs médicaux à risque : Prothèses totales de genou (PTG) [WWW Document]. URL (accessed 2.21.18).
- Banerjee, S., Cherian, J.J., Elmallah, R.K., Jauregui, J.J., Pierce, T.P., Mont, M.A., 2015. Robotic-assisted knee arthroplasty. *Expert Review of Medical Devices* 12, 727–735. <https://doi.org/10.1586/17434440.2015.1086264>
- Barrack, R.L., Barnes, C.L., Burnett, R.S.J., Miller, D., Clohisy, J.C., Maloney, W.J., 2009. Minimal incision surgery as a risk factor for early failure of total knee arthroplasty. *J Arthroplasty* 24, 489–498. <https://doi.org/10.1016/j.arth.2009.02.004>
- Bastian, H., Glasziou, P., Chalmers, I., 2010. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 7, e1000326. <https://doi.org/10.1371/journal.pmed.1000326>
- Beecher, H.K., 1955. THE POWERFUL PLACEBO. *Journal of the American Medical Association* 159, 1602–1606. <https://doi.org/10.1001/jama.1955.02960340022006>
- Berger, R.A., Crossett, L.S., Jacobs, J.J., Rubash, H.E., 1998. Malrotation causing patellofemoral complications after total knee arthroplasty. *Clin. Orthop. Relat. Res.* 144–153.
- Berstock, J.R., Murray, J.R., Whitehouse, M.R., Blom, A.W., Beswick, A.D., 2018a. Medial subvastus versus the medial parapatellar approach for total knee replacement: A systematic review and meta-analysis of randomized controlled trials. *EFORT Open Reviews* 3, 78–84. <https://doi.org/10.1302/2058-5241.3.170030>
- Biau, D.J., Porcher, R., 2010. A method for monitoring a process from an out of control to an in control state: Application to the learning curve. *Stat Med* 29, 1900–1909. <https://doi.org/10.1002/sim.3947>

- Biau, D.J., Resche-Rigon, M., Godiris-Petit, G., Nizard, R.S., Porcher, R., 2007. Quality control of surgical and interventional procedures: a review of the CUSUM. *Qual Saf Health Care* 16, 203–207. <https://doi.org/10.1136/qshc.2006.020776>
- Blyth, M.J.G., Anthony, I., Rowe, P., Banger, M.S., MacLean, A., Jones, B., 2017. Robotic arm-assisted versus conventional unicompartmental knee arthroplasty: Exploratory secondary analysis of a randomised controlled trial. *Bone Joint Res* 6, 631–639. <https://doi.org/10.1302/2046-3758.611.BJR-2017-0060.R1>
- Bristol case surgeon claimed to have been on “learning curve,” 1999. *BMJ* 319, 1456.
- Brooks, S.P., Gelman, A., 1998. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics* 7, 434. <https://doi.org/10.2307/1390675>
- Burns, P.B., Rohrich, R.J., Chung, K.C., 2011. The Levels of Evidence and their role in Evidence-Based Medicine. *Plast Reconstr Surg* 128, 305–310. <https://doi.org/10.1097/PRS.0b013e318219c171>
- Chen, K., Dai, X., Li, L., Chen, Z., Cui, H., Lv, S., 2021. Patellar resurfacing versus nonresurfacing in total knee arthroplasty: an updated meta-analysis of randomized controlled trials. *J Orthop Surg Res* 16, 83. <https://doi.org/10.1186/s13018-020-02185-5>
- Cheng, T., Liu, T., Zhang, G., Peng, X., Zhang, X., 2010. Does Minimally Invasive Surgery Improve Short-term Recovery in Total Knee Arthroplasty? *Clinical Orthopaedics and Related Research*® 468, 1635. <https://doi.org/10.1007/s11999-010-1285-9>
- Cherian, J.J., Kapadia, B.H., Banerjee, S., Jauregui, J.J., Issa, K., Mont, M.A., 2014. Mechanical, Anatomical, and Kinematic Axis in TKA: Concepts and Practical Applications. *Curr Rev Musculoskelet Med* 7, 89–95. <https://doi.org/10.1007/s12178-014-9218-y>
- Cooper Samuel, 1838. *A Dictionary Of Practical Surgery*.
- Copay, A.G., Subach, B.R., Glassman, S.D., Polly, D.W., Schuler, T.C., 2007. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 7, 541–546. <https://doi.org/10.1016/j.spinee.2007.01.008>
- Dalury, D.F., Dennis, D.A., 2005. Mini-incision total knee arthroplasty can increase risk of component malalignment. *Clin. Orthop. Relat. Res.* 440, 77–81. <https://doi.org/10.1097/01.blo.0000185757.17401.7b>

- De Valk, E.J., Noorduyn, J.C.A., Mutsaerts, E.L.A.R., 2016. How to assess femoral and tibial component rotation after total knee arthroplasty with computed tomography: a systematic review. *Knee Surg Sports Traumatol Arthrosc* 24, 3517–3528. <https://doi.org/10.1007/s00167-016-4325-5>
- Dhebar, M.I., 2012. Comparison of the minimally invasive and standard medial parapatellar approaches for primary total knee arthroplasty. *Knee Surgery, Sports Traumatology, Arthroscopy* 20, 2502–2512. <https://doi.org/10.1007/s00167-012-1944-3>
- Deep, K., Shankar, S., Mahendra, A., 2017. Computer assisted navigation in total knee and hip arthroplasty. *SICOT-J* 3, 50. <https://doi.org/10.1051/sicotj/2017034>
- Delacour, H., Servonnet, A., Perrot, A., Vigezzi, J., Ramirez, J., 2005. La courbe ROC (receiver operating characteristic) : principes et principales applications en biologie clinique. undefined.
- Delaunay, C., 2014. Registres en orthopédie: Registries in orthopaedics, in: Hutten, D. (Ed.), *Conférences d'enseignement 2014*. Elsevier Masson, Paris, pp. 99–109. <https://doi.org/10.1016/B978-2-294-74506-5.00009-6>
- Dessolle, L., Fréour, T., Barrière, P., Jean, M., Ravel, C., Daraï, E., Biau, D.J., 2010a. How soon can I be proficient in embryo transfer? Lessons from the cumulative summation test for learning curve (LC-CUSUM). *Hum Reprod* 25, 380–386. <https://doi.org/10.1093/humrep/dep391>
- Dias, S., Welton, N.J., Caldwell, D.M., Ades, A.E., 2010b. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 29, 932–944. <https://doi.org/10.1002/sim.3767>
- Dias, S., Welton, N.J., Sutton, A.J., Caldwell, D.M., Lu, G., Ades, A.E., 2013. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Making* 33, 641–656. <https://doi.org/10.1177/0272989X12455847>
- Engh, G.A., Ammeen, D.J., 2014. Unicondylar arthroplasty in knees with deficient anterior cruciate ligaments. *Clin Orthop Relat Res* 472, 73–77. <https://doi.org/10.1007/s11999-013-2982-y>
- Engh, G.A., Holt, B.T., Parks, N.L., 1997. A midvastus muscle-splitting approach for total knee arthroplasty. *J Arthroplasty* 12, 322–331.
- Felson, D.T., 1988. Epidemiology of hip and knee osteoarthritis. *Epidemiol Rev* 10, 1–28.

- Fu, Y., Wang, M., Liu, Y., Fu, Q., 2012a. Alignment outcomes in navigated total knee arthroplasty: a meta-analysis. *Knee Surgery, Sports Traumatology, Arthroscopy* 20, 1075–1082. <https://doi.org/10.1007/s00167-011-1695-6>
- Fu, Y., Wang, M., Liu, Y., Fu, Q., 2012b. Alignment outcomes in navigated total knee arthroplasty: a meta-analysis. *Knee Surg Sports Traumatol Arthrosc* 20, 1075–1082. <https://doi.org/10.1007/s00167-011-1695-6>
- Gandhi, R., Smith, H., Lefavre, K.A., Davey, J.R., Mahomed, N.N., 2011. Complications After Minimally Invasive Total Knee Arthroplasty as Compared With Traditional Incision Techniques. *The Journal of Arthroplasty* 26, 29–35. <https://doi.org/10.1016/j.arth.2009.11.022>
- Gargon, E., Williamson, P.R., Altman, D.G., Blazeby, J.M., Tunis, S., Clarke, M., 2017a. The COMET Initiative database: progress and activities update (2015). *Trials* 18. <https://doi.org/10.1186/s13063-017-1788-8>
- Garstang, S.V., Stitik, T.P., 2006. Osteoarthritis: Epidemiology, Risk Factors, and Pathophysiology. *American Journal of Physical Medicine & Rehabilitation* 85, S2–S11. <https://doi.org/10.1097/01.phm.0000245568.69434.1a>
- Gharaibeh M.A., Solayar G.N., Harris I.A., Chen D.B., MacDessi S.J., 2017. Accelerometer-Based, Portable Navigation (KneeAlign) vs Conventional Instrumentation for Total Knee Arthroplasty: A Prospective Randomized Comparative Trial. *J. Arthroplasty* 32, 777–782. <https://doi.org/10.1016/j.arth.2016.08.025>
- Goble, E.M., Justin, D.F., 2004. Minimally invasive total knee replacement: principles and technique. *Orthopedic Clinics of North America* 35, 235–245. [https://doi.org/10.1016/S0030-5898\(03\)00113-5](https://doi.org/10.1016/S0030-5898(03)00113-5)
- Gouvernement français, n.d. Arrêté du 6 novembre 1995 relatif au Comité national des registres – Légifrance. URL <https://circulaires.legifrance.gouv.fr/jorf/id/JORF-TEXT000000738272> (accessed 8.6.21).
- Green, M.L., Aagaard, E.M., Caverzagie, K.J., Chick, D.A., Holmboe, E., Kane, G., Smith, C.D., Iobst, W., 2009. Charting the road to competence: developmental milestones for internal medicine residency training. *J Grad Med Educ* 1, 5–20. <https://doi.org/10.4300/01.01.0003>

Gregory, T.M., Gregory, J., Sledge, J., Allard, R., Mir, O., 2018. Surgery guided by mixed reality: presentation of a proof of concept. *Acta Orthop* 89, 480–483. <https://doi.org/10.1080/17453674.2018.1506974>

Guillemin, F., Rat, A.C., Mazieres, B., Pouchot, J., Fautrel, B., Euller-Ziegler, L., Fardellone, P., Morvan, J., Roux, C.H., Verrouil, E., Saraux, A., Coste, J., 3000 Osteoarthritis group, 2011. Prevalence of symptomatic hip and knee osteoarthritis: a two-phase population-based survey. *Osteoarthritis Cartilage* 19, 1314–1322. <https://doi.org/10.1016/j.joca.2011.08.004>

Gunston, F.H., 2006. Polycentric knee arthroplasty: prosthetic simulation of normal knee movement. 1971. *Clin Orthop Relat Res* 446, 11–12. <https://doi.org/10.1097/01.blo.0000214423.59829.04>

Haglin, J.M., Eltorai, A.E.M., Gil, J.A., Marcaccio, S.E., Botero-Hincapie, J., Daniels, A.H., 2016. Patient-Specific Orthopaedic Implants. *Orthop Surg* 8, 417–424. <https://doi.org/10.1111/os.12282>

Hailer, N.P., 2015. Orthopedic registry research — limitations and future perspectives. *Acta Orthop* 86, 1–2. <https://doi.org/10.3109/17453674.2014.1002183>

Hall, R.W., 2002. Lessons from the Learning Curve Part 2 of Out of the Box : Overcoming Our Mind Set Legacies [WWW Document]. URL / paper/Lessons-from-the-Learning-Curve-Part-2-of-Out-of-%3A-Hall/ 82149c48a0a38ef6ca1a56999db3781fb7e4baa8 (accessed 7.24.20).

Hardoon, S.L., Lewsey, J.D., Gregg, P.J., Reeves, B.C., van der Meulen, J.H.P., 2006. Continuous monitoring of the performance of hip prostheses. *The Journal of Bone and Joint Surgery. British volume* 88-B, 716–720. <https://doi.org/10.1302/0301-620X.88B6.17334>

Heijink, A., Gomoll, A.H., Madry, H., Drobnič, M., Filardo, G., Espregueira-Mendes, J., Van Dijk, C.N., 2012a. Biomechanical considerations in the pathogenesis of osteoarthritis of the knee. *Knee Surg Sports Traumatol Arthrosc* 20, 423–435. <https://doi.org/10.1007/s00167-011-1818-0>

Held, M.B., Grosso, M.J., Gazgalis, A., Sarpong, N.O., Boddapati, V., Neuwirth, A., Geller, J.A., 2021. Improved Compartment Balancing Using a Robot-Assisted Total Knee Arthroplasty. *Arthroplast Today* 7, 130–134. <https://doi.org/10.1016/j.artd.2020.12.022>

- Higgins, J.P.T., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., Savović, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., 2011b. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 343. <https://doi.org/10.1136/bmj.d5928>
- Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. *BMJ* 327, 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hofmann, A.A., Plaster, R.L., Murdock, L.E., 1991. Subvastus (Southern) approach for primary total knee arthroplasty. *Clin. Orthop. Relat. Res.* 70–77.
- Howell, S.M., Howell, S.J., Kuznik, K.T., Cohen, J., Hull, M.L., 2013. Does A Kinetically Aligned Total Knee Arthroplasty Restore Function Without Failure Regardless of Alignment Category? *Clinical Orthopaedics and Related Research*® 471, 1000–1007. <https://doi.org/10.1007/s11999-012-2613-z>
- Howell, S.M., Roth, J.D., Hull, M.L., 2014. Kinematic Alignment in Total Knee Arthroplasty 10.
- Hozo, S.P., Djulbegovic, B., Hozo, I., 2005a. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology* 5. <https://doi.org/10.1186/1471-2288-5-13>
- Huang, E.H., Copp, S.N., Bugbee, W.D., 2015. Accuracy of A Handheld Accelerometer-Based Navigation System for Femoral and Tibial Resection in Total Knee Arthroplasty. *The Journal of Arthroplasty* 30, 1906–1910. <https://doi.org/10.1016/j.arth.2015.05.055>
- Huijbregts, H.J.T.A.M., Khan, R.J.K., Sorensen, E., Fick, D.P., Haebich, S., 2016a. Patient-specific instrumentation does not improve radiographic alignment or clinical outcomes after total knee arthroplasty. *Acta Orthop* 87, 386–394. <https://doi.org/10.1080/17453674.2016.1193799>
- Hungerford, D.S., Krackow, K.A., 1985. Total joint arthroplasty of the knee. *Clin Orthop Relat Res* 23–33.
- Husson, F., Josse, J., Le, S., Mazet, J., 2020. FactoMineR: Multivariate Exploratory Data Analysis and Data Mining.
- Hutt, J.R.B., LeBlanc, M.-A., Massé, V., Lavigne, M., Vendittoli, P.-A., 2016. Kinematic TKA using navigation: Surgical technique and initial results. *Orthop Traumatol Surg Res* 102, 99–104. <https://doi.org/10.1016/j.otsr.2015.11.010>

Ikawa T., Takemura S., Kim M., Takaoka K., Minoda Y., Kadoya Y., 2017. Usefulness of an accelerometer-based portable navigation system in total knee arthroplasty. *Bone Jt. J.* 99B, 1047–1052. <https://doi.org/10.1302/0301620X.99B8.BJJ20160596.R3>

Insall, J., 1971. A midline approach to the knee. *J Bone Joint Surg Am* 53, 1584–1586.

Insall, J.N., Binazzi, R., Soudry, M., Mestriner, L.A., 1985. Total knee arthroplasty. *Clin Orthop Relat Res* 13–22.

Insall, J.N., Lachiewicz, P.F., Burstein, A.H., 1982. The posterior stabilized condylar prosthesis: a modification of the total condylar design. Two to four-year clinical experience. *J Bone Joint Surg Am* 64, 1317–1323.

Inserm arthrose, 2012. Arthrose Inserm, La science pour la santé. Inserm. URL <https://www.inserm.fr/dossier/arthrose/> (accessed 9.4.21).

Jackson, J.P., Elson, R.A., 1973. Evaluation of the Walldius and other prostheses for knee arthroplasty. *Clin Orthop Relat Res* 104–114. <https://doi.org/10.1097/00003086-197307000-00013>

Jacofsky, D.J., Allen, M., 2016. Robotics in Arthroplasty: A Comprehensive Review. *The Journal of Arthroplasty* 31, 2353–2363. <https://doi.org/10.1016/j.arth.2016.05.026>

Jeffery, R.S., Morris, R.W., Denham, R.A., 1991. Coronal alignment after total knee replacement. *J Bone Joint Surg Br* 73, 709–714.

Karl Pearson, 1904. Report on Certain Enteric Fever Inoculation Statistics. *Br Med J* 2, 1243–1246. <https://doi.org/10.1136/bmj.2.2288.1243>

Kassambara, A., Mundt, F., 2020. *Facto extra: Extract and Visualize the Results of Multivariate Data Analyses*.

Kayani, B., Tahmassebi, J., Ayuob, A., Konan, S., Oussedik, S., Haddad, F.S., 2021. A prospective randomized controlled trial comparing the systemic inflammatory response in conventional jig-based total knee arthroplasty versus robotic-arm assisted total knee arthroplasty. *Bone Joint J* 103-B, 113–122. <https://doi.org/10.1302/0301-620X.103B1.BJJ-2020-0602.R2>

Kazarian, G.S., Siow, M.Y., Chen, A.F., Deirmengian, C.A., 2018a. Comparison of Quadriceps-Sparing and Medial Parapatellar Approaches in Total Knee Arthroplasty: A Meta-Analysis of Randomized Controlled Trials. *The Journal of Arthroplasty* 33, 277–283. <https://doi.org/10.1016/j.arth.2017.08.025>

Kim, Y.-H., Kim, J.-S., Yoon, S.-H., 2007b. Alignment and orientation of the components in total knee replacement with and without navigation support: a prospective, randomised study. *J Bone Joint Surg Br* 89, 471–476. <https://doi.org/10.1302/0301-620X.89B4.18878>

Kim, Y.-H., Yoon, S.-H., Park, J.-W., 2020. Does Robotic-assisted TKA Result in Better Outcome Scores or Long-Term Survivorship Than Conventional TKA? A Randomized, Controlled Trial. *Clin Orthop Relat Res* 478, 266–275. <https://doi.org/10.1097/CORR.0000000000000916>

King, J., Stamper, D.L., Schaad, D.C., Leopold, S.S., 2007. Minimally invasive total knee arthroplasty compared with traditional total knee arthroplasty. Assessment of the learning curve and the postoperative recuperative period. *J Bone Joint Surg Am* 89, 1497–1503. <https://doi.org/10.2106/JBJS.F.00867>

Kinney M.C., Cidambi K.R., Severns D.L., Gonzales F.B., 2018. Comparison of the iAssist Handheld Guidance System to Conventional Instruments for Mechanical Axis Restoration in Total Knee Arthroplasty. *J. Arthroplasty* 33, 61–66. <https://doi.org/10.1016/j.arth.2017.06.004>

Kurtz, S., Ong, K., Lau, E., Mowat, F., Halpern, M., 2007a. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *J Bone Joint Surg Am* 89, 780–785. <https://doi.org/10.2106/JBJS.F.00222>

Kurtz, S.M., Ong, K.L., Lau, E., Widmer, M., Maravic, M., Gómez-Barrena, E., de Pina, M. de F., Manno, V., Torre, M., Walter, W.L., de Steiger, R., Geesink, R.G.T., Peltola, M., Röder, C., 2011. International survey of primary and revision total knee replacement. *Int Orthop* 35, 1783–1789. <https://doi.org/10.1007/s00264-011-1235-5>

Kuzyk, P.R.T., Higgins, G.A., Tunggal, J.A.W., Sellan, M.E., Waddell, J.P., Schemitsch, E.H., 2012a. Computer Navigation vs Extramedullary Guide for Sagittal Alignment of Tibial Components: Radiographic Study and Meta-Analysis. *The Journal of Arthroplasty* 27, 630–637. <https://doi.org/10.1016/j.arth.2011.07.001>

Laurent, P., 2016. Patient-Specific Instruments in Orthopedics, in: *Computer-Assisted Musculoskeletal Surgery*. Springer, Cham, pp. 163–179. https://doi.org/10.1007/978-3-319-12943-3_13

Lawrence, R.C., Felson, D.T., Helmick, C.G., Arnold, L.M., Choi, H., Deyo, R.A., Gabriel, S., Hirsch, R., Hochberg, M.C., Hunder, G.G., Jordan, J.M., Katz, J.N., Kremers, H.M., Wolfe, F., National Arthritis Data Workgroup, 2008. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. *Arthritis Rheum* 58, 26–35. <https://doi.org/10.1002/art.23176>

Leopold, S.S., 2015a. Editorial: “Pencil and Paper” Research? Network Meta-analysis and Other Study Designs That Do Not Enroll Patients. *Clin Orthop Relat Res* 473, 2163–2165. <https://doi.org/10.1007/s11999-015-4329-3>

Lexer, E., 1914. FREE TRANSPLANTATION. *Ann Surg* 60, 166–194. <https://doi.org/10.1097/00000658-191408000-00005>

Liddle, A.D., Pandit, H., Judge, A., Murray, D.W., 2016. Effect of Surgical Caseload on Revision Rate Following Total and Unicompartmental Knee Replacement. *J Bone Joint Surg Am* 98, 1–8. <https://doi.org/10.2106/JBJS.N.00487>

Liow, M.H.L., Goh, G.S.-H., Wong, M.K., Chin, P.L., Tay, D.K.-J., Yeo, S.-J., 2017. Robotic-assisted total knee arthroplasty may lead to improvement in quality-of-life measures: a 2-year follow-up of a prospective randomized trial. *Knee Surg Sports Traumatol Arthrosc* 25, 2942–2951. <https://doi.org/10.1007/s00167-016-4076-3>

Liu, H.-W., Gu, W.-D., Xu, N.-W., Sun, J.-Y., 2014a. Surgical Approaches in Total Knee Arthroplasty: A Meta-Analysis Comparing the Midvastus and Subvastus to the Medial Peripatellar Approach. *The Journal of Arthroplasty* 29, 2298–2304. <https://doi.org/10.1016/j.arth.2013.10.023>

Longstaff, L.M., Sloan, K., Stamp, N., Scaddan, M., Beaver, R., 2009. Good Alignment After Total Knee Arthroplasty Leads to Faster Rehabilitation and Better Function. *The Journal of Arthroplasty* 24, 570–578. <https://doi.org/10.1016/j.arth.2008.03.002>

Lu, G., Ades, A.E., 2004a. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 23, 3105–3124. <https://doi.org/10.1002/sim.1875>

Luo, Z., Zhou, K., Peng, L., Shang, Q., Pei, F., Zhou, Z., 2020. Similar results with kinematic and mechanical alignment applied in total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 28, 1720–1735. <https://doi.org/10.1007/s00167-019-05584-2>

Lyman, S., Lee, Y.-Y., McLawhorn, A.S., Islam, W., MacLean, C.H., 2018. What Are the Minimal and Substantial Improvements in the HOOS and KOOS and JR Versions After Total Joint Replacement? *Clin Orthop Relat Res* 476, 2432–2441. <https://doi.org/10.1097/CORR.0000000000000456>

- Mannan, A., Akinyooye, D., Hossain, F., 2017. A Meta-analysis of Functional Outcomes in Patient-Specific Instrumented Knee Arthroplasty. *J Knee Surg* 30, 668–674. <https://doi.org/10.1055/s-0036-1593869>
- Mannan A., Smith T.O., 2016. Favourable rotational alignment outcomes in PSI knee arthroplasty: A Level 1 systematic review and meta-analysis. *Knee* 23, 186–190. <https://doi.org/10.1016/j.knee.2015.08.006>
- Maredupaka, S., Meshram, P., Chatte, M., Kim, W.H., Kim, T.K., 2020. Minimal clinically important difference of commonly used patient-reported outcome measures in total knee arthroplasty: review of terminologies, methods and proposed values. *Knee Surg Relat Res* 32, 19. <https://doi.org/10.1186/s43019-020-00038-3>
- Mason, J.B., Fehring, T.K., Estok, R., Banel, D., Fahrback, K., 2007. Meta-analysis of alignment outcomes in computer-assisted total knee arthroplasty surgery. *J Arthroplasty* 22, 1097–1106. <https://doi.org/10.1016/j.arth.2007.08.001>
- Mattei, L., Pellegrino, P., Calò, M., Bistolfi, A., Castoldi, F., 2016. Patient specific instrumentation in total knee arthroplasty: a state of the art. *Ann Transl Med* 4. <https://doi.org/10.21037/atm.2016.03.33>
- Mishra, S., Vanli, O.A., Park, C., n.d. A Multivariate Cumulative Sum Method for Continuous Damage Monitoring with Lamb-wave Sensors 11.
- Mizner, R.L., Petterson, S.C., Stevens, J.E., Vandenborne, K., Snyder-Mackler, L., 2005. Early quadriceps strength loss after total knee arthroplasty. The contributions of muscle atrophy and failure of voluntary muscle activation. *J Bone Joint Surg Am* 87, 1047–1053. <https://doi.org/10.2106/JBJS.D.01992>
- Moskal, J., Capps, S., Mann, J., Scanelli, J., 2013. Navigated versus Conventional Total Knee Arthroplasty. *Journal of Knee Surgery* 27, 235–248. <https://doi.org/10.1055/s-0033-1360659>
- Murray, D.W., Price, A.J., 2018. The impact of patient-specific instrumentation on unicompartamental knee arthroplasty: a prospective randomised controlled study. *Knee Surg Sports Traumatol Arthrosc* 26, 1662–1670. <https://doi.org/10.1007/s00167-017-4677-5>
- Nam D., Cody E.A., Nguyen J.T., Figgie M.P., Mayman D.J., 2014. Extra-medullary guides versus portable, accelerometer-based navigation for tibial alignment in total knee arthroplasty: A randomized, controlled trial: Winner of the 2013 HAP PAUL award. *J Arthroplasty* 29, 288–294. <https://doi.org/10.1016/j.arth.2013.06.006>

- Newman, J.M., Sodhi, N., Dekis, J.C., Khlopas, A., Piuzzi, N.S., Sultan, A.A., Levin, J.M., Mont, M.A., 2020. Survivorship and Functional Outcomes of Cementless versus Cemented Total Knee Arthroplasty: A Meta-Analysis. *J Knee Surg* 33, 270–278. <https://doi.org/10.1055/s-0039-1678525>
- Oliveria, S.A., Felson, D.T., Reed, J.I., Cirillo, P.A., Walker, A.M., 1995a. Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization. *Arthritis Rheum* 38, 1134–1141. <https://doi.org/10.1002/art.1780380817>
- Ollier, L., University of Glasgow. Library, University of Glasgow. Library, 1885. *Traité des résections et des opérations conservatrices qu'on peut pratiquer sur le système osseux* [electronic resource]. Paris : G. Masson.
- Ollivier, M., Parratte, S., Lunebourg, A., Viehweger, E., Argenson, J.-N., 2016. The John Insall Award: No Functional Benefit After Unicompartmental Knee Arthroplasty Performed With Patient-specific Instrumentation: A Randomized Trial. *Clinical Orthopaedics and Related Research*® 474, 60–68. <https://doi.org/10.1007/s11999-015-4259-0>
- Onggo, James, Onggo, Jason, Phan, K., Wilson, C., 2020. Comparison of infection in cemented, cementless and hybrid primary total knee arthroplasty: a network meta-analysis and systematic review of randomized clinical trials. *ANZ J Surg* 90, 1289–1298. <https://doi.org/10.1111/ans.16078>
- Oussedik, S., Abdel, M.P., Victor, J., Pagnano, M.W., Haddad, F.S., 2020. Alignment in total knee arthroplasty. *Bone Joint J* 102-B, 276–279. <https://doi.org/10.1302/0301-620X.102B3.BJJ-2019-1729>
- Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* n71. <https://doi.org/10.1136/bmj.n711>.
- Page Es. Continuous inspection. *Biometrika*. 1954;41(1-2):100-115. doi:10.1093/biomet/41.1-2.100
- Pagès, J., 2004. Analyse factorielle de données mixtes. *Revue de Statistique Appliquée* 52, 93–111.

Parratte, Sébastien, Blanc, G., Boussemart, T., Ollivier, M., Corroller, T.L., Argenson, J.-N., 2013. Rotation in total knee arthroplasty: no difference between patient-specific and conventional instrumentation. *Knee Surg Sports Traumatol Arthrosc* 21, 2213–2219. <https://doi.org/10.1007/s00167-013-2623-8>

Parratte, S., Pagnano, M.W., Trousdale, R.T., Berry, D.J., 2010. Effect of postoperative mechanical axis alignment on the fifteen-year survival of modern, cemented total knee replacements. *J Bone Joint Surg Am* 92, 2143–2149. <https://doi.org/10.2106/JBJS.I.01398>

Pickering, S., Armstrong, D., 2012. Focus on alignment in total knee replacement. *J Bone Joint Surg* 1–3.

Piriou, P., Biau, D., Judet, T., 2006. [Every two year follow-up of arthroplasties: myth or reality?]. *Rev Chir Orthop Reparatrice Appar Mot* 92, 290–292. [https://doi.org/10.1016/s0035-1040\(06\)75739-7](https://doi.org/10.1016/s0035-1040(06)75739-7)

Putman, S., Dartus, J., Migaud, H., Pasquier, G., Girard, J., Preda, C., Duhamel, A., 2021. Can the minimal clinically important difference be determined in a French-speaking population with primary hip replacement using one PROM item and the Anchor strategy? *Orthopaedics & Traumatology: Surgery & Research* 107, 102830. <https://doi.org/10.1016/j.otsr.2021.102830>

Rabenda, V., Manette, C., Lemmens, R., Mariani, A.-M., Struvay, N., Reginster, J.-Y., 2007. Prevalence and impact of osteoarthritis and osteoporosis on health-related quality of life among active subjects. *Aging Clin Exp Res* 19, 55–60.

Rajgopal, A., Ahuja, N., Dolai, B., 2005. Total knee arthroplasty in stiff and ankylosed knees. *J Arthroplasty* 20, 585–590. <https://doi.org/10.1016/j.arth.2005.04.002>

Ramsay, C.R., Wallace, S.A., Garthwaite, P.H., Monk, A.F., Russell, I.T., Grant, A.M., 2002. Assessing the learning curve effect in health technologies. Lessons from the nonclinical literature. *Int J Technol Assess Health Care* 18, 1–10.

Ranawat, A.S., Ranawat, C.S., 2012. The history of total knee arthroplasty, in: Bonnin, M., Amendola, A., Bellemans, J., MacDonald, S., Ménétrey, J. (Eds.), *The Knee Joint: Surgical Techniques and Strategies*. Springer, Paris, pp. 699–707. https://doi.org/10.1007/978-2-287-99353-4_63

- Rebal, B.A., Babatunde, O.M., Lee, J.H., Geller, J.A., Patrick, D.A.J., Macaulay, W., 2014. Imageless computer navigation in total knee arthroplasty provides superior short term functional outcomes: a meta-analysis. *J Arthroplasty* 29, 938–944. <https://doi.org/10.1016/j.arth.2013.09.018>
- Riddle, D.L., 2012. Validity of clinical measures of frontal plane knee alignment: data from the Osteoarthritis Initiative. *Man Ther* 17, 459–465. <https://doi.org/10.1016/j.math.2012.05.006>
- Riley, R.D., Higgins, J.P.T., Deeks, J.J., 2011. Interpretation of random effects meta-analyses. *BMJ* 342, d549. <https://doi.org/10.1136/bmj.d549>
- Ritter, M.A., Faris, P.M., Keating, E.M., Meding, J.B., 1994. Postoperative alignment of total knee replacement. Its effect on survival. *Clin. Orthop. Relat. Res.* 153–156.
- Rivière, E., Quinton, A., Dehail, P., 2019. Analyse de la discrimination des notes aux ECN informatisées de 2016. *La Revue de Médecine Interne* 40, 286–290. <https://doi.org/10.1016/j.revmed.2018.10.386>
- Robinson, R.P., 2005. The early innovators of today's resurfacing condylar knees. *J Arthroplasty* 20, 2–26. <https://doi.org/10.1016/j.arth.2004.11.002>
- Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B., Richardson, W.S., 1996. Evidence based medicine: what it is and what it isn't. *BMJ* 312, 71–72.
- Salanti, G., Giovane, C.D., Chaimani, A., Caldwell, D.M., Higgins, J.P.T., 2014a. Evaluating the Quality of Evidence from a Network Meta-Analysis. *PLOS ONE* 9, e99682. <https://doi.org/10.1371/journal.pone.0099682>
- Schulz, K.F., Altman, D.G., Moher, D., CONSORT Group, 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 340, c332. <https://doi.org/10.1136/bmj.c332>
- Scuderi, G.R., Tenholder, M., Capeci, C., 2004. Surgical Approaches in Mini-incision Total Knee Arthroplasty: *Clinical Orthopaedics and Related Research* 428, 61–67. <https://doi.org/10.1097/01.blo.0000148574.79874.d0>
- Sharma, L., Lou, C., Cahue, S., Dunlop, D.D., 2000. The mechanism of the effect of obesity in knee osteoarthritis: The mediating role of malalignment. *Arthritis & Rheumatism* 43, 568. [https://doi.org/10.1002/1529-0131\(200003\)43:3<568::AID-ANR13>3.0.CO;2-E](https://doi.org/10.1002/1529-0131(200003)43:3<568::AID-ANR13>3.0.CO;2-E)

- Sikorski, J.M., 2008. Alignment in total knee replacement. *J Bone Joint Surg Br* 90, 1121–1127. <https://doi.org/10.1302/0301-620X.90B9.20793>
- Silva, M., Shepherd, E.F., Jackson, W.O., Pratt, J.A., McClung, C.D., Schmalzried, T.P., 2003. Knee strength after total knee arthroplasty. *J Arthroplasty* 18, 605–611. [https://doi.org/10.1016/s0883-5403\(03\)00191-8](https://doi.org/10.1016/s0883-5403(03)00191-8)
- Smith, A.F., Eccles, C.J., Bhimani, S.J., Denehy, K.M., Bhimani, R.B., Smith, L.S., Malkani, A.L., 2021. Improved Patient Satisfaction following Robotic-Assisted Total Knee Arthroplasty. *J Knee Surg* 34, 730–738. <https://doi.org/10.1055/s-0039-1700837>
- Smith, T.O., King, J.J., Hing, C.B., 2012. A meta-analysis of randomised controlled trials comparing the clinical and radiological outcomes following minimally invasive to conventional exposure for total knee arthroplasty. *The Knee* 19, 1–7. <https://doi.org/10.1016/j.knee.2010.12.001>
- Song, E.K., N, M., Lee, S.-H., Na, B.-R., Seon, J.-K., 2016. Comparison of Outcome and Survival After Unicompartmental Knee Arthroplasty Between Navigation and Conventional Techniques With an Average 9-Year Follow-Up. *The Journal of Arthroplasty* 31, 395–400. <https://doi.org/10.1016/j.arth.2015.09.012>
- Soomro, N.A., Hashimoto, D.A., Porteous, A.J., Ridley, C.J.A., Marsh, W.J., Ditto, R., Roy, S., 2019. Systematic review of learning curves in robot-assisted surgery. *BJS Open* 4, 27–44. <https://doi.org/10.1002/bjs5.50235>
- Specogna, A.V., Birmingham, T.B., Hunt, M.A., Jones, I.C., Jenkyn, T.R., Fowler, P.J., Giffin, J.R., 2007. Radiographic measures of knee alignment in patients with varus gonarthrosis: effect of weightbearing status and associations with dynamic joint load. *Am J Sports Med* 35, 65–70. <https://doi.org/10.1177/0363546506293024>
- Steiner, S.H., Cook, R.J., Farewell, V.T., 2001. Risk-adjusted monitoring of binary surgical outcomes. *Med Decis Making* 21, 163–169. <https://doi.org/10.1177/0272989X0102100301>
- Stern, S.H., Moeckel, B.H., Insall, J.N., 1991. Total knee arthroplasty in valgus knees. *Clin. Orthop. Relat. Res.* 5–8.
- Stulberg, S.D., Loan, P., Sarin, V., 2002. Computer-assisted navigation in total knee replacement: results of an initial experience in thirty-five patients. *JBJS* 84, 90–98.

- Sun, Xuedong, Gao, X., Sun, Xiaohui, Su, Z., 2021b. Comparison of clinical and radiographic results between total knee arthroplasties using medial pivot and posterior-stabilized prosthesis: A meta-analysis. *Medicine (Baltimore)* 100, e23809. <https://doi.org/10.1097/MD.00000000000023809>
- Teel, A.J., Esposito, J.G., Lanting, B.A., Howard, J.L., Schemitsch, E.H., 2019. Patellar Resurfacing in Primary Total Knee Arthroplasty: A Meta-Analysis of Randomized Controlled Trials. *J Arthroplasty* 34, 3124–3132. <https://doi.org/10.1016/j.arth.2019.07.019>
- Thiengwittayaporn, S., Fusakul, Y., Kangkano, N., Jarupongprapa, C., Charoenphandhu, N., 2016. Hand-held navigation may improve accuracy in minimally invasive total knee arthroplasty: a prospective randomized controlled trial. *Int Orthop* 40, 51–57. <https://doi.org/10.1007/s00264-015-2848-x>
- Thienpont E., Schwab P.-E., Fennema P., 2017. Efficacy of patient-specific instruments in total knee arthroplasty: A systematic review and meta-analysis. *J. Bone Jt. Surg. Am.* Vol. 99, 521–530. <https://doi.org/10.2106/JBJS.16.00496>
- Tria, A.J., Coon, T.M., 2003. Minimal Incision Total Knee Arthroplasty: Early Experience: Early Experience. *Clinical Orthopaedics and Related Research* 416, 185–190. <https://doi.org/10.1097/01.blo.0000093030.56370.d9>
- Trillat, A., Dejour, H., Bousquet, G., Grammont, L., 1973. [The rotational knee prosthesis]. *Acta Orthop Belg* 39, 296–319.
- Trinquart, L., Attiche, N., Bafeta, A., Porcher, R., Ravaud, P., 2016a. Uncertainty in Treatment Rankings: Reanalysis of Network Meta-analyses of Randomized Trials. *Annals of Internal Medicine*. 2016;164(10):666. doi:10.7326/M15-2521
- Vail, T.P., 2004. Minimally Invasive Knee Arthroplasty: Clinical Orthopaedics and Related Research 428, 51–52. <https://doi.org/10.1097/01.blo.0000147706.53226.53>
- Valkenhoef, G. van, Kuiper, J., 2020. gemtc: Network Meta-Analysis Using Bayesian Methods.
- Valsamis, E.M., Chouari, T., O'Dowd-Booth, C., Rogers, B., Ricketts, D., 2018a. Learning curves in surgery: variables, analysis and applications.
- Valsamis, E.M., Golubic, R., Glover, T.E., Husband, H., Hussain, A., Jenabzadeh, A.-R., 2018b. Modeling Learning in Surgical Practice. *Journal of Surgical Education* 75, 78–87. <https://doi.org/10.1016/j.jsurg.2017.06.015>

Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., Vandenbroucke, J.P., STROBE Initiative, 2014. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg* 12, 1495–1499. <https://doi.org/10.1016/j.ijisu.2014.07.013>

Wang, K., Sun, H., Zhang, K., Li, S., Wu, G., Zhou, J., Sun, X., 2020. Better outcomes are associated with cementless fixation in primary total knee arthroplasty in young patients: A systematic review and meta-analysis of randomized controlled trials. *Medicine (Baltimore)* 99, e18750. <https://doi.org/10.1097/MD.00000000000018750>

Weil, G., Motamed, C., Biau, D.J., Guye, M.L., 2017a. Learning curves for three specific procedures by anesthesiology residents using the learning curve cumulative sum (LC-CUSUM) test. *Korean J Anesthesiol* 70, 196–202. <https://doi.org/10.4097/kjae.2017.70.2.196>

Weir, C.J., Butcher, I., Assi, V., Lewis, S.C., Murray, G.D., Langhorne, P., Brady, M.C., 2018a. Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review. *BMC Medical Research Methodology* 18. <https://doi.org/10.1186/s12874-018-0483-0>

WHO Scientific Group on the Burden of Musculoskeletal Conditions at the Start of the New Millennium, 2003a. The burden of musculoskeletal conditions at the start of the new millennium. *World Health Organ Tech Rep Ser* 919, i–x, 1–218, back cover.

Wilcoxon, F., 1945. Individual Comparisons by Ranking Methods. *Biometrics Bulletin* 1, 80–83. <https://doi.org/10.2307/3001968>

Wittenaar, R., Smith, L., Aden, K., 2013. Background paper 6.12 osteoarthritis. World Health Organisation.

Wright, T.P., 1936. Factors Affecting the Cost of Airplanes. *Journal of the Aeronautical Sciences* 3, 122–128. <https://doi.org/10.2514/8.155>

Zabell, S.L., 2008. On Student's 1908 Article "The Probable Error of a Mean." *Journal of the American Statistical Association* 103, 1–7. <https://doi.org/10.1198/016214508000000030>

Zeng, H.B., Ying, X.Z., Chen, G.J., Yang, X.Q., Lin, D.D., Li, Z.J., Liu, H.X., 2015. Extramedullary versus intramedullary tibial alignment technique in total knee arthroplasty: A meta-analysis of randomized controlled trials. *Clinics (Sao Paulo)* 70, 714–719. [https://doi.org/10.6061/clinics/2015\(10\)10](https://doi.org/10.6061/clinics/2015(10)10)

Appendices

Appendix 1. Search terms

- PubMed

((((((((((arthroplasty, replacement, knee[MeSH Terms]) OR knee prosthesis[MeSH Terms]) OR ((knee[Title/Abstract]) AND replacement[Title/Abstract]) AND arthroplasty[Title/Abstract])) OR ((total[Title/Abstract]) AND knee[Title/Abstract]) AND arthroplasty[Title/Abstract])) OR ((total[Title/Abstract]) AND knee[Title/Abstract]) AND replacement[Title/Abstract])) OR TKA[Title/Abstract]) OR TKR[Title/Abstract]) OR ((knee[Title/Abstract]) AND prosthesis[Title/Abstract])) AND (((((((((((navigation[Title/Abstract]) OR navigated[Title/Abstract]) OR ((imageless[Title/Abstract]) AND computer[Title/Abstract]) AND assisted[Title/Abstract])) OR ((computer[Title/Abstract]) AND assisted[Title/Abstract])) OR ((computer[Title/Abstract]) AND navigation[Title/Abstract])) OR (((computer[Title/Abstract]) AND assisted[Title/Abstract]) AND techniques[Title/Abstract])) OR ((image[Title/Abstract]) AND free[Title/Abstract])))) OR (((Accelerometric[Title/Abstract]) OR iAssist[Title/Abstract]) OR accelerometers[Title/Abstract]) OR accelerometer-based[Title/Abstract])) OR (((iblock[Title/Abstract]) OR mako[Title/Abstract]) OR (((robot assisted[Title/Abstract]) OR robotic guidance[Title/Abstract])) OR ((RIO[Title/Abstract]) OR NAVIO[Title/Abstract]) OR ROBODOC[Title/Abstract])))) OR (((((((((PSI[Title/Abstract]) OR (((patient[Title/Abstract]) AND specific[Title/Abstract]) AND cutting[Title/Abstract]) AND blocks[Title/Abstract])) OR ((patient[Title/Abstract]) AND specific[Title/Abstract]) AND instrumentation[Title/Abstract])) OR ((patient[Title/Abstract]) AND specific[Title/Abstract]) AND instruments[Title/Abstract])) OR ((patient[Title/Abstract]) AND specific[Title/Abstract]) AND guides[Title/Abstract])) OR ((custom[Title/Abstract]) AND fit[Title/Abstract]) AND instrumentation[Title/Abstract])) OR ((custom[Title/Abstract]) AND made[Title/Abstract]) AND guides[Title/Abstract])) OR ((patient[Title/Abstract]) AND matched[Title/Abstract]) AND instrumentation[Title/Abstract])) OR ((customized[Title/Abstract]) AND instrumentation[Title/Abstract])) OR (((patient[Title/Abstract]) AND specific[Title/Abstract]) AND templating[Title/Abstract])))) AND (((((((((((randomized controlled trial[Publication Type]) OR controlled clinical trial[Publication Type]) OR randomized[Title/Abstract]) OR placebo[Title/Abstract]) OR drug therapy[MeSH Subheading]) OR randomly[Title/Abstract]) OR trial[Title/Abstract]) OR groups[Title/Abstract])) NOT ((animals[MeSH Terms]) NOT humans[MeSH Terms]))

- Embase

#24 #9 AND #23
#23 #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22
#22 'standard':ab,ti AND 'approach':ab,ti AND [randomized controlled trial]/lim
#21 'quadriceps':ab,ti AND 'sparring':ab,ti AND [randomized controlled trial]/lim
#20 'midvastus':ab,ti AND [randomized controlled trial]/lim
#19 'subvastus':ab,ti AND [randomized controlled trial]/lim
#18 'southern':ab,ti AND 'approach':ab,ti AND [randomized controlled trial]/lim

- #17 'median':ab,ti AND 'parapatellar':ab,ti AND [randomized controlled trial]/lim
- #16 'medial':ab,ti AND 'parapatellar':ab,ti AND [randomized controlled trial]/lim
- #15 'less':ab,ti AND 'invasive':ab,ti AND [randomized controlled trial]/lim
- #14 'exposure':ab,ti AND [randomized controlled trial]/lim
- #13 'surgical':ab,ti AND 'approach':ab,ti AND [randomized controlled trial]/lim
- #12 'dissection':ab,ti AND [randomized controlled trial]/lim
- #11 'incision':ab,ti AND [randomized controlled trial]/lim
- #10 'minimally':ab,ti AND 'invasive':ab,ti AND [randomized controlled trial]/lim
- #9 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8
- #8 'knee':ab,ti AND 'arthroplasty':ab,ti AND 'replacement':ab,ti
- #7 'knee':ti,ab,kw AND 'prosthesis':ti,ab,kw
- #6 'tkr':ti,ab,kw
- #5 'tka':ti,ab,kw
- #4 'total':ti,ab,kw AND 'knee':ti,ab,kw AND 'replacement':ti,ab,kw
- #3 'total':ti,ab,kw AND 'knee':ti,ab,kw AND 'arthroplasty':ti,ab,kw
- #2 'total knee arthroplasty'/exp
- #1 'knee prosthesis'/exp

Central

- #1 MeSH descriptor: [Arthroplasty, Replacement, Knee] explode all trees
- #2 MeSH descriptor: [Knee Prosthesis] explode all trees
- #3 total:ti,ab,kw and knee and arthroplasty (Word variations have been searched)
- #4 total:ti,ab,kw and knee and replacement (Word variations have been searched)
- #5 tka:ti,ab,kw (Word variations have been searched)
- #6 tkr:ti,ab,kw (Word variations have been searched)
- #7 knee:ti,ab,kw and prosthesis:ti,ab,kw (Word variations have been searched)
- #8 knee:ti,ab,kw and arthroplasty:ti,ab,kw and replacement:ti,ab,kw (Word variations have been searched)
- #9 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 in Trials
- #10 (minimally):ti,ab,kw AND (invasive):ti,ab,kw in Trials (Word variations have been searched)
- #11 (incision):ti,ab,kw in Trials (Word variations have been searched)
- #12 (dissection):ti,ab,kw in Trials (Word variations have been searched)
- #13 (surgical):ti,ab,kw AND (approach):ti,ab,kw in Trials (Word variations have been searched)
- #14 (exposure):ti,ab,kw in Trials (Word variations have been searched)
- #15 (less):ti,ab,kw AND (invasive):ti,ab,kw in Trials (Word variations have been searched)
- #16 (medial):ti,ab,kw AND (parapatellar):ti,ab,kw in Trials (Word variations have been searched)
- #17 (median):ti,ab,kw AND (parapatellar):ti,ab,kw in Trials (Word variations have been searched)
- #18 (southern):ti,ab,kw AND (approach):ti,ab,kw in Trials (Word variations have been searched)
- #19 (subvastus):ti,ab,kw in Trials (Word variations have been searched)
- #20 (midvastus):ti,ab,kw in Trials (Word variations have been searched)

- #21 (quadriceps):ti,ab,kw AND (sparring):ti,ab,kw in Trials (Word variations have been searched)
- #22 (standard):ti,ab,kw AND (approach):ti,ab,kw in Trials (Word variations have been searched)
- #23 #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22
- #24 #9 AND #23

Appendix 2. Outcomes definition

Radiological outcomes:

Three radiological data were extracted:

- The proportion at 6 months post-surgery of malalignments in the hip-knee-ankle (HKA) angle, which measures the mechanical axis of the lower limb. Malalignments were defined by an HKA angle higher or lower than $180^{\circ} \pm 3^{\circ}$.
- The proportion at 6 months post-surgery of malalignments for tibial/femoral implant in sagittal plane, defined by an angle higher or lower than $90 \pm 3^{\circ}$.

HKA is the angle formed by the lines uniting at the center of the femoral head (hip) to the middle of the knee (knee) and then from the center of the knee to the middle of the ankle (ankle) (1). HKA is measured using an x-ray of the lower limb.

Femoral and tibial components alignment in the sagittal plane are evaluated with a classical lateral radiograph or using CT-scan. For the tibia component, the critical feature is the tibial slope, defined by the angle between the tibial shaft axis and the tangent to the tibial plateau. Regarding the femoral component, the main parameter in the sagittal plane is the angle between the femoral shaft axis and the tangent to the femoral condyles. Tibial and femoral implant rotations are analyzed on a postoperative CT-scan.

1. Luo C-F. Reference axes for reconstruction of the knee. *Knee*. 2004;11:251–7.
2. Abdel MP, Oussedik S, Parratte S, Lustig S, Haddad FS. Coronal alignment in total knee replacement: historical review, contemporary analysis, and future direction. *Bone Joint J*. 2014;96:857–862.
3. Fang DM, Ritter MA, Davis KE. Coronal alignment in total knee arthroplasty: just how important is it? *J Arthroplasty*. 2009;24(6 Suppl):39–43.

B. KSS score

The Knee Society Score (KSS) is composed of a clinical part and a functional part (1). In the clinical part, seven items are evaluated by the clinician during the last 4 weeks (about 100 points): pain, flexion contracture, extension lag, total range of flexion, alignment (varus and valgus), AP stability, and mediolateral stability. In the functional part, three items are evaluated by the patient: walking, stairs, and walking aids used (about 100 points). The total score is calculated by summing the results of clinical part and functional part: a high score is related to little knee joint damage.

KSS		
Part 1- Knee score :		
Pain : <ul style="list-style-type: none"> • None • Mild/occasional • Mild (stairs only) • Mild (Walking and Stairs) • Moderate-occasional • Moderate-Continual • Severe 	Flexion contracture (if present) : <ul style="list-style-type: none"> • 5-10° • 10-15° • 16-20° • >20° 	
	Extension lag : <ul style="list-style-type: none"> • <10° • 10-20° • >20° 	
Total range of flexion : <ul style="list-style-type: none"> <li style="width: 33%;">• 0-5° <li style="width: 33%;">• 51-55° <li style="width: 33%;">• 101-105° <li style="width: 33%;">• 6-10° <li style="width: 33%;">• 56-60° <li style="width: 33%;">• 106-110° <li style="width: 33%;">• 11-15° <li style="width: 33%;">• 61-65° <li style="width: 33%;">• 111-115° <li style="width: 33%;">• 16-20° <li style="width: 33%;">• 66-70° <li style="width: 33%;">• 116-120° <li style="width: 33%;">• 21-25° <li style="width: 33%;">• 71-75° <li style="width: 33%;">• 121-125° <li style="width: 33%;">• 26-30° <li style="width: 33%;">• 76-80° <li style="width: 33%;">• 31-35° <li style="width: 33%;">• 81-85° <li style="width: 33%;">• 36-40° <li style="width: 33%;">• 86-90° <li style="width: 33%;">• 41-45° <li style="width: 33%;">• 91-95° <li style="width: 33%;">• 46-50° <li style="width: 33%;">• 96-100° 		
Alignement (Varus-Valgus) : <ul style="list-style-type: none"> <li style="width: 50%;">• 0 <li style="width: 50%;">• 11 <li style="width: 50%;">• 1 <li style="width: 50%;">• 12 <li style="width: 50%;">• 2 <li style="width: 50%;">• 13 <li style="width: 50%;">• 3 <li style="width: 50%;">• 14 <li style="width: 50%;">• 4 <li style="width: 50%;">• 15 <li style="width: 50%;">• 5-10 <li style="width: 50%;">• Over 15° 		
Stability : <ul style="list-style-type: none"> <li style="width: 50%;">• Antero-posterior : <li style="width: 50%;">• Mediolateral : 		

<ul style="list-style-type: none"> ○ <5mm ○ 5-10mm ○ 10+mm 	<ul style="list-style-type: none"> ○ <5° ○ 6-9° ○ 10-14° ○ 15° 	
Part 2- Function score :		
Walking : <ul style="list-style-type: none"> ○ Unlimited ○ >10 blocks ○ 5-10 blocks ○ <5 blocks ○ Housebound ○ Unable 	Stairs : <ul style="list-style-type: none"> ○ Normal Up and down ○ Normal Up down with rail ○ Up and down with rail ○ Up with rail, down unable ○ Unable 	Walking aides used : <ul style="list-style-type: none"> ○ None used ○ Use of Cane/Walking stick deduct ○ Two canes/sticks ○ Crutches frame

1. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res.* 1989;248:13-4.

C. WOMAC score

The WOMAC is a self-questionnaire completed by the patient [1]. It evaluates pain, stiffness and function in patients suffering from lower limb osteoarthritis. There are 24 questions rated from 0 to 4: five about the pain, two about the stiffness and 17 about the function. The total score is calculated by summing the results of each question: 0 is a knee or a hip with normal function, and the maximum score of 96 corresponds to a nonfunctional and very painful joint.

Instructions: please rate the activities in each category according to the following scale of difficulty: 0=none, 1=Slight, 2=moderate, 3=very, 4=extremely.

Circle one number for each activity.

Pain	1. Walking	0	1	2	3	4
	2. Stair Climbing	0	1	2	3	4
	3. Nocturnal	0	1	2	3	4
	4. Rest	0	1	2	3	4
	5. Weight bearing	0	1	2	3	4
Stiffness	1. Morning stiffness	0	1	2	3	4
	2. Stiffness occurring later in the day	0	1	2	3	4
Physical function	1. Descending stairs	0	1	2	3	4
	2. Ascending stairs	0	1	2	3	4
	3. Rising from sitting	0	1	2	3	4
	4. Standing	0	1	2	3	4
	5. Bending to floor	0	1	2	3	4
	6. Walking on flat surface	0	1	2	3	4
	7. Getting in/out of car	0	1	2	3	4
	8. Going shopping	0	1	2	3	4
	9. Putting on socks	0	1	2	3	4
	10. Lying in bed	0	1	2	3	4
	11. Taking off socks	0	1	2	3	4
	12. Rising from bed	0	1	2	3	4
	13. Getting in/out of bath	0	1	2	3	4
	14. Sitting	0	1	2	3	4
	15. Getting on/off toilet	0	1	2	3	4
	16. Heavy domestic duties	0	1	2	3	4
	17. Light domestic duties	0	1	2	3	4
Total Score	/96 = %					

1. McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. *Arthritis Rheum.* 2001;45:453–61.

D. SF-36 score:

The SF-36 (Short Form 36 health survey) is a questionnaire which measures the patient's quality of life [1]. It focuses on the overall quality of life and not on quality of life the operated knee. A short version exists: the SF-12. It evaluates 8 health concepts: physical functioning, pain, role limitations due to physical health problems, role limitations due to personal or emotional problems, emotional well-being, social functioning, energy/fatigue, and general health perceptions. An average is made on each subscale and is compared to the central tendency of the population.

1. In general, would you say your health is:

1 - Excellent

2 - Very good

3 - Good

4 - Fair

5 - Poor

2. **Compared to one year ago**, how would you rate your health in general **now**?

1 - Much better now than one year ago

2 - Somewhat better now than one year ago

3 - About the same

4 - Somewhat worse now than one year ago

5 - Much worse now than one year ago

The following items are about activities you might do during a typical day. Does **your health now limit you** in these activities? If so, how much?

	Yes, limited a lot	Yes, limited a little	No, not limited at all
3. Vigorous activities , such as running, lifting heavy objects, participating in strenuous sports	1	2	3
4. Moderate activities , such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	1	2	3
5. Lifting or carrying groceries	1	2	3
6. Climbing several flights of stairs	1	2	3
7. Climbing one flight of stairs	1	2	3
8. Bending, kneeling, or stooping	1	2	3
9. Walking more than a mile	1	2	3
10. Walking several blocks	1	2	3
11. Walking one block	1	2	3
12. Bathing or dressing yourself	1	2	3

During the **past 4 weeks**, have you had any of the following problems with your work or other regular daily activities **as a result of your physical health**?

	Yes	No
13. Cut down the amount of time you spent on work or other activities	1	2
14. Accomplished less than you would like	1	2
15. Were limited in the kind of work or other activities	1	2
16. Had difficulty performing the work or other activities (for example, it took extra effort)	1	2

During the **past 4 weeks**, have you had any of the following problems with your work or other regular daily activities **as a result of any emotional problems** (such as feeling depressed or anxious)?

	Yes	No
17. Cut down the amount of time you spent on work or other activities	1	2
18. Accomplished less than you would like	1	2
19. Didn't do work or other activities as carefully as usual	1	2

20. During the **past 4 weeks**, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?

1 - Not at all

2 - Slightly

3 - Moderately

4 - Quite a bit

5 - Extremely

21. How much **bodily** pain have you had during the **past 4 weeks**?

1 - None

2 - Very mild

3 - Mild

4 - Moderate

5 - Severe

6 - Very severe

22. During the **past 4 weeks**, how much did **pain** interfere with your normal work (including both work outside the home and housework)?

1 - Not at all

2 - A little bit

3 - Moderately

4 - Quite a bit

5 - Extremely

These questions are about how you feel and how things have been with you **during the past 4 weeks**. For each question, please give the one answer that comes closest to the way you have been feeling.

How much of the time during the **past 4 weeks**...

	All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time
23. Did you feel full of pep?	1	2	3	4	5	6
24. Have you been a very nervous person?	1	2	3	4	5	6
25. Have you felt so down in the dumps that nothing could cheer you up?	1	2	3	4	5	6
26. Have you felt calm and peaceful?	1	2	3	4	5	6
27. Did you have a lot of energy?	1	2	3	4	5	6
28. Have you felt downhearted and blue?	1	2	3	4	5	6
29. Did you feel worn out?	1	2	3	4	5	6
30. Have you been a happy person?	1	2	3	4	5	6
31. Did you feel tired?	1	2	3	4	5	6

32. During the **past 4 weeks**, how much of the time has **your physical health or emotional problems** interfered with your social activities (like visiting with friends, relatives, etc.)?

1 - All of the time

2 - Most of the time

3 - Some of the time

4 - A little of the time

5 - None of the time

How TRUE or FALSE is **each** of the following statements for you.

	Definitely true	Mostly true	Don't know	Mostly false	Definitely false
33. I seem to get sick a little easier than other people	1	2	3	4	5
34. I am as healthy as anybody I know	1	2	3	4	5
35. I expect my health to get worse	1	2	3	4	5
36. My health is excellent	1	2	3	4	5

Appendix 3. Characteristics and references of studies

study	Country	funding	surgical approach	Age(mean)	BMI(mean)	Gender(woman)	Numbers of patients	Numbers of knees	Prosthesis
Aglietti 2006 [1]	Italy	Public	Subvastus QS	70.5	28.75	39	60	60	NexGen <i>(Zimmer Biomet, USA, Warsaw)</i>
Aslam 2016 [2]	India	Public	Midvastus parapatellar	68.7	30.6	30	60	84	Scorpio <i>(Stryker, USA, Kalamazoo)</i>
Bathis 2005 [3]	Germany	No specific funding	Midvastus parapatellar	69.6	29.75	38	50	50	DePuy <i>(DePuy Synthes, USA, Massschusetts)</i>
Bourke 2011 [4]	Australia	No specific funding	Subvastus parapatellar	67.9		45	76	76	Multiple
Bridgman 2008 [5]	England	No specific funding	Subvastus parapatellar	70.5		112	231	231	DePuy <i>(DePuy Synthes, USA, Massschusetts)</i>
Chalidis 2010 [6]	Greece	No specific funding	Midvastus parapatellar	70.7	34.4	90	100	100	Genesis II <i>(Smith & Nephew, UK, London)</i>
Chiang 2012 [7]	Taiwan	No specific funding	QS parapatellar	69.8	29.1	54	60	80	NexGen <i>(Zimmer Biomet, USA, Warsaw)</i>
Chin 2007 [8]	Singapore	No specific funding	Midvastus QS parapatellar	66.6	28.49	75	90	90	Multiple
Cho 2013 [9]	Korea	No specific funding	Midvastus mini parapatellar	66.3	28.6	63	66	66	Stryker Triathlon <i>(Stryker, USA, Kalamazoo)</i>
Chotanaphuti 2018 [10]	Thailand	No specific funding	QS parapatellar	68		33	40	40	-
Dalury 1999 [11]	USA	No specific funding	Midvastus parapatellar	70		16	24	48	-
Dalury 2008 [12]	USA	No specific funding	Midvastus parapatellar	67	30	14	20	40	-
Dutton 2008 [13]	Singapore	No specific funding	mini parapatellar parapatellar	67.5	27.6	88	108	108	DePuy <i>(DePuy Synthes, USA, Massschusetts)</i>
Faure 1993 [14]	USA	No specific funding	Subvastus parapatellar				14	28	-
Feczko 2016 [15]	International	No specific funding	Midvastus parapatellar	65	28.4	45	69	69	Scorpio <i>(Stryker, USA, Kalamazoo)</i>
Guy 2012 [16]	England	No specific funding	Midvastus parapatellar	70.1	28.55	42	80	80	NexGen <i>(Zimmer Biomet, USA, Warsaw)</i>
Halder 2009 [17]	Germany	No specific funding	Subvastus parapatellar				100	100	-
Han 2008 [18]	Korea	Public	mini parapatellar parapatellar	65	26.7	28	30	60	NexGen <i>(Zimmer Biomet, USA, Warsaw)</i>

Heekin 2013 [19]	USA	Private	Midvastus mini parapatellar	65.13	30.98	14	40	80	Stryker Triathlon (Stryker, USA, Kalamazoo)
Hernandez-Vaquero 2010 [20]	Spain	Public	Midvastus parapatellar	70.7	31.4	51	62	62	Stryker Triathlon (Stryker, USA, Kalamazoo)
Jhurani 2019 [21]	India	No specific funding	Subvastus parapatellar	64.6	29.5	124	186	186	Genesis II (Smith & Nephew, UK, London)
Jung 2005 [22]	South Korea	No specific funding	Subvastus parapatellar			24	26	40	DePuy (DePuy Synthes, USA, Massschussetts)
Juosponis 2008 [23]	Lithuania	No specific funding	Midvastus mini parapatellar	71.7	28.6	60	70	70	DePuy (DePuy Synthes, USA, Massschussetts)
Karachalios 2008 [24]	Greece	No specific funding	Midvastus parapatellar	70.9	31.7	66	100	100	Genesis II (Smith & Nephew, UK, London)
Karpman 2009 [25]	USA	Private	QS Midvastus parapatellar	73.3	29	35	59	59	NexGen (Zimmer Biomet, USA, Warsaw)
Keating 1999 [26]	USA	No specific funding	Midvastus parapatellar	70.23	38.8	60	100	200	-
Kim 2007 [27]	Korea	Public	QS parapatellar	65.4	28.1	93	120	240	NexGen (Zimmer Biomet, USA, Warsaw)
Koh 2016 [28]	Korea	Public	Subvastus parapatellar	65	26.8	47	50	100	LOSPA
Kolisek 2007 [29]	USA	Private	Midvastus parapatellar	68.5	31	26	80	80	Scorpio (Stryker, USA, Kalamazoo)
Layher 2015 [30]	Germany	No specific funding	Midvastus parapatellar	64.6	33.1	12	21	21	Multiple
Lee 2010 [31]	Korea	Public	Midvastus mini parapatellar	67	27.5	28	30	60	E-motion (B. Braun, Germany, Melsungen)
Li 2017 [32]	China	No specific funding	Subvastus parapatellar	69	25.66	32	50	50	-
Lin 2012 [33]	Taiwan	No specific funding	QS parapatellar	68.7	25.97	86	100	100	NexGen (Zimmer Biomet, USA, Warsaw)
Liu 2015 [34]	China	Public	Midvastus mini parapatellar				45	90	Genesis II (Smith & Nephew, UK, London)
Luring 2008 [35]	Germany	No specific funding	mini parapatellar parapatellar	69.3	31.3		90	90	DePuy (DePuy Synthes, USA, Massschussetts)
Mehta 2016 [36]	India	No specific funding	QS parapatellar	60.6		40	55	55	-
Nestor 2010 [37]	USA	Private	Midvastus parapatellar	66.7	29.6	18	27	54	-
Nutton 2014 [38]	UK	Private	Midvastus parapatellar	71.3	30.1	12	28	28	DePuy (DePuy Synthes, USA, Massschussetts)
Pan 2010 [39]	China	No specific funding	Subvastus parapatellar	62.8	24.7	48	68	68	NexGen (Zimmer Biomet, USA, Warsaw)

Pescador 2011 [40]	Spain	No specific funding	Subvastus parapatellar	70.2	28.1	68	96	96	Zimmer <i>(Zimmer Biomet, USA, Warsaw)</i>
Pongcharoen 2012 [41]	Thailand	Public	Midvastus mini parapatellar	67	26.4	48	59	60	NexGen <i>(Zimmer Biomet, USA, Warsaw)</i>
Roysam 1999 [42]	UK	No specific funding	Subvastus parapatellar	70		42	89	89	Insall-Burstein II prosthesis <i>(Zimmer Biomet, USA, Warsaw)</i>
Seon 2005 [43]	Korea	No specific funding	Midvastus parapatellar	66	26.25	91	89	102	E-motion <i>(B.Braun, Germany, Melsungen)</i>
Seon 2007 [44]	Korea	No specific funding	Midvastus parapatellar	64.2		33	42	84	E-motion <i>(B.Braun, Germany, Melsungen)</i>
Shukla 2016 [45]	India	No specific funding	Midvastus parapatellar	61.4		30	52	52	Genesis II <i>(Smith & Nephew, UK, London)</i>
Stevens Lapsley 2012 [46]	USA	No specific funding	mini parapatellar parapatellar	64.3	30.9	22	44	44	DePuy <i>(DePuy Synthes, USA, Massschusetts)</i>
Stevens-Lapsley 2012 [47]	USA	Public	mini parapatellar parapatellar	64.3	30.9	22	42	42	DePuy <i>(DePuy Synthes, USA, Massschusetts)</i>
Thienpont 2012 [48]	Belgium	No specific funding	mini parapatellar parapatellar	68.5	30.1	205	300	300	-
Tomek 2014 [49]	USA	Public and private	Subvastus parapatellar	64.3	30.5	83	127	127	LPS legacy <i>(Zimmer Biomet, USA, Warsaw)</i>
Van Hemert 2010 [50]	Netherlands	No specific funding	Subvastus parapatellar	70.6	29.75	27	40	40	Scorpio <i>(Stryker, USA, Kalamazoo)</i>
Varela-Egocheaga 2009 [51]	Spain	No specific funding	Subvastus parapatellar	69.3	30.8	73	100	100	NexGen <i>(Zimmer Biomet, USA, Warsaw)</i>
Verburg 2016 [52]	Netherlands	No specific funding	Midvastus parapatellar	67	28.6	67	97	100	NexGen <i>(Zimmer Biomet, USA, Warsaw)</i>
Walter 2006 [53]	USA	No specific funding	Midvastus parapatellar	69	33.4	89	122	122	Multiple
Wegrzyn 2012 [54]	USA	No specific funding	Subvastus parapatellar	65.5	30.6	28	36	36	Sigma RP knee system
Weinhardt 2004 [55]	Germany	No specific funding	Subvastus parapatellar	71.7		19	52	52	Genesis II <i>(Smith & Nephew, UK, London)</i>
Wulker 2010 [56]	International	Private	Midvastus parapatellar	70.1	29.3	101	134	134	Genesis II <i>(Smith & Nephew, UK, London)</i>
Yang 2010 [57]	Korea	No specific funding	QS parapatellar	68.3	27.8	43	48	75	-
Zhang 2012 [58]	China	No specific funding	Midvastus mini parapatellar	63.1	24.6	61	89	89	Genesis II <i>(Smith & Nephew, UK, London)</i>
Zhang 2013 [59]	China	No specific funding	mini parapatellar parapatellar	65.2	25.5	42	82	82	DePuy <i>(DePuy Synthes, USA, Massschusetts)</i>

Zhu 2015 [60]	Singapore	Public and private	mini parapatellar parapatellar	66.5	27.7	59	67	67	DePuy (DePuy Synthes, USA, Massachusetts)
------------------	-----------	--------------------------	--------------------------------------	------	------	----	----	----	--

References:

1. Aglietti P, Baldini A, Sensi L. Quadriceps-sparing versus mini-subvastus approach in total knee arthroplasty. *Clin Orthop Relat Res.* 2006;452:106-111. doi:[10.1097/01.blo.0000238789.51972.16](https://doi.org/10.1097/01.blo.0000238789.51972.16)
2. Aslam M, Sabir A, Tiwari V, Abbas S, Tiwari A, Singh P. Approach to Total Knee Replacement: a Randomized Double Blind Study between Medial Parapatellar and Midvastus Approach in the Early Postoperative Period in Asian Population. *Journal of knee surgery.* 2016;(no pagination). doi:[10.1055/s-0036-1597978](https://doi.org/10.1055/s-0036-1597978)
3. B athis H, Perlick L, Blum C, L uring C, Perlick C, Grifka J. Midvastus approach in total knee arthroplasty: a randomized, double-blinded study on early rehabilitation. *Knee surgery, sports traumatology, arthroscopy.* 2005;13(7):545-550. doi:[10.1007/s00167-004-0574-9](https://doi.org/10.1007/s00167-004-0574-9)
4. Bourke MG, Sclavos EK, Jull GA, Buttrum PJ, Dalton PA, Russell TG. A comparison of patellar vascularity between the medial parapatellar and subvastus approaches in total knee arthroplasty. *J Arthroplasty.* 2012;27(6):1123-1127.e1. doi:[10.1016/j.arth.2012.01.013](https://doi.org/10.1016/j.arth.2012.01.013)
5. Bridgman S, Walley G, Clement D, Griffiths D, MacKenzie G, Maffulli N. Sub-vastus versus medial parapatellar arthrotomy for total knee arthroplasty (SMAK): a randomised controlled trial. *The journal of bone and joint surgery (proceedings).* 2009;91-B(SUPP_I):31-3e.
6. Chalidis BE, Petsatodis G, Christodoulou AG, Christoforidis J, Papadopoulos PP, Pournaras J. Is obesity a contraindication for minimal invasive total knee replacement? A prospective randomized control trial. *Obes Surg.* 2010;20(12):1633-1641. doi:[10.1007/s11695-009-9968-6](https://doi.org/10.1007/s11695-009-9968-6)
7. Chiang H, Lee C, Lin W, Jiang C. Comparison of quadriceps-sparing minimally invasive and medial parapatellar total knee arthroplasty: a 2-year follow-up study. *Journal of the Formosan Medical Association / Taiwan yi zhi.* 2012;111(12):698-704. doi:[10.1016/j.jfma.2011.11.025](https://doi.org/10.1016/j.jfma.2011.11.025)
8. Chin P.L., Foo L.S.S., Yang K.Y., Yeo S.J., Lo N.N. Randomized Controlled Trial Comparing the Radiologic Outcomes of Conventional and Minimally Invasive Techniques for Total Knee Arthroplasty. *J Arthroplasty.* 2007;22(6):800-806. doi:[10.1016/j.arth.2006.10.009](https://doi.org/10.1016/j.arth.2006.10.009)
9. Cho K-Y, Kim K-I, Umrani S, Kim S-H. Better quadriceps recovery after minimally invasive total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc.* 2014;22(8):1759-1764. doi:[10.1007/s00167-013-2556-2](https://doi.org/10.1007/s00167-013-2556-2)
10. Chotanaphuti T, Ongnamthip P, Karnchanalerk K, Udombuathong P. Comparative study between 2 cm limited quadriceps exposure minimal invasive surgery and conventional total

knee arthroplasty in quadriceps function: prospective randomized controlled trial. *J Med Assoc Thai*. 2008;91(2):203-207.

11. Dalury D, Jiranek W. A comparison of the midvastus and paramedian approaches for total knee arthroplasty. *Journal of arthroplasty*. 1999;14(1):33-37.

12. Dalury DF, Snow RG, Adams MJ. Electromyographic evaluation of the midvastus approach. *J Arthroplasty*. 2008;23(1):136-140. doi:[10.1016/j.arth.2007.01.020](https://doi.org/10.1016/j.arth.2007.01.020)

13. Dutton AQ, Yeo S-J, Yang K-Y, Lo N-N, Chia K-U, Chong H-C. Computer-assisted minimally invasive total knee arthroplasty compared with standard total knee arthroplasty. A prospective, randomized study. *J Bone Joint Surg Am*. 2008;90(1):2-9. doi:[10.2106/JBJS.F.01148](https://doi.org/10.2106/JBJS.F.01148)

14. Faure BT, Benjamin JB, Lindsey B, Volz RG, Schutte D. Comparison of the subvastus and paramedian surgical approaches in bilateral knee arthroplasty. *J Arthroplasty*. 1993;8(5):511-516.

15. Feczko P., Engelmann L., Arts J.J., Campbell D. Computer-assisted total knee arthroplasty using mini midvastus or medial parapatellar approach technique: A prospective, randomized, international multicentre trial Orthopedics and biomechanics. *BMC Musculoskelet Disord*. 2016;17(1). doi:[10.1186/s12891-016-0872-7](https://doi.org/10.1186/s12891-016-0872-7)

16. Guy S.P., Farndon M.A., Conroy J.L., Bennett C., Grainger A.J., London N.J. A prospective randomised study of minimally invasive midvastus total knee arthroplasty compared with standard total knee arthroplasty. *Knee*. 2012;19(6):866-871. doi:[10.1016/j.knee.2012.04.009](https://doi.org/10.1016/j.knee.2012.04.009)

17. Halder A., Beier A., Neumann W. Mini-Subvastus approach for total knee replacement. *Oper Orthop Traumatol*. 2009;21(1):14-24. doi:[10.1007/s00064-009-1602-1](https://doi.org/10.1007/s00064-009-1602-1)

18. Han I, Seong S, Lee S, Yoo J, Lee M. Simultaneous bilateral MIS-TKA results in faster functional recovery. *Clinical orthopaedics and related research*. 2008;466(6):1449-1453. doi:[10.1007/s11999-008-0216-5](https://doi.org/10.1007/s11999-008-0216-5)

19. Heekin RD, Fokin AA. Mini-midvastus versus mini-medial parapatellar approach for minimally invasive total knee arthroplasty: outcomes pendulum is at equilibrium. *J Arthroplasty*. 2014;29(2):339-342. doi:[10.1016/j.arth.2013.05.016](https://doi.org/10.1016/j.arth.2013.05.016)

20. Hernandez-Vaquero D, Noriega-Fernandez A, Suarez-Vazquez A. Total knee arthroplasties performed with a mini-incision or a standard incision. Similar results at six months follow-up. *BMC musculoskeletal disorders*. 2010;11:27. doi:[10.1186/1471-2474-11-27](https://doi.org/10.1186/1471-2474-11-27)

21. Jhurani A, Agarwal P, Aswal M, Rasquinha C, Srivastava M. Subvastus Exposure Compared to Parapatellar Approach in Navigated Sequential Bilateral Total Knee Arthroplasty (SBTKA): A Prospective Randomized Study. *J Knee Surg*. November 2019;s-0039-1700496. doi:[10.1055/s-0039-1700496](https://doi.org/10.1055/s-0039-1700496)

22. Jung W, Ha Y, Cha M. A Comparison of the Midvastus and Median Parapatellar Surgical Approaches in Total Knee Arthroplasty. *Journal of the Korean Orthopaedic Association*. 2007;42(3):354-359. doi:[10.4055/jkoa.2007.42.3.354](https://doi.org/10.4055/jkoa.2007.42.3.354)
23. Juosponis R., Tarasevicius S., Smailys A., Kalesinskas R.J. Functional and radiological outcome after total knee replacement performed with mini-midvastus or conventional arthrotomy: Controlled randomised trial. *Int Orthop*. 2009;33(5):1233-1237. doi:[10.1007/s00264-008-0630-z](https://doi.org/10.1007/s00264-008-0630-z)
24. Karachalios T, Giotikas D, Roidis N, Poultsides L, Bargiotas K, Malizos K. Total knee replacement performed with either a mini-midvastus or a standard approach: a prospective randomised clinical and radiological trial. *Journal of bone and joint surgery British volume*. 2008;90(5):584-591. doi:[10.1302/0301-620X.90B5.20122](https://doi.org/10.1302/0301-620X.90B5.20122)
25. Karpman R.R., Smith H.L. Comparison of the Early Results of Minimally Invasive Vs Standard Approaches to Total Knee Arthroplasty. A Prospective, Randomized Study. *J Arthroplasty*. 2009;24(5):681-688. doi:[10.1016/j.arth.2008.03.011](https://doi.org/10.1016/j.arth.2008.03.011)
26. Keating EM, Faris PM, Meding JB, Ritter MA. Comparison of the midvastus muscle-splitting approach with the median parapatellar approach in total knee arthroplasty. *J Arthroplasty*. 1999;14(1):29-32.
27. Kim Y-H, Kim J-S, Kim D-Y. Clinical outcome and rate of complications after primary total knee replacement performed with quadriceps-sparing or standard arthrotomy. *J Bone Joint Surg Br*. 2007;89(4):467-470. doi:[10.1302/0301-620X.89B4.18663](https://doi.org/10.1302/0301-620X.89B4.18663)
28. Koh IJ, Kim MW, Kim MS, Jang SW, Park DC, In Y. The Patient's Perception Does Not Differ Following Subvastus and Medial Parapatellar Approaches in Total Knee Arthroplasty: A Simultaneous Bilateral Randomized Study. *J Arthroplasty*. 2016;31(1):112-117. doi:[10.1016/j.arth.2015.08.004](https://doi.org/10.1016/j.arth.2015.08.004)
29. Kolisek FR, Bonutti PM, Hozack WJ, et al. Clinical experience using a minimally invasive surgical approach for total knee arthroplasty: early results of a prospective randomized study compared to a standard approach. *J Arthroplasty*. 2007;22(1):8-13. doi:[10.1016/j.arth.2006.06.004](https://doi.org/10.1016/j.arth.2006.06.004)
30. Layher F, Zipfel M, Sander K, Matziolis G, Roth A. Functional Comparison of the Outcome after Midvastus and Medial Parapatellar Surgical Approach in Total Knee Arthroplasty. *Zeitschrift für Orthopädie und Unfallchirurgie*. 2016;154(1):50-57. doi:[10.1055/s-0035-1558074](https://doi.org/10.1055/s-0035-1558074)
31. Lee D, Choi J, Nha K, Kim H, Han S. No difference in early functional outcomes for mini-midvastus and limited medial parapatellar approaches in navigation-assisted total knee arthroplasty: a prospective randomized clinical trial. *Knee surgery, sports traumatology, arthroscopy*. 2011;19(1):66-73. doi:[10.1007/s00167-010-1130-4](https://doi.org/10.1007/s00167-010-1130-4)
32. Li Z, Cheng W, Sun L, et al. Mini-subvastus versus medial parapatellar approach for total knee arthroplasty: a prospective randomized controlled study. *International Orthopaedics*. 2017;1-7. doi:[10.1007/s00264-017-3703-z](https://doi.org/10.1007/s00264-017-3703-z)

33. Lin S-Y, Chen C-H, Fu Y-C, et al. Comparison of the clinical and radiological outcomes of three minimally invasive techniques for total knee replacement at two years. *Bone Joint J.* 2013;95-B(7):906-910. doi:[10.1302/0301-620X.95B7.29694](https://doi.org/10.1302/0301-620X.95B7.29694)
34. Liu H-C, Kuo F-C, Huang C-C, Wang J-W. Mini-midvastus total knee arthroplasty in patients with severe varus deformity. *Orthopedics.* 2015;38(2):e112-117. doi:[10.3928/01477447-20150204-58](https://doi.org/10.3928/01477447-20150204-58)
35. Luring C, Beckmann J, Haibock P, Perlick L, Grifka J, Tingart M. Minimal invasive and computer assisted total knee replacement compared with the conventional technique: a prospective, randomised trial. *Knee Surg Sports Traumatol Arthrosc.* 2008;16(10):928-934. doi:[10.1007/s00167-008-0582-2](https://doi.org/10.1007/s00167-008-0582-2)
36. Mehta N., Bhat M.S., Goyal A., Mishra P., Joshi D., Chaudhary D. Quadriceps sparing (subvastus/midvastus) approach versus the conventional medial parapatellar approach in primary knee arthroplasty. *J Arthrosc Jt Surg.* 2017;4(1):15-20. doi:[10.1016/j.jajs.2017.02.004](https://doi.org/10.1016/j.jajs.2017.02.004)
37. Nestor BJ, Toulson CE, Backus SI, Lyman SL, Foote KL, Windsor RE. Mini-midvastus vs standard medial parapatellar approach: a prospective, randomized, double-blinded study in patients undergoing bilateral total knee arthroplasty. *J Arthroplasty.* 2010;25(6 Suppl):5-11, 11.e1. doi:[10.1016/j.arth.2010.04.003](https://doi.org/10.1016/j.arth.2010.04.003)
38. Nutton R, Wade F, Coutts F, Van Der Linden M. Short term recovery of function following total knee arthroplasty: a randomised study of the medial parapatellar and midvastus approaches. *Arthritis.* 2014. doi:[10.1155/2014/173857](https://doi.org/10.1155/2014/173857)
39. Pan W, Li X, Tang T, Qian Z, Zhang Q, Zhang C. Mini-subvastus versus a standard approach in total knee arthroplasty: a prospective, randomized, controlled study. *Journal of international medical research.* 2010;38(3):890-900. doi:[10.1177/147323001003800315](https://doi.org/10.1177/147323001003800315)
40. Pescador D, Moreno A, Blanco J, García I. Long-term analysis of minimally invasive surgery in knee arthroplasty. *Acta ortopedica mexicana.* 2011;25(6):353-358.
41. Pongcharoen B, Pongcharoen B, Yakampor T, Charoencholvanish K. Patellar tracking and anterior knee pain are similar after medial parapatellar and midvastus approaches in minimally invasive TKA. *Clinical orthopaedics and related research.* 2013;471(5):1654-1660. doi:[10.1007/s11999-012-2778-5](https://doi.org/10.1007/s11999-012-2778-5)
42. Roysam G.S., Oakley M.J. Subvastus approach for total knee arthroplasty: A prospective, randomized, and observer-blinded trial. *J Arthroplasty.* 2001;16(4):454-457. doi:[10.1054/arth.2001.22388](https://doi.org/10.1054/arth.2001.22388)
43. Seon JK, Song EK, Yoon TR, Park SJ, Bae BH, Cho SG. Comparison of functional results with navigation-assisted minimally invasive and conventional techniques in bilateral total knee arthroplasty. *Comput Aided Surg.* 2007;12(3):189-193. doi:[10.3109/10929080701311861](https://doi.org/10.3109/10929080701311861)

44. Seon J, Song E. Navigation-assisted less invasive total knee arthroplasty compared with conventional total knee arthroplasty: a randomized prospective trial. *Journal of arthroplasty*. 2006;21(6):777-782. doi:[10.1016/j.arth.2005.08.024](https://doi.org/10.1016/j.arth.2005.08.024)
45. Shukla R, Mahajan P, Singh M, Jain R, Kumar R. Outcome of Total Knee Replacement via Two Approaches in Indian Scenario. *Journal of knee surgery*. 2017;30(2):174-178. doi:[10.1055/s-0036-1584192](https://doi.org/10.1055/s-0036-1584192)
46. Stevens-Lapsley J, Bade M, Shulman B, Kohrt W, Dayton M. Minimally invasive total knee arthroplasty improves early knee strength but not functional performance: a randomized controlled trial. *Journal of arthroplasty*. 2012;27(10):1812-1819.e2. doi:[10.1016/j.arth.2012.02.016](https://doi.org/10.1016/j.arth.2012.02.016)
47. Dayton MR, Bade MJ, Muratore T, Shulman BC, Kohrt WM, Stevens-Lapsley JE. Minimally invasive total knee arthroplasty: surgical implications for recovery. *J Knee Surg*. 2013;26(3):195-201. doi:[10.1055/s-0032-1327449](https://doi.org/10.1055/s-0032-1327449)
48. Thienpont E. Faster recovery after minimally invasive surgery in total knee arthroplasty. *Knee surgery, sports traumatology, arthroscopy*. 2013;21(10):2412-2417. doi:[10.1007/s00167-012-1978-6](https://doi.org/10.1007/s00167-012-1978-6)
49. Tomek I, Kantor S, Cori L, et al. Early Patient Outcomes After Primary Total Knee Arthroplasty with Quadriceps-Sparing Subvastus and Medial Parapatellar Techniques: a Randomized, Double-Blind Clinical Trial. *Journal of bone and joint surgery American volume*. 2014;96(11):907-915. doi:[10.2106/JBJS.L.01578](https://doi.org/10.2106/JBJS.L.01578)
50. Van Hemert W, Senden R, Grimm B, van der Linde M, Lataster A, Heyligers I. Early functional outcome after subvastus or parapatellar approach in knee arthroplasty is comparable. *Knee surgery, sports traumatology, arthroscopy*. 2011;19(6):943-951. doi:[10.1007/s00167-010-1292-0](https://doi.org/10.1007/s00167-010-1292-0)
51. Varela-Egocheaga JR, Suarez-Suarez MA, Fernandez-Villan M, Gonzalez-Sastre V, Varela-Gomez JR, Rodriguez-Merchan C. Minimally invasive subvastus approach: improving the results of total knee arthroplasty: a prospective, randomized trial. *Clin Orthop Relat Res*. 2010;468(5):1200-1208. doi:[10.1007/s11999-009-1160-8](https://doi.org/10.1007/s11999-009-1160-8)
52. Verburg H, Mathijssen NMC, Niesten D-D, Verhaar JAN, Pilot P. Comparison of Mini-Midvastus and Conventional Total Knee Arthroplasty with Clinical and Radiographic Evaluation: A Prospective Randomized Clinical Trial with 5-Year Follow-up. *J Bone Joint Surg Am*. 2016;98(12):1014-1022. doi:[10.2106/JBJS.15.00654](https://doi.org/10.2106/JBJS.15.00654)
53. Walter F, Haynes M, Markel D. A randomized prospective study evaluating the effect of patellar eversion on the early functional outcomes in primary total knee arthroplasty. *Journal of arthroplasty*. 2007;22(4):509-514. doi:[10.1016/j.arth.2006.06.015](https://doi.org/10.1016/j.arth.2006.06.015)
54. Wegrzyn J, Parratte S, Coleman-Wood K, Kaufman KR, Pagnano MW. The John Insall award: no benefit of minimally invasive TKA on gait and strength outcomes: a randomized controlled trial. *Clin Orthop Relat Res*. 2013;471(1):46-55. doi:[10.1007/s11999-012-2486-1](https://doi.org/10.1007/s11999-012-2486-1)

55. Weinhardt C, Barisic M, Bergmann EG, Heller KD. Early results of subvastus versus medial parapatellar approach in primary total knee arthroplasty. *Arch Orthop Trauma Surg.* 2004;124(6):401-403. doi:[10.1007/s00402-004-0692-3](https://doi.org/10.1007/s00402-004-0692-3)
56. Wulker N, Lambermont JP, Sacchetti L, Lazaro JG, Nardi J. A prospective randomized study of minimally invasive total knee arthroplasty compared with conventional surgery. *J Bone Joint Surg Am.* 2010;92(7):1584-1590. doi:[10.2106/JBJS.H.01070](https://doi.org/10.2106/JBJS.H.01070)
57. Yang J, Yoon J, Pandher D, Oh K. Clinical and radiologic outcomes of contemporary 3 techniques of TKA. *Orthopedics.* 2010;33(10 Suppl):76-81. doi:[10.3928/01477447-20100510-59](https://doi.org/10.3928/01477447-20100510-59)
58. Zhang Z, Gu B, Zhu W, Zhu L, Li Q, Du Y. Minimally invasive and computer-assisted total knee arthroplasty versus conventional technique: a prospective, randomized study. *European journal of orthopaedic surgery & traumatology : orthopedie traumatologie.* 2013;24(8):1475-1479. doi:[10.1007/s00590-013-1313-z](https://doi.org/10.1007/s00590-013-1313-z)
59. Zhang Z, Zhu W, Gu B, Zhu L, Chen C. Mini-midvastus versus mini-medial parapatellar approach in total knee arthroplasty: a prospective, randomized study. *Archives of orthopaedic and trauma surgery.* 2013;133(3):389-395. doi:[10.1007/s00402-012-1645-x](https://doi.org/10.1007/s00402-012-1645-x)
60. Zhu M, Ang C, Yeo S, Lo N, Chia S, Chong H. Minimally Invasive Computer-Assisted Total Knee Arthroplasty Compared With Conventional Total Knee Arthroplasty: a Prospective 9-Year Follow-Up. *Journal of arthroplasty.* 2016;31(5):1000-1004. doi:[10.1016/j.arth.2015.11.023](https://doi.org/10.1016/j.arth.2015.11.023)

Heekin 2013 [19]	moderate	moderate	low	high	moderate	NA	low	low	NA	low	low	low
Hernandez-Vaquero 2010 [20]	low	low	moderate	high	high	NA	low	low	NA	low	low	low
Jhurani 2019 [21]	low	moderate	moderate	high	low	low	low	low	low	low	low	low
Jung 2005 [22]	moderate	moderate	low	high	low	NA	low	NA	NA	low	NA	low
Juosponis 2008 [23]	moderate	moderate	moderate	high	high	NA	low	low	NA	low	low	low
Karachalios 2008 [24]	low	low	low	high	low	low	low	low	low	low	low	low
Karpman 2009 [25]	low	moderate	low	high	low	low	low	low	low	low	low	low
Keating 1999 [26]	moderate	moderate	low	high	high	NA	low	NA	NA	low	NA	low
Kim 2007 [27]	low	moderate	low	high	low	low	low	low	low	low	low	low
Koh 2016 [28]	low	low	low	high	low	NA	low	low	NA	low	low	low
Kolisek 2007 [29]	low	moderate	moderate	high	moderate	low	low	low	low	low	low	low
Layher 2015 [30]	low	low	low	high	high	high	high	high	low	low	low	low
Lee 2010 [31]	low	low	low	high	low	low	NA	low	low	NA	low	low
Li 2017 [32]	low	moderate	moderate	high	low	low	low	low	low	low	low	low
Lin 2012 [33]	low	low	high	high	low	low	low	low	low	low	low	low
Liu 2015 [34]	high	high	high	high	high	low	low	low	low	low	low	low
Luring 2008 [35]	low	low	moderate	high	low	low	low	low	low	low	low	low
Mehta 2016 [36]	low	low	moderate	high	moderate	low	low	low	low	low	low	low
Nestor 2010 [37]	moderate	moderate	low	high	low	low	low	low	low	low	low	low
Nutton 2014 [38]	low	low	low	high	low	NA	low	low	NA	low	low	low
Pan 2010 [39]	low	low	moderate	high	low	low	low	low	low	low	low	low
Pescador 2011 [40]	moderate	moderate	moderate	high	moderate	NA	NA	low	NA	NA	low	low

References:

1. Aglietti P, Baldini A, Sensi L. Quadriceps-sparing versus mini-subvastus approach in total knee arthroplasty. *Clin Orthop Relat Res*. 2006;452:106-111. doi:[10.1097/01.blo.0000238789.51972.16](https://doi.org/10.1097/01.blo.0000238789.51972.16)
2. Aslam M, Sabir A, Tiwari V, Abbas S, Tiwari A, Singh P. Approach to Total Knee Replacement: a Randomized Double Blind Study between Medial Parapatellar and Midvastus Approach in the Early Postoperative Period in Asian Population. *Journal of knee surgery*. 2016;(no pagination). doi:[10.1055/s-0036-1597978](https://doi.org/10.1055/s-0036-1597978)
3. B athis H, Perlick L, Blum C, L uring C, Perlick C, Grifka J. Midvastus approach in total knee arthroplasty: a randomized, double-blinded study on early rehabilitation. *Knee surgery, sports traumatology, arthroscopy*. 2005;13(7):545-550. doi:[10.1007/s00167-004-0574-9](https://doi.org/10.1007/s00167-004-0574-9)
4. Bourke MG, Sclavos EK, Jull GA, Buttrum PJ, Dalton PA, Russell TG. A comparison of patellar vascularity between the medial parapatellar and subvastus approaches in total knee arthroplasty. *J Arthroplasty*. 2012;27(6):1123-1127.e1. doi:[10.1016/j.arth.2012.01.013](https://doi.org/10.1016/j.arth.2012.01.013)
5. Bridgman S, Walley G, Clement D, Griffiths D, MacKenzie G, Maffulli N. Sub-vastus versus medial parapatellar arthrotomy for total knee arthroplasty (SMAK): a randomised controlled trial. *The journal of bone and joint surgery (proceedings)*. 2009;91-B(SUPP_I):31-3e.
6. Chalidis BE, Petsatodis G, Christodoulou AG, Christoforidis J, Papadopoulos PP, Pournaras J. Is obesity a contraindication for minimal invasive total knee replacement? A prospective randomized control trial. *Obes Surg*. 2010;20(12):1633-1641. doi:[10.1007/s11695-009-9968-6](https://doi.org/10.1007/s11695-009-9968-6)
7. Chiang H, Lee C, Lin W, Jiang C. Comparison of quadriceps-sparing minimally invasive and medial parapatellar total knee arthroplasty: a 2-year follow-up study. *Journal of the Formosan Medical Association / Taiwan yi zhi*. 2012;111(12):698-704. doi:[10.1016/j.jfma.2011.11.025](https://doi.org/10.1016/j.jfma.2011.11.025)
8. Chin P.L., Foo L.S.S., Yang K.Y., Yeo S.J., Lo N.N. Randomized Controlled Trial Comparing the Radiologic Outcomes of Conventional and Minimally Invasive Techniques for Total Knee Arthroplasty. *J Arthroplasty*. 2007;22(6):800-806. doi:[10.1016/j.arth.2006.10.009](https://doi.org/10.1016/j.arth.2006.10.009)
9. Cho K-Y, Kim K-I, Umrani S, Kim S-H. Better quadriceps recovery after minimally invasive total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc*. 2014;22(8):1759-1764. doi:[10.1007/s00167-013-2556-2](https://doi.org/10.1007/s00167-013-2556-2)
10. Chotaphuti T, Ongnamthip P, Karnchanalerk K, Udombuathong P. Comparative study between 2 cm limited quadriceps exposure minimal invasive surgery and conventional total knee arthroplasty in quadriceps function: prospective randomized controlled trial. *J Med Assoc Thai*. 2008;91(2):203-207.
11. Dalury D, Jiranek W. A comparison of the midvastus and paramedian approaches for total knee arthroplasty. *Journal of arthroplasty*. 1999;14(1):33-37.

12. Dalury DF, Snow RG, Adams MJ. Electromyographic evaluation of the midvastus approach. *J Arthroplasty*. 2008;23(1):136-140. doi:[10.1016/j.arth.2007.01.020](https://doi.org/10.1016/j.arth.2007.01.020)
13. Dutton AQ, Yeo S-J, Yang K-Y, Lo N-N, Chia K-U, Chong H-C. Computer-assisted minimally invasive total knee arthroplasty compared with standard total knee arthroplasty. A prospective, randomized study. *J Bone Joint Surg Am*. 2008;90(1):2-9. doi:[10.2106/JBJS.F.01148](https://doi.org/10.2106/JBJS.F.01148)
14. Faure BT, Benjamin JB, Lindsey B, Volz RG, Schutte D. Comparison of the subvastus and paramedian surgical approaches in bilateral knee arthroplasty. *J Arthroplasty*. 1993;8(5):511-516.
15. Feczko P., Engelmann L., Arts J.J., Campbell D. Computer-assisted total knee arthroplasty using mini midvastus or medial parapatellar approach technique: A prospective, randomized, international multicentre trial Orthopedics and biomechanics. *BMC Musculoskelet Disord*. 2016;17(1). doi:[10.1186/s12891-016-0872-7](https://doi.org/10.1186/s12891-016-0872-7)
16. Guy S.P., Farndon M.A., Conroy J.L., Bennett C., Grainger A.J., London N.J. A prospective randomised study of minimally invasive midvastus total knee arthroplasty compared with standard total knee arthroplasty. *Knee*. 2012;19(6):866-871. doi:[10.1016/j.knee.2012.04.009](https://doi.org/10.1016/j.knee.2012.04.009)
17. Halder A., Beier A., Neumann W. Mini-Subvastus approach for total knee replacement. *Oper Orthop Traumatol*. 2009;21(1):14-24. doi:[10.1007/s00064-009-1602-1](https://doi.org/10.1007/s00064-009-1602-1)
18. Han I, Seong S, Lee S, Yoo J, Lee M. Simultaneous bilateral MIS-TKA results in faster functional recovery. *Clinical orthopaedics and related research*. 2008;466(6):1449-1453. doi:[10.1007/s11999-008-0216-5](https://doi.org/10.1007/s11999-008-0216-5)
19. Heekin RD, Fokin AA. Mini-midvastus versus mini-medial parapatellar approach for minimally invasive total knee arthroplasty: outcomes pendulum is at equilibrium. *J Arthroplasty*. 2014;29(2):339-342. doi:[10.1016/j.arth.2013.05.016](https://doi.org/10.1016/j.arth.2013.05.016)
20. Hernandez-Vaquero D, Noriega-Fernandez A, Suarez-Vazquez A. Total knee arthroplasties performed with a mini-incision or a standard incision. Similar results at six months follow-up. *BMC musculoskeletal disorders*. 2010;11:27. doi:[10.1186/1471-2474-11-27](https://doi.org/10.1186/1471-2474-11-27)
21. Jhurani A, Agarwal P, Aswal M, Rasquinha C, Srivastava M. Subvastus Exposure Compared to Parapatellar Approach in Navigated Sequential Bilateral Total Knee Arthroplasty (SBTKA): A Prospective Randomized Study. *J Knee Surg*. November 2019;s-0039-1700496. doi:[10.1055/s-0039-1700496](https://doi.org/10.1055/s-0039-1700496)
22. Jung W, Ha Y, Cha M. A Comparison of the Midvastus and Median Parapatellar Surgical Approaches in Total Knee Arthroplasty. *Journal of the korean orthopaedic association*. 2007;42(3):354-359. doi:[10.4055/jkoa.2007.42.3.354](https://doi.org/10.4055/jkoa.2007.42.3.354)
23. Juosponis R., Tarasevicius S., Smailys A., Kalesinskas R.J. Functional and radiological outcome after total knee replacement performed with mini-midvastus or conventional

arthrotomy: Controlled randomised trial. *Int Orthop*. 2009;33(5):1233-1237.
doi:[10.1007/s00264-008-0630-z](https://doi.org/10.1007/s00264-008-0630-z)

24. Karachalios T, Giotikas D, Roidis N, Poultzides L, Bargiotas K, Malizos K. Total knee replacement performed with either a mini-midvastus or a standard approach: a prospective randomised clinical and radiological trial. *Journal of bone and joint surgery British volume*. 2008;90(5):584-591. doi:[10.1302/0301-620X.90B5.20122](https://doi.org/10.1302/0301-620X.90B5.20122)

25. Karpman R.R., Smith H.L. Comparison of the Early Results of Minimally Invasive Vs Standard Approaches to Total Knee Arthroplasty. A Prospective, Randomized Study. *J Arthroplasty*. 2009;24(5):681-688. doi:[10.1016/j.arth.2008.03.011](https://doi.org/10.1016/j.arth.2008.03.011)

26. Keating EM, Faris PM, Meding JB, Ritter MA. Comparison of the midvastus muscle-splitting approach with the median parapatellar approach in total knee arthroplasty. *J Arthroplasty*. 1999;14(1):29-32.

27. Kim Y-H, Kim J-S, Kim D-Y. Clinical outcome and rate of complications after primary total knee replacement performed with quadriceps-sparing or standard arthrotomy. *J Bone Joint Surg Br*. 2007;89(4):467-470. doi:[10.1302/0301-620X.89B4.18663](https://doi.org/10.1302/0301-620X.89B4.18663)

28. Koh IJ, Kim MW, Kim MS, Jang SW, Park DC, In Y. The Patient's Perception Does Not Differ Following Subvastus and Medial Parapatellar Approaches in Total Knee Arthroplasty: A Simultaneous Bilateral Randomized Study. *J Arthroplasty*. 2016;31(1):112-117.
doi:[10.1016/j.arth.2015.08.004](https://doi.org/10.1016/j.arth.2015.08.004)

29. Kolisek FR, Bonutti PM, Hozack WJ, et al. Clinical experience using a minimally invasive surgical approach for total knee arthroplasty: early results of a prospective randomized study compared to a standard approach. *J Arthroplasty*. 2007;22(1):8-13.
doi:[10.1016/j.arth.2006.06.004](https://doi.org/10.1016/j.arth.2006.06.004)

30. Layher F, Zipfel M, Sander K, Matziolis G, Roth A. Functional Comparison of the Outcome after Midvastus and Medial Parapatellar Surgical Approach in Total Knee Arthroplasty. *Zeitschrift fur orthopadie und unfallchirurgie*. 2016;154(1):50-57.
doi:[10.1055/s-0035-1558074](https://doi.org/10.1055/s-0035-1558074)

31. Lee D, Choi J, Nha K, Kim H, Han S. No difference in early functional outcomes for mini-midvastus and limited medial parapatellar approaches in navigation-assisted total knee arthroplasty: a prospective randomized clinical trial. *Knee surgery, sports traumatology, arthroscopy*. 2011;19(1):66-73. doi:[10.1007/s00167-010-1130-4](https://doi.org/10.1007/s00167-010-1130-4)

32. Li Z, Cheng W, Sun L, et al. Mini-subvastus versus medial parapatellar approach for total knee arthroplasty: a prospective randomized controlled study. *International orthopaedics*. 2017;1-7. doi:[10.1007/s00264-017-3703-z](https://doi.org/10.1007/s00264-017-3703-z)

33. Lin S-Y, Chen C-H, Fu Y-C, et al. Comparison of the clinical and radiological outcomes of three minimally invasive techniques for total knee replacement at two years. *Bone Joint J*. 2013;95-B(7):906-910. doi:[10.1302/0301-620X.95B7.29694](https://doi.org/10.1302/0301-620X.95B7.29694)

34. Liu H-C, Kuo F-C, Huang C-C, Wang J-W. Mini-midvastus total knee arthroplasty in patients with severe varus deformity. *Orthopedics*. 2015;38(2):e112-117. doi:[10.3928/01477447-20150204-58](https://doi.org/10.3928/01477447-20150204-58)
35. Luring C, Beckmann J, Haibock P, Perlick L, Grifka J, Tingart M. Minimal invasive and computer assisted total knee replacement compared with the conventional technique: a prospective, randomised trial. *Knee Surg Sports Traumatol Arthrosc*. 2008;16(10):928-934. doi:[10.1007/s00167-008-0582-2](https://doi.org/10.1007/s00167-008-0582-2)
36. Mehta N., Bhat M.S., Goyal A., Mishra P., Joshi D., Chaudhary D. Quadriceps sparing (subvastus/midvastus) approach versus the conventional medial parapatellar approach in primary knee arthroplasty. *J Arthrosc Jt Surg*. 2017;4(1):15-20. doi:[10.1016/j.jajs.2017.02.004](https://doi.org/10.1016/j.jajs.2017.02.004)
37. Nestor BJ, Toulson CE, Backus SI, Lyman SL, Foote KL, Windsor RE. Mini-midvastus vs standard medial parapatellar approach: a prospective, randomized, double-blinded study in patients undergoing bilateral total knee arthroplasty. *J Arthroplasty*. 2010;25(6 Suppl):5-11, 11.e1. doi:[10.1016/j.arth.2010.04.003](https://doi.org/10.1016/j.arth.2010.04.003)
38. Nutton R, Wade F, Coutts F, Van Der Linden M. Short term recovery of function following total knee arthroplasty: a randomised study of the medial parapatellar and midvastus approaches. *Arthritis*. 2014. doi:[10.1155/2014/173857](https://doi.org/10.1155/2014/173857)
39. Pan W, Li X, Tang T, Qian Z, Zhang Q, Zhang C. Mini-subvastus versus a standard approach in total knee arthroplasty: a prospective, randomized, controlled study. *Journal of international medical research*. 2010;38(3):890-900. doi:[10.1177/147323001003800315](https://doi.org/10.1177/147323001003800315)
40. Pescador D, Moreno A, Blanco J, García I. Long-term analysis of minimally invasive surgery in knee arthroplasty. *Acta ortopedica mexicana*. 2011;25(6):353-358.
41. Pongcharoen B, Pongcharoen B, Yakampor T, Charoencholvanish K. Patellar tracking and anterior knee pain are similar after medial parapatellar and midvastus approaches in minimally invasive TKA. *Clinical orthopaedics and related research*. 2013;471(5):1654-1660. doi:[10.1007/s11999-012-2778-5](https://doi.org/10.1007/s11999-012-2778-5)
42. Roysam G.S., Oakley M.J. Subvastus approach for total knee arthroplasty: A prospective, randomized, and observer-blinded trial. *J Arthroplasty*. 2001;16(4):454-457. doi:[10.1054/arth.2001.22388](https://doi.org/10.1054/arth.2001.22388)
43. Seon JK, Song EK, Yoon TR, Park SJ, Bae BH, Cho SG. Comparison of functional results with navigation-assisted minimally invasive and conventional techniques in bilateral total knee arthroplasty. *Comput Aided Surg*. 2007;12(3):189-193. doi:[10.3109/10929080701311861](https://doi.org/10.3109/10929080701311861)
44. Seon J, Song E. Navigation-assisted less invasive total knee arthroplasty compared with conventional total knee arthroplasty: a randomized prospective trial. *Journal of arthroplasty*. 2006;21(6):777-782. doi:[10.1016/j.arth.2005.08.024](https://doi.org/10.1016/j.arth.2005.08.024)

45. Shukla R, Mahajan P, Singh M, Jain R, Kumar R. Outcome of Total Knee Replacement via Two Approaches in Indian Scenario. *Journal of knee surgery*. 2017;30(2):174-178. doi:[10.1055/s-0036-1584192](https://doi.org/10.1055/s-0036-1584192)
46. Stevens-Lapsley J, Bade M, Shulman B, Kohrt W, Dayton M. Minimally invasive total knee arthroplasty improves early knee strength but not functional performance: a randomized controlled trial. *Journal of arthroplasty*. 2012;27(10):1812-1819.e2. doi:[10.1016/j.arth.2012.02.016](https://doi.org/10.1016/j.arth.2012.02.016)
47. Dayton MR, Bade MJ, Muratore T, Shulman BC, Kohrt WM, Stevens-Lapsley JE. Minimally invasive total knee arthroplasty: surgical implications for recovery. *J Knee Surg*. 2013;26(3):195-201. doi:[10.1055/s-0032-1327449](https://doi.org/10.1055/s-0032-1327449)
48. Thienpont E. Faster recovery after minimally invasive surgery in total knee arthroplasty. *Knee surgery, sports traumatology, arthroscopy*. 2013;21(10):2412-2417. doi:[10.1007/s00167-012-1978-6](https://doi.org/10.1007/s00167-012-1978-6)
49. Tomek I, Kantor S, Cori L, et al. Early Patient Outcomes After Primary Total Knee Arthroplasty with Quadriceps-Sparing Subvastus and Medial Parapatellar Techniques: a Randomized, Double-Blind Clinical Trial. *Journal of bone and joint surgery American volume*. 2014;96(11):907-915. doi:[10.2106/JBJS.L.01578](https://doi.org/10.2106/JBJS.L.01578)
50. Van Hemert W, Senden R, Grimm B, van der Linde M, Lataster A, Heyligers I. Early functional outcome after subvastus or parapatellar approach in knee arthroplasty is comparable. *Knee surgery, sports traumatology, arthroscopy*. 2011;19(6):943-951. doi:[10.1007/s00167-010-1292-0](https://doi.org/10.1007/s00167-010-1292-0)
51. Varela-Egocheaga JR, Suarez-Suarez MA, Fernandez-Villan M, Gonzalez-Sastre V, Varela-Gomez JR, Rodriguez-Merchan C. Minimally invasive subvastus approach: improving the results of total knee arthroplasty: a prospective, randomized trial. *Clin Orthop Relat Res*. 2010;468(5):1200-1208. doi:[10.1007/s11999-009-1160-8](https://doi.org/10.1007/s11999-009-1160-8)
52. Verburg H, Mathijssen NMC, Niesten D-D, Verhaar JAN, Pilot P. Comparison of Mini-Midvastus and Conventional Total Knee Arthroplasty with Clinical and Radiographic Evaluation: A Prospective Randomized Clinical Trial with 5-Year Follow-up. *J Bone Joint Surg Am*. 2016;98(12):1014-1022. doi:[10.2106/JBJS.15.00654](https://doi.org/10.2106/JBJS.15.00654)
53. Walter F, Haynes M, Markel D. A randomized prospective study evaluating the effect of patellar eversion on the early functional outcomes in primary total knee arthroplasty. *Journal of arthroplasty*. 2007;22(4):509-514. doi:[10.1016/j.arth.2006.06.015](https://doi.org/10.1016/j.arth.2006.06.015)
54. Wegrzyn J, Parratte S, Coleman-Wood K, Kaufman KR, Pagnano MW. The John Insall award: no benefit of minimally invasive TKA on gait and strength outcomes: a randomized controlled trial. *Clin Orthop Relat Res*. 2013;471(1):46-55. doi:[10.1007/s11999-012-2486-1](https://doi.org/10.1007/s11999-012-2486-1)
55. Weinhardt C, Barisic M, Bergmann EG, Heller KD. Early results of subvastus versus medial parapatellar approach in primary total knee arthroplasty. *Arch Orthop Trauma Surg*. 2004;124(6):401-403. doi:[10.1007/s00402-004-0692-3](https://doi.org/10.1007/s00402-004-0692-3)

56. Wulker N, Lambermont JP, Sacchetti L, Lazaro JG, Nardi J. A prospective randomized study of minimally invasive total knee arthroplasty compared with conventional surgery. *J Bone Joint Surg Am.* 2010;92(7):1584-1590. doi:[10.2106/JBJS.H.01070](https://doi.org/10.2106/JBJS.H.01070)
57. Yang J, Yoon J, Pandher D, Oh K. Clinical and radiologic outcomes of contemporary 3 techniques of TKA. *Orthopedics.* 2010;33(10 Suppl):76-81. doi:[10.3928/01477447-20100510-59](https://doi.org/10.3928/01477447-20100510-59)
58. Zhang Z, Gu B, Zhu W, Zhu L, Li Q, Du Y. Minimally invasive and computer-assisted total knee arthroplasty versus conventional technique: a prospective, randomized study. *European journal of orthopaedic surgery & traumatology : orthopedie traumatologie.* 2013;24(8):1475-1479. doi:[10.1007/s00590-013-1313-z](https://doi.org/10.1007/s00590-013-1313-z)
59. Zhang Z, Zhu W, Gu B, Zhu L, Chen C. Mini-midvastus versus mini-medial parapatellar approach in total knee arthroplasty: a prospective, randomized study. *Archives of orthopaedic and trauma surgery.* 2013;133(3):389-395. doi:[10.1007/s00402-012-1645-x](https://doi.org/10.1007/s00402-012-1645-x)
60. Zhu M, Ang C, Yeo S, Lo N, Chia S, Chong H. Minimally Invasive Computer-Assisted Total Knee Arthroplasty Compared With Conventional Total Knee Arthroplasty: a Prospective 9-Year Follow-Up. *Journal of arthroplasty.* 2016;31(5):1000-1004. doi:[10.1016/j.arth.2015.11.023](https://doi.org/10.1016/j.arth.2015.11.023)

Appendix 5. Formulas used

When we had only the first quartile and the third quartile, we used the estimator of the Cochrane Handbook to obtain a estimation of the SD:

$$SD \approx \frac{q3-q1}{1.35}$$

When we had only the range, we used this formula to obtain a estimation of the SD [1]:

Sample size	$n \leq 15$	$15 < n \leq 70$	$70 < n$
Estimate SD	$s^2 \approx \frac{1}{12} \left(\frac{(a-2m+b)^2}{4} + (b-a)^2 \right)$	Range/4	Range/6

When we had, only, the median, we used this formula to obtain a estimation of the mean [1]:

Sample size	$15 < n \leq 25$	$25 < n$
Estimate mean	$\bar{x} = \frac{a+2m+b}{4}$	Median

With:

m = Median

a = The smallest value (minimum)

b = The largest value (maximum)

n = The size of the sample.

To estimate the SD of the difference between baseline score and post-surgery score, we used this formula:

$$sd_t = \sqrt{sd_1^2 + sd_2^2}$$

With:

- sd_t : SD of the difference between standard deviation of the preoperative score and SD of the postoperative score.
- sd_1 : SD of the preoperative score.
- sd_2 : SD of the post-surgery score.

1. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005;5:13.

Appendix 6. Pairwise comparisons; standard meta-analyses

KSS at the first 6 months post-surgery: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar	
Trials included/total available trials	3 of 7		6 of 22		3 of 8		4 of 8		6 of 16	
Patient included in Meta-Analysis (MA)	216		394		457		470		663	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	3.6 (-21.0 to 28.0)	2.5 (-1.4 to 6.3)	-4.0 (-22.0 to 14.0)	-4.0 (-8.9 to 0.9)	-12.0 (-36.0 to 13.0)	-17.0 (-21.0 to -14.0)	-10.0 (-31.0 to 11.0)	2.0 (-0.13 to 4.1)	4.3 (-13.0 to 22.0)	3.6 (0.4 to 6.7)
I ²	0.0%	0.0%	45.7%	45.4%	99.4%	99.3%	99.1%	98.6%	83.9%	83.7%

KSS within 1 and 2 years post-surgery: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar	
Trials included/total available trials	4 of 7		3 of 22		2 of 8		2 of 8		6 of 16	
Patient included in Meta-Analysis (MA)	296		286		382		340		698	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	-1.0 (-11.0 to 9.4)	-0.6 (-2.0 to 0.7)	-4.2 (-17.0 to 8.6)	-4.1 (-9.5 to 1.4)	-20.0 (-35.0 to -5.8)	-24.0 (-27.0 to -21.0)	3.8 (-11.0 to 18.0)	5.2 (2.7 to 7.7)	2.2 (-7.1 to 11.0)	1.0 (-2.1 to 4.1)
I ²	0.0%	0.0%	0.0%	0.0%	99.2%	99.1%	69.4%	49.5%	68.9%	67.8%

WOMAC at the first 6 months post-surgery: Pairwise

Treatment comparisons	Parapatellar vs midvastus		Parapatellar vs Mini-parapatellar		Subvastus vs parapatellar	
Trials included/total available trials	2 of 22		2 of 8		2 of 16	
Patient included in Meta-Analysis (MA)	286		382		698	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	2.0 (-9.1 to 13.0)	1.8 (-5.5 to 9.1)	1.5 (-11.0 to 11.0)	3.3 (-2.5 to 9.1)	-1.2 (-2.8 to 0.4)	1.0 (-2.1 to 4.1)
I ²	0.0%	0.0%	78.8%	76.9%	62.2%	67.8%

WOMAC within one- and two-years post-surgery: Pairwise

Treatment comparisons	Subvastus vs parapatellar	
Trials included/total available trials	1 of 16	
Patient included in Meta-Analysis (MA)	222	
Fixed or random effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	3.1 (-3.0 to 9.4)	3.1 (-1.9 to 8.2)
I ²	50.0%	50.0%

SF36 physical scale at the first 6 months: Pairwise

Treatment comparisons	Mini-parapatellar vs parapatellar	Subvastus vs parapatellar
Trials included/total available trials	2 of 7	2 of 16
Patient included in Meta-Analysis (MA)	98	255

Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	7.5 (-5.9 to 19.0)	11.0 (8.5 to 14.0)	1.0 (-12.0 to 14.0)	0.6 (-4.7 to 5.9)
I ²	94.5%	90.8%	0.0%	0.0%

SF36 physical scale within one- and two-years post-surgery: Pairwise

Treatment comparisons	Subvastus vs parapatellar	
Trials included/total available trials	1 of 16	
Patient included in Meta-Analysis (MA)	222	
Fixed or random effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	7.8 (-4.6 to 20.0)	7.9 (-0.1 to 15.9)
I ²	50.0%	50.0%

SF36 health perception scale at the first 6 months: Pairwise

Treatment comparisons	Subvastus vs parapatellar	
Trials included/total available trials	1 of 16	
Patient included in Meta-Analysis (MA)	219	
Fixed or random effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	-1.0 (-8.3 to 6.1)	-0.9 (-8.2 to 6.4)
I ²	47.0%	48.0%

SF36 health perception scale within one- and two-years post-surgery: Pairwise

Treatment comparisons	Subvastus vs parapatellar	
Trials included/total available trials	1 of 16	
Patient included in Meta-Analysis (MA)	222	
Fixed or random effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	7.1 (-4.8 to 18.9)	7.2 (-1.2 to 15.6)
I ²	50.0%	50.0%

ROM at 6 months post-surgery: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar	
Trials included/total available trials	3 of 7		5 of 22		2 of 8		3 of 8		4 of 16	
Patient included in MA	230		370		71		395		264	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	-3.5 (-9.0 to 1.9)	-3.5 (-7.0 to 0.1)	-1.9 (-6.9 to 3.3)	-1.9 (-5.5 to 1.6)	3.6 (-7.0 to 13.0)	3.4 (-6.0 to 13.0)	-0.4 (-5.5 to 4.7)	-0.3 (-3.2 to 2.6)	8.8 (2.8 to 14.0)	9.2 (4.8 to 14.0)
I ²	0.0%	0.0%	0.0%	0.0%	89.2%	89.3%	0.0%	0.0%	65.1%	65.4%

ROM within 1 and 2 years post-surgery: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus	Parapatellar vs midvastus	Parapatellar vs Mini-parapatellar	Quadriceps sparring vs parapatellar	Subvastus vs parapatellar
Trials included/total available trials	1 of 7	2 of 22	1 of 8	2 of 8	5 of 16

Patient included in MA	80		218		82		340		442	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	-2.1 (-18.0 to 14.0)	-2.2 (-8.4 to 3.9)	-2.8 (-18.0 to 13.0)	-3.1 (-14.0 to 7.5)	-15.0 (-30.0 to 0.2)	-15.0 (-19.0 to -11.0)	-0.7 (-12.0 to 10.0)	-0.6 (-4.5 to 3.3)	5.8 (-1.7 to 13.0)	5.7 (2.8 to 8.6)
I ²	-	-	0.0%	0.0%	-	-	0.0%	0.0%	72.4%	72.4%

HKA outliers +/- 3°: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Quadriceps sparring vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar 6 of 16 663	
Trials included/total available trials	2 of 7		5 of 22		1 of 2		3 of 8		4 of 8			
Patient included in Meta-Analysis (MA)	149		366		90		280		340		336	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	0.0 (0.0 to 0.5)	0.0 (0.0 to 0.3)	1.0 (0.4 to 12.6)	1.0 (0.6 to 1.6)	2.4 (0.4 to 14.0)	2.3 (0.9 to 6.0)	1.6 (0.5 to 4.3)	1.7 (1.0 to 2.9)	1.9 (0.7 to 6.2)	1.7 (1.0 to 3.1)	0.3 (0.1 to 1.5)	0.4 (0.1 to 1.1)
I ²	0.0%	0.0%	39.2%	39.8%	-	-	61.0%	62.3%	5.0%	0.0%	43.0%	38.0%

Sagittal frontal outliers +/-3°: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus	Parapatellar vs midvastus	Parapatellar vs Mini-parapatellar	Quadriceps sparring vs parapatellar
Trials included/total available trials	2 of 7	2 of 22	1 of 8	1 of 8
Patient included in Meta-Analysis (MA)	149	134	108	240

Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	0.0 (0.0 to 0.5)	0.0 (0.0 to 0.5)	0.4 (0.0 to 16.5)	0.4 (0.0 to 5.4)	1.4 (0.3 to 6.4)	1.4 (0.8 to 2.8)	1.2 (0.3 to 5.9)	1.2 (0.5 to 2.8)
I ²	0.0%	0.0%	42.3%	41.6%	-	-	-	-

Sagittal tibial outliers +/-3°: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar	
Trials included/total available trials	2 of 7		3 of 22		1 of 8		1 of 8	
Patient included in Meta-Analysis (MA)	149		236		108		240	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	0.2 (0.0 to 2.6)	0.0 (0.0 to 0.4)	1.1 (0.1 to 6.6)	1.5 (0.4 to 6.0)	3.6 (0.5 to 27.0)	3.6 (1.7 to 9.1)	1.3 (0.2 to 9.2)	1.3 (0.6 to 2.7)
I ²	0.0%	0.0%	60.3%	55.0%	-	-	-	-

Surgical time: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Quadriceps sparring Vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar		Subvastus vs quadriceps sparring	
Trials included/total available trials	4 of 7		12 of 22		2 of 2		1 of 8		3 of 8		4 of 16		1 of 1	
Patient included in MA	270		948		100		90		99		454		60	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	-8.0 (-19.0 to 3.2)	-3.5 (-7.0 to 0.1)	-9.3 (-16.0 to -2.9)	-1.9 (-5.5 to 1.6)	6.7 (-11.0 to 25.0)	6.6 (-2.7 to 16.0)	-5.0 (-27.0 to 17.0)	-5.0 (-11.0 to 1.3)	12.0 (1.08 to 122.0)	9.7 (7.5 to 12.0)	11.0 (0.2 to 21.0)	17.0 (14.0 to 19.0)	-1.1 (-24.0 to 21.0)	-1.0 (-6.6 to 4.5)
I ²	93.0%	90.6%	94.1%	93.9%	54.7%	54.1%	-	-	91.3%	91.0%	94.0%	90.4%	-	-

Blood Loss: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Quadriceps sparring Vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar		Subvastus vs quadriceps sparring	
Trials included/total available trials	5 of 7		11 of 22		2 of 2		4 of 8		5 of 8		5 of 16		1 of 1	
Patient included in MA	365		898		200		456		510		396		60	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	15.0 (-150.0 to 180)	-4.5 (-6.5 to -2.6)	-15.0 (-99.0 to -130.0)	9.7 (-3.3 to 23.0)	-22.0 (-320 to 280)	-32.0 (-190 to 130.0)	230 (48.0 to 420)	310 (290.0 to 330.0)	84.0 (-80.0 to 250)	50 (25.0 to 74.0)	-120.0 (-290.0 to 46.0)	-71.0 (120.0 to -28.0)	-25.0 (-380.0 to 330.0)	-25.0 (-31.0 to -19.0)
I ²	99.0%	0.0%	92.6%	92.6%	0.0%	0.0%	98.4%	97.8%	84.4%	77.9%	78.5%	72.0%	-	-

Length of stay: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Quadriceps sparring Vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar	
Trials included/total available trials	1 of 7		7 of 22		2 of 2		4 of 8		3 of 8		4 of 16	
Patient included in MA	80		479		100		300		139		412	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	-0.3 (-1.6 to 1.1)	-0.3 (-0.8 to -0.3)	0.2 (-0.4 to 0.8)	0.1 (-0.1 to 0.4)	0.0 (-1.2 to 1.2)	0.0 (-0.6 to 0.5)	5.0 (3.7 to 6.3)	5.0 (4.6 to 5.4)	-0.5 (-1.4 to 0.3)	-0.5 (-0.9 to -0.1)	-0.4 (-1.1 to 0.3)	-0.4 (-0.7 to -0.1)
I ²	-	-	0.0%	0.0%	0.0%	0.0%	-	97.8%	83.9%	83.8%	85.8%	85.6%

VAS at day one post-surgery: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Quadriceps sparring Vs midvastus		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar		Subvastus vs quadriceps sparring	
Trials included/total available trials	2 of 7		1 of 22		1 of 2		2 of 8		4 of 16		1 of 1	
Patient included in MA	149		39		40		100		322		60	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	-0.0 (-1.6 to 1.5)	-0.0 (-0.3 to -0.3)	-0.2(-3.1 to 2.7)	-0.2 (-2.2 to 1.8)	-0.5 (-3.3 to 2.4)	-0.5 (-2.4 to 1.4)	-0.1 (-1.9 to 1.7)	-0.0 (-0.6 to 0.6)	-1.5 (-2.6 to -0.3)	-0.9(-1.0 to -0.8)	-0.2(-2.5 to 2.1)	-0.2(-0.8 to 0.4)
I ²	0.0%	0.0%	-	-	-	-	0.0%	0.0%	98.2%	92.9%	-	-

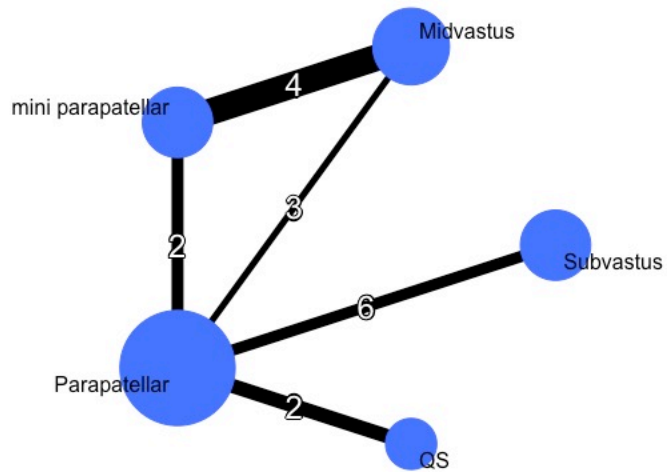
Complications: Pairwise

Treatment comparisons	Mini-parapatellar vs Midvastus		Parapatellar vs midvastus		Quadriceps sparring Vs midvastus		Parapatellar vs Mini-parapatellar		Quadriceps sparring vs parapatellar		Subvastus vs parapatellar		Subvastus vs quadriceps sparring	
Trials included/total available trials	5 of 7		18 of 22		2 of 2		6 of 8		6 of 8		13 of 16		1 of 1	
Patient included in MA	369		1469		60		654		551		1247		60	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise MA RR (95% CI)	0.0 (0.0 to 0.1)	0.0 (0.0 to 0.0)	0.7 (0.5 to 0.9)	0.7 (0.5 to 0.9)	1.5 (0.3 to 11.0)	1.6 (0.3 to 9.8)	0.8 (0.3 to 2.2)	0.9 (0.3 to 2.3)	2.2 (1.0 to 5.2)	2.1 (1.1 to 4.3)	0.8 (0.5 to 1.2)	0.8 (0.6 to 1.0)	1.3 (0.3 to 6.4)	1.3 (0.4 to 5.0)
I ²	0.0%	99.5%	13.4%	12.9%	0.0%	0.0%	37.9%	35.5%	0.0%	0.0%	12.2%	19.6%	-	-

Appendix 7. Secondary outcomes: clinical outcomes

A. Network diagrams

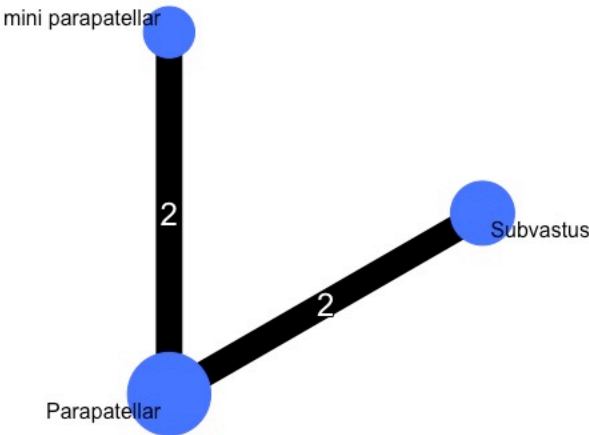
KSS within One and two years post surgery network



WOMAC within one and two years post surgery network



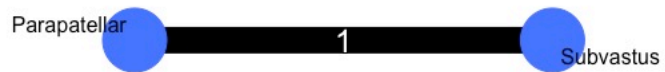
Delta SF 36 physical at the first six months network



Delta SF 36 physical within 1 and 2 years post operative



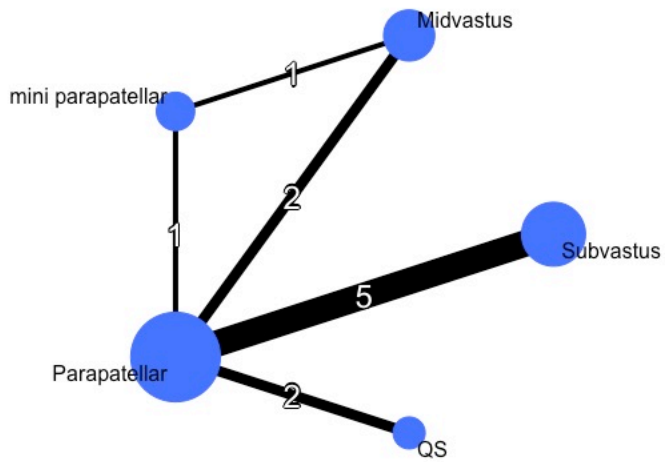
Delta SF 36 health at the first six months network



Delta SF 36 health within 1 and 2 years post operative

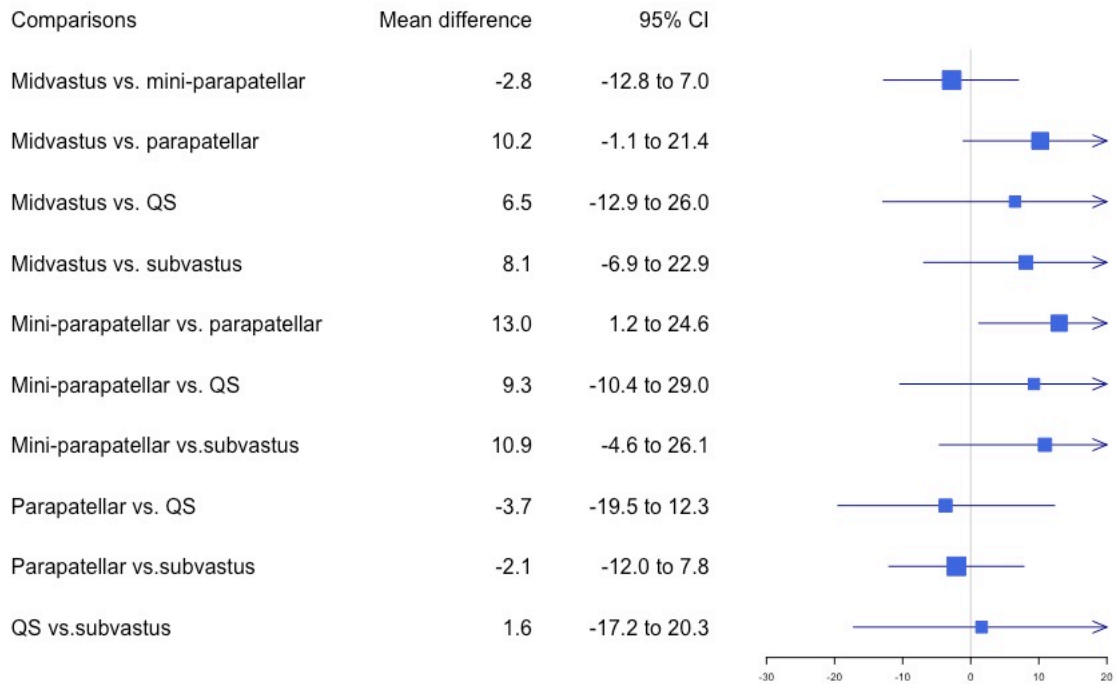


Delta ROM within 1 and 2 years post operative

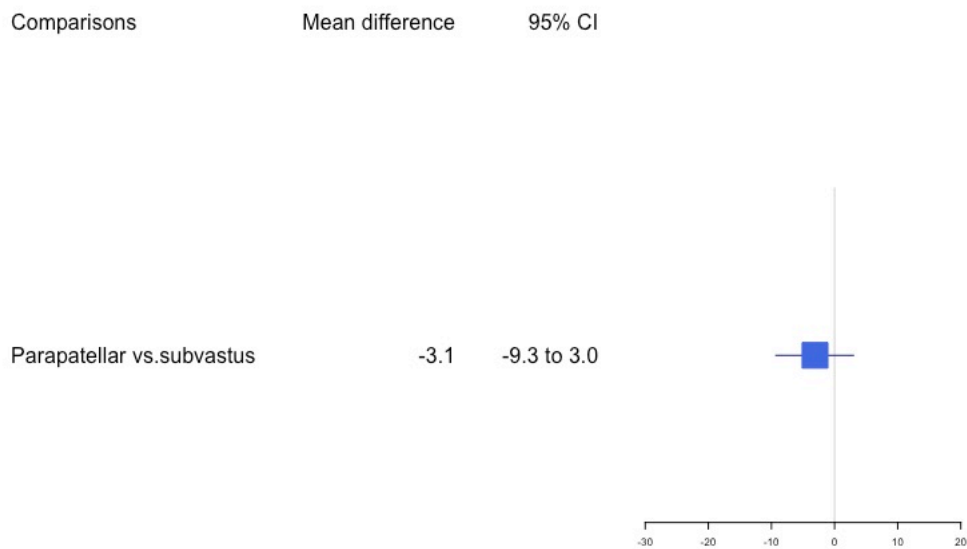


B. Relative effects

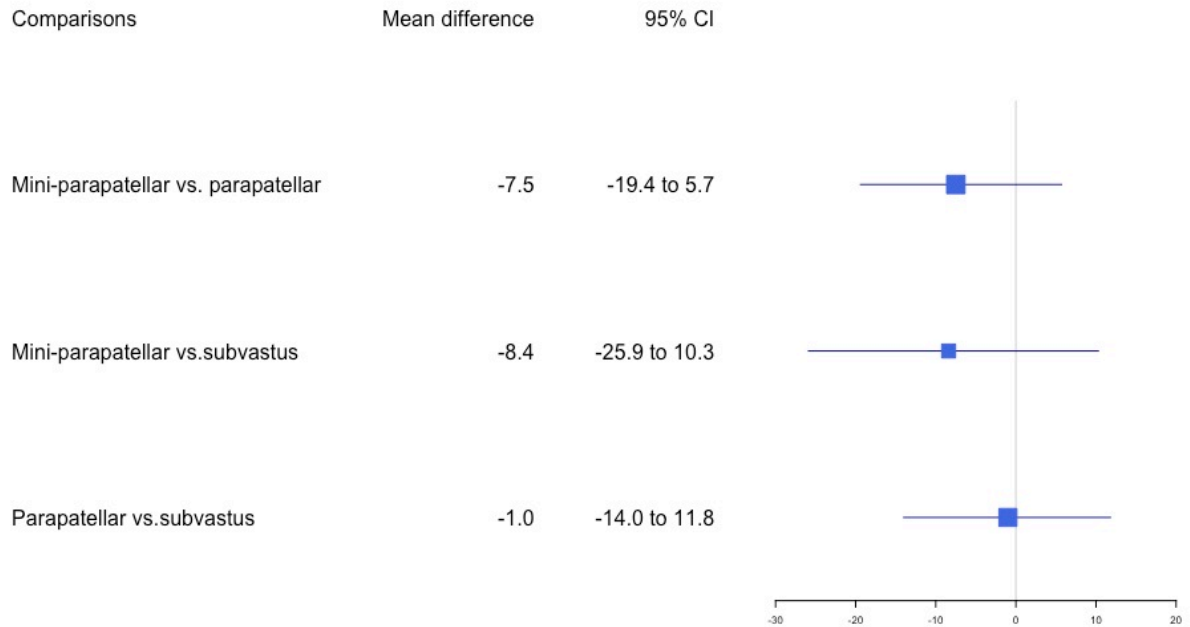
Forest plots in the network meta-analysis of the mean difference of KSS within 1 and 2 years postoperatively. QS= quadriceps sparing.



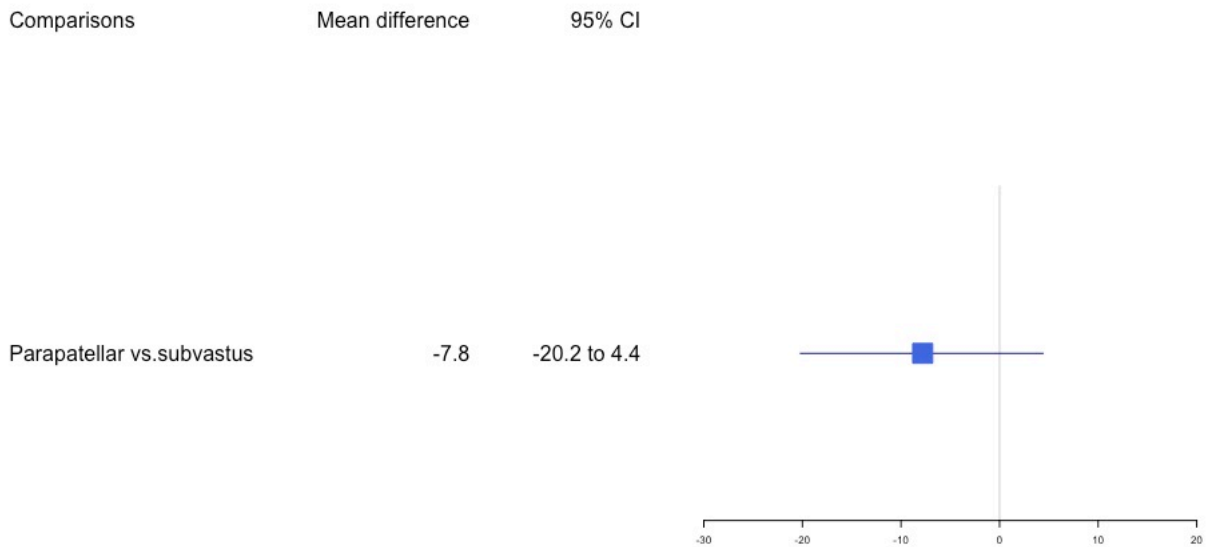
Forest plots in the network meta-analysis of the mean difference of WOMAC within 1 and 2 years postoperatively. QS= quadriceps sparing.



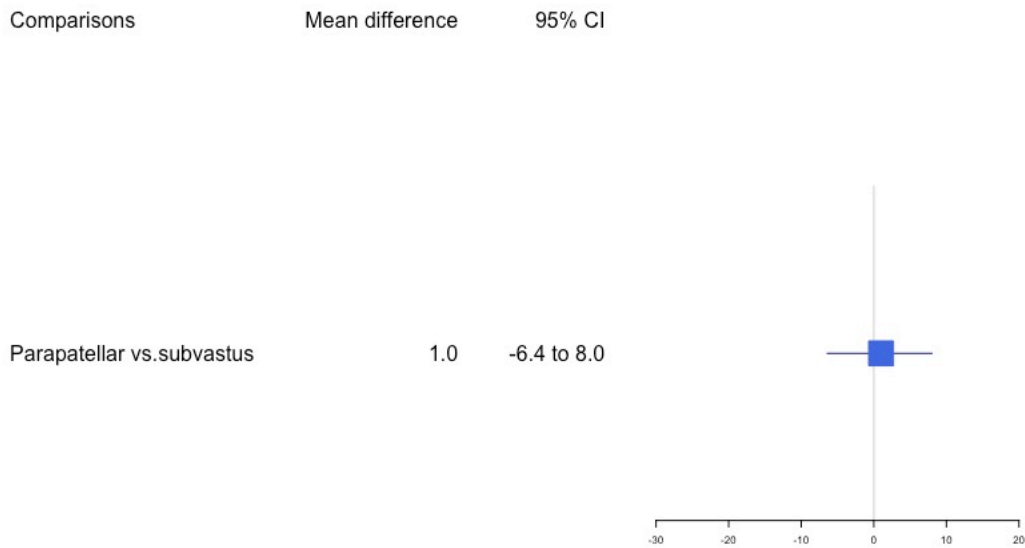
Forest plots in the network meta-analysis of the mean difference of SF 36 physical scale at 6 months postoperatively. QS= quadriceps sparring.



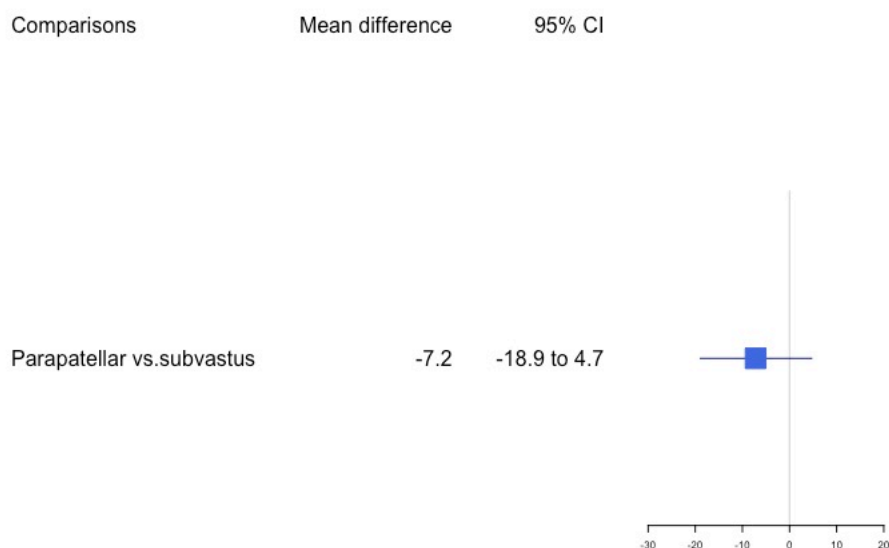
Forest plots in the network meta-analysis of the mean difference of SF 36 physical scale within 1 and 2 years postoperatively. QS= quadriceps sparring.



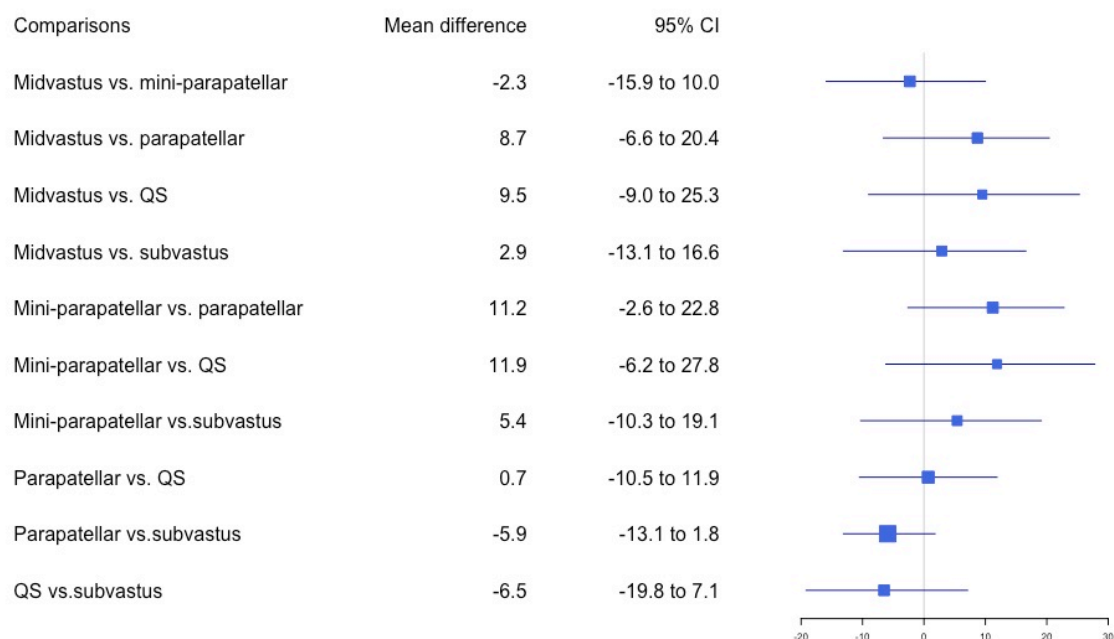
Forest plots in the network meta-analysis of the mean difference of SF 36 health perception scale at 6 months postoperatively. QS= quadriceps sparring.



Forest plots in the network meta-analysis of the mean difference of SF 36 health perception scale within 1 and 2 years postoperatively. QS= quadriceps sparring.



Forest plots in the network meta-analysis of the mean difference of ROM within 1 and 2 years post-surgery. QS= quadriceps sparring.

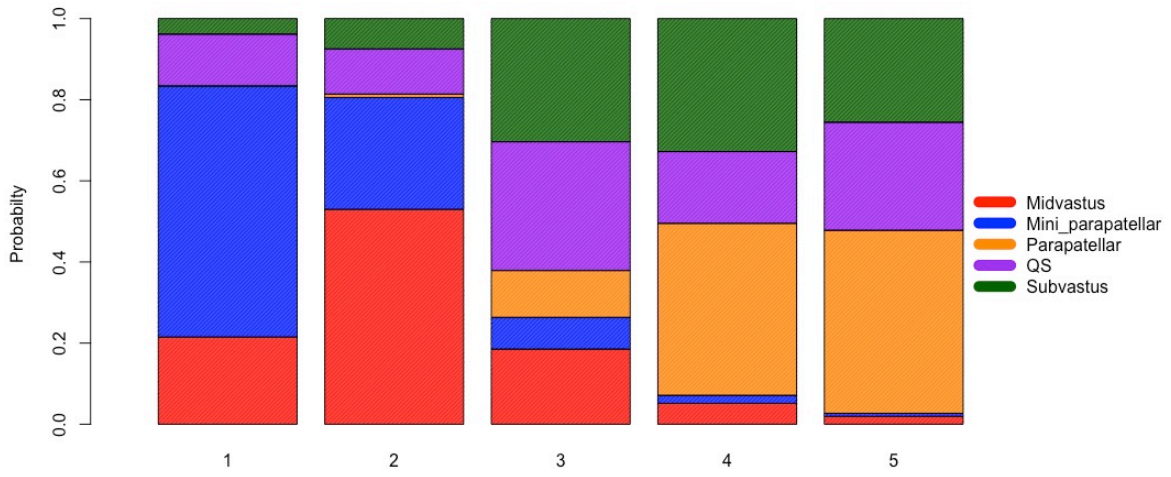


Ranking

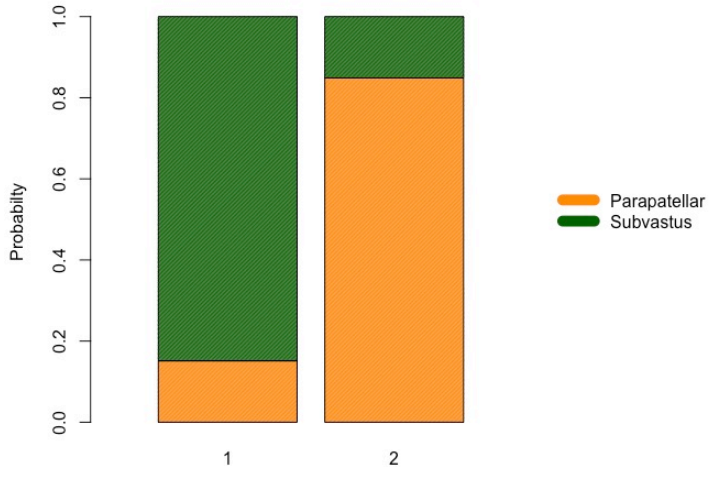
Mean rank

These images show rank probabilities of secondary clinical outcomes for each cutting guide. The x axis represents the ranking. The y axis represents the probability of rank: 0% to 100%. Each color represents one surgical approach: red for midvastus, blue for mini-parapatellar, orange for parapatellar, purple for quadriceps sparring, and green for the subvastus.

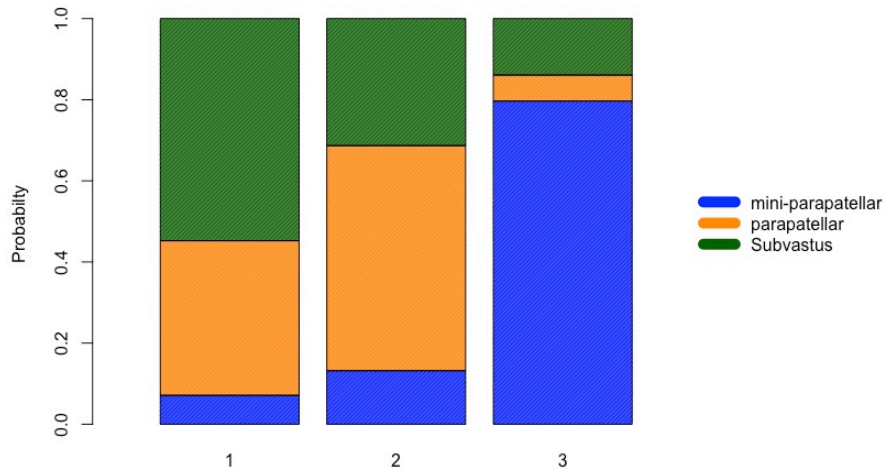
Ranking for KSS within one and two years months PO



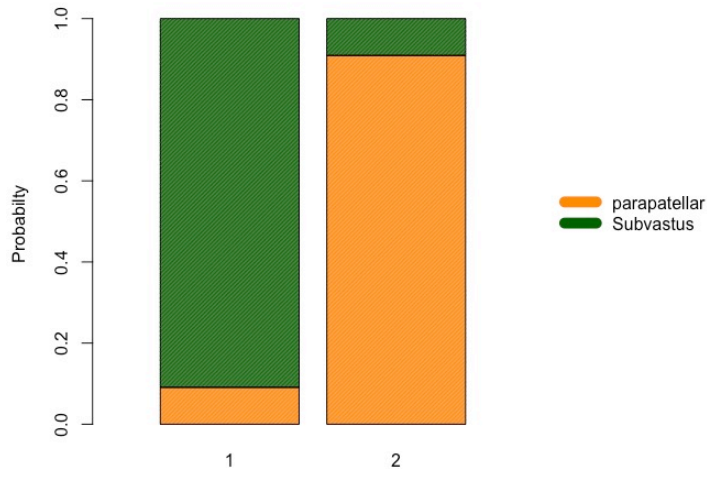
Ranking for WOMAC within one and two years months PO



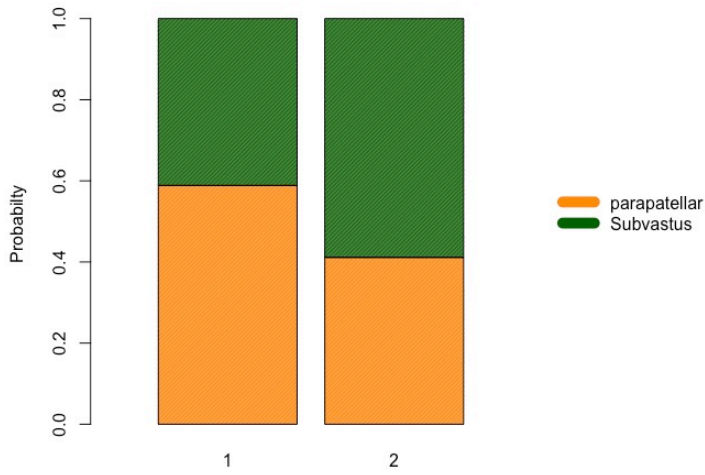
Ranking for SF 36 physical functioning at the first 6 months post-surgery



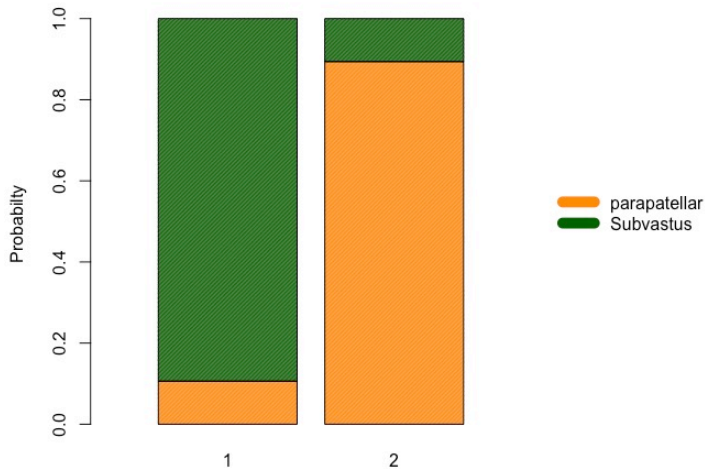
Ranking for SF 36 physical functioning within one and two years months PO



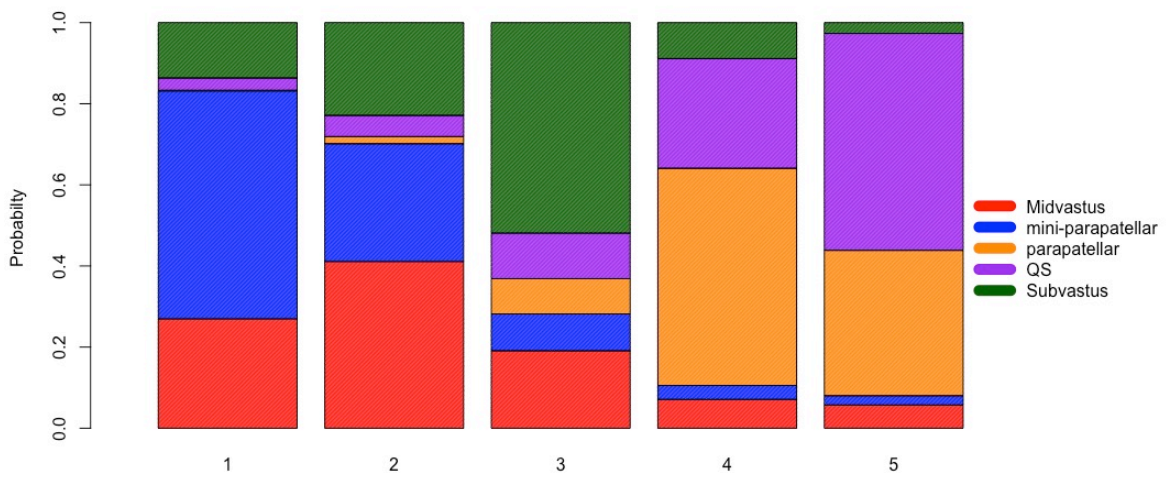
Ranking for SF 36 Health perception at the first 6 months post-surgery



Ranking for SF 36 Health perception within one and two years months PO

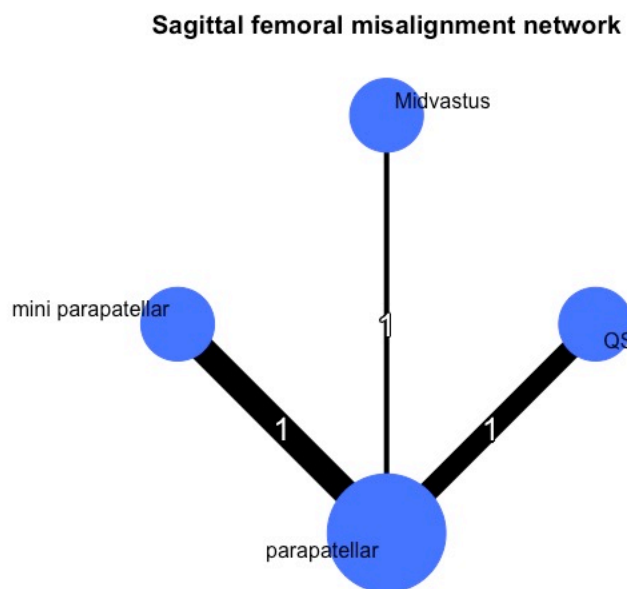
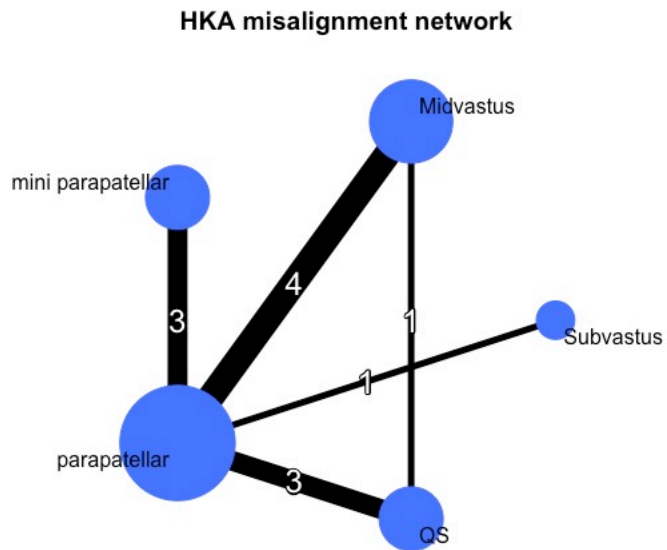


Ranking for Range of Motion within one and two years months PO

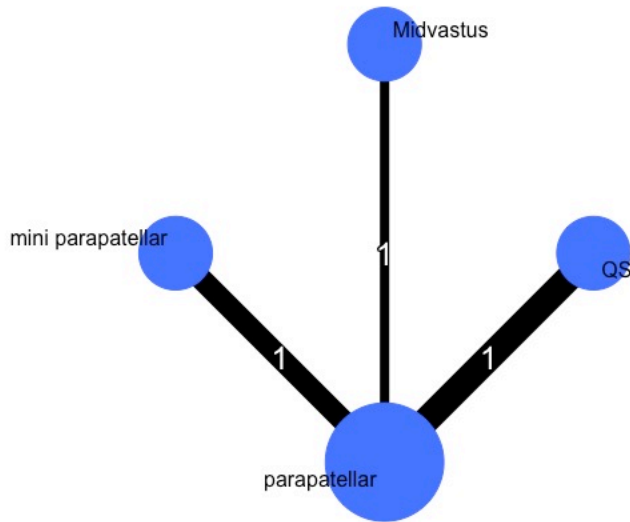


Appendix 8. Secondary outcomes: radiological outcomes

A. Network diagrams

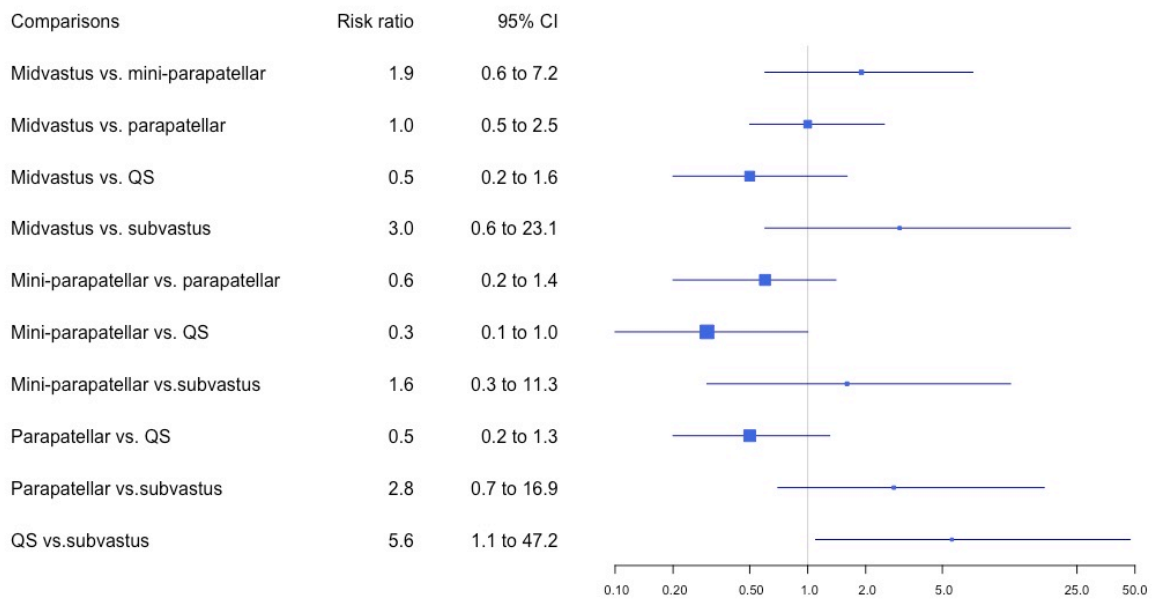


Sagittal tibial misalignment network

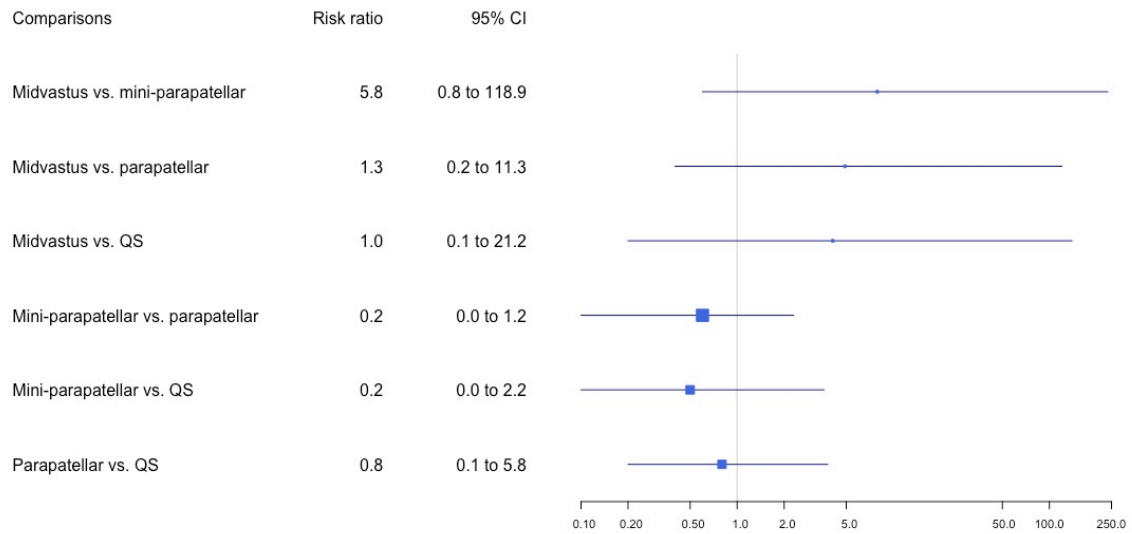


B. Relative effects

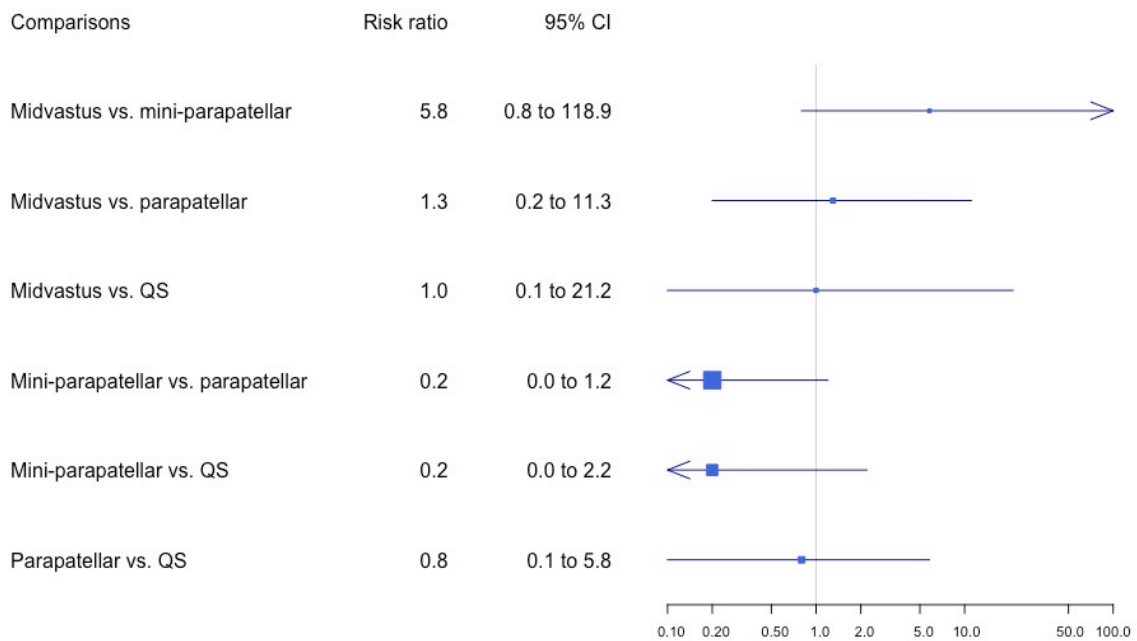
Forest plots in the network meta-analysis of the relative risks of HKA outliers between the different surgical approaches. QS= quadriceps sparing.



Forest plots in the network meta-analysis of the relative risks of sagittal femoral outliers between the different surgical approaches. QS= quadriceps sparing.



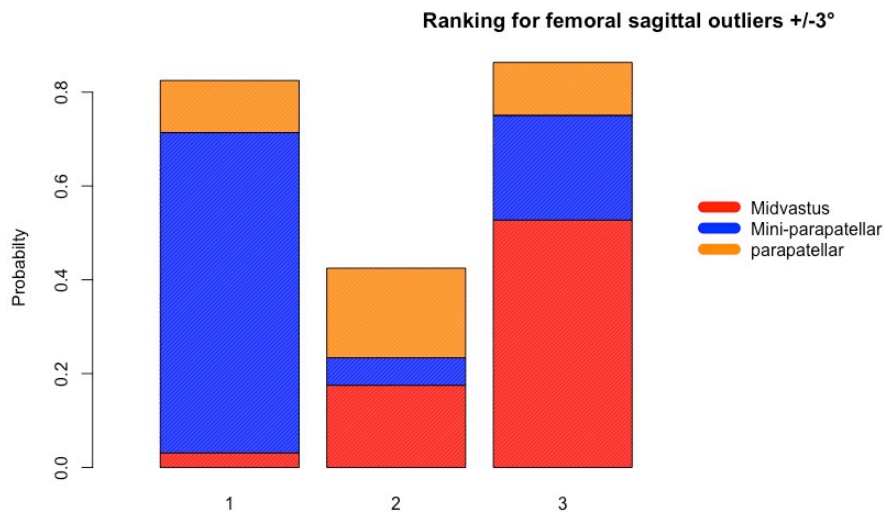
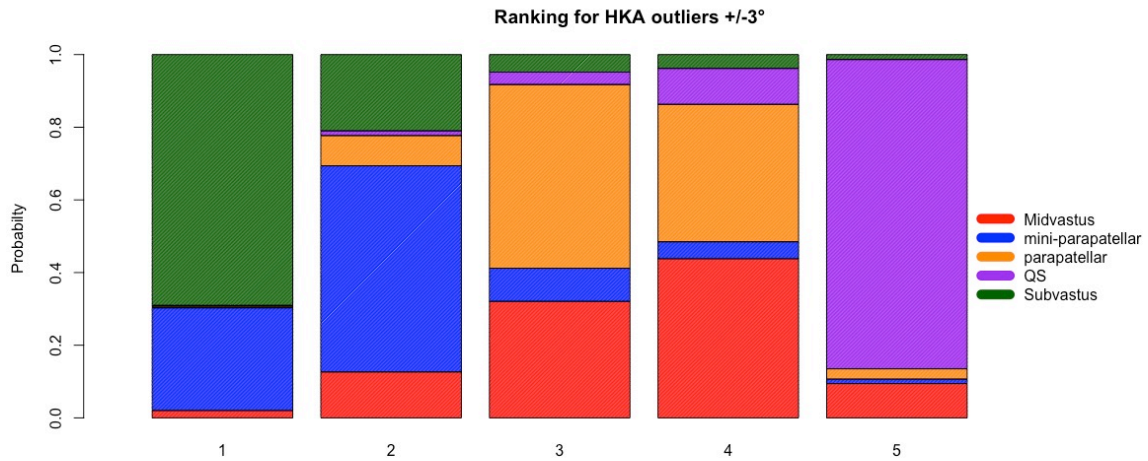
Forest plots in the network meta-analysis of the relative risks of sagittal tibial outliers between the different surgical approaches. QS= quadriceps sparing.



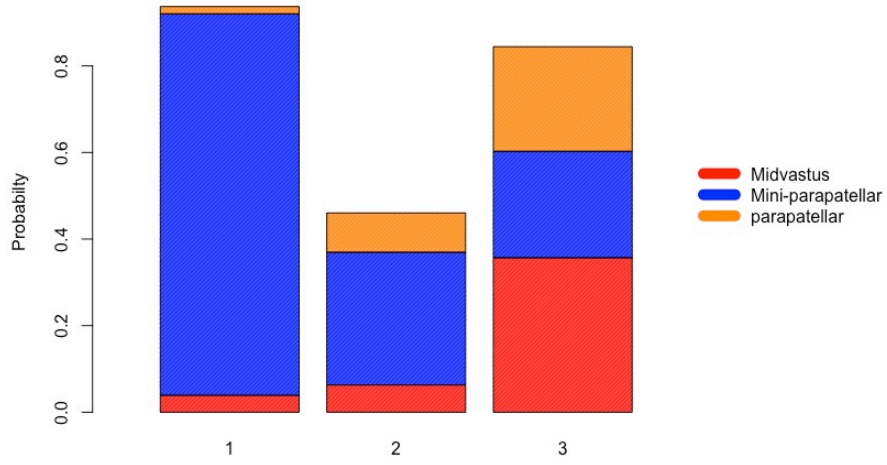
Ranking

Mean rank

These images show rank probabilities of secondary radiological outcomes for each cutting guide. The x axis represents the ranking. The y axis represents the probability of rank: 0% to 100%. Each color represents one surgical approach: red for midvastus, blue for mini-parapatellar, orange for parapatellar, purple for quadriceps sparing, and green for the subvastus.

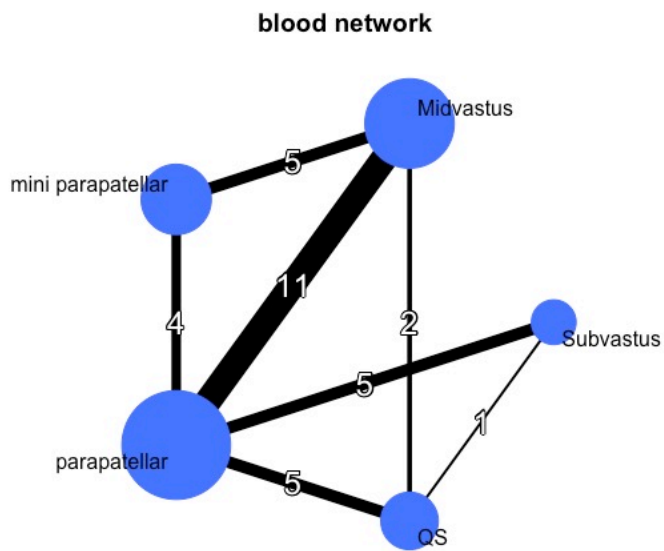
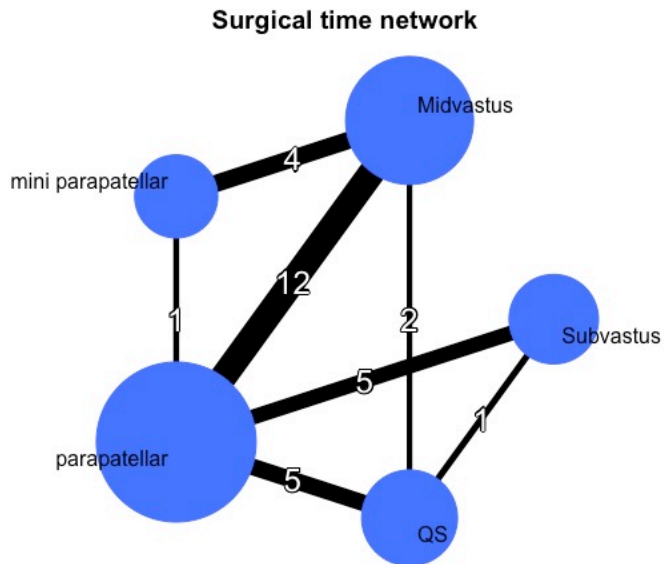


Ranking for tibial sagittal outliers +/-3°

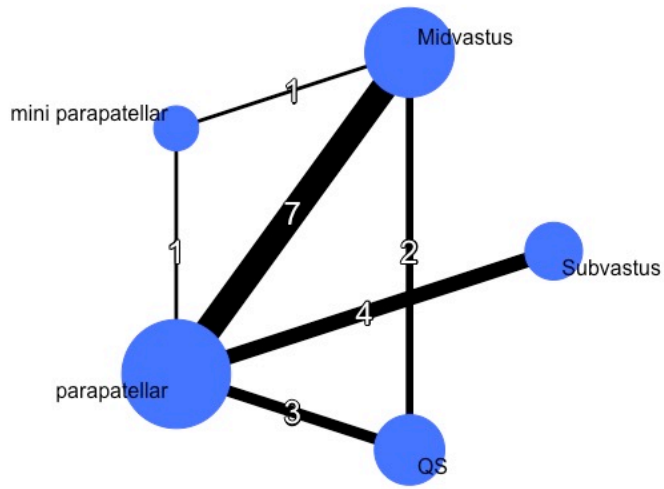


Appendix 9. Secondary outcomes: per operative outcomes

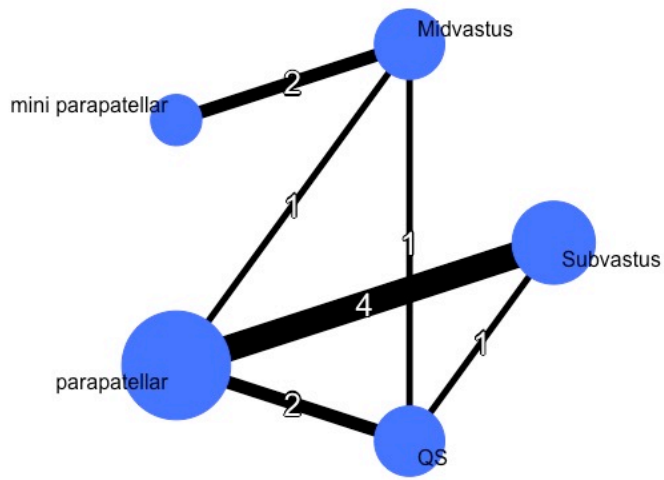
A. Network diagrams

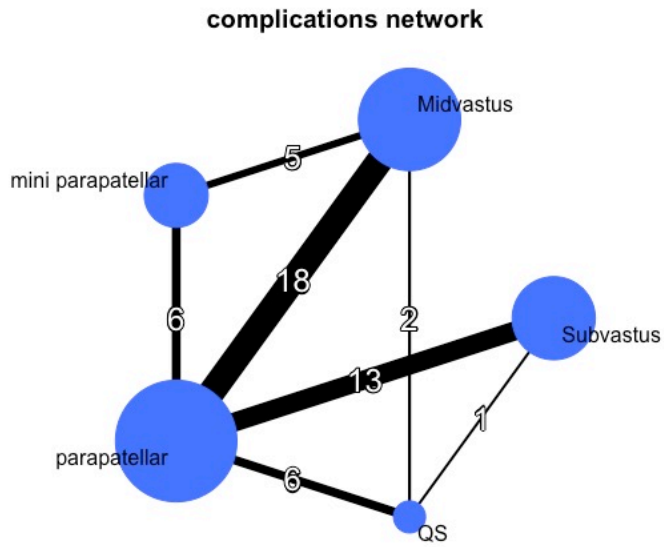


length network



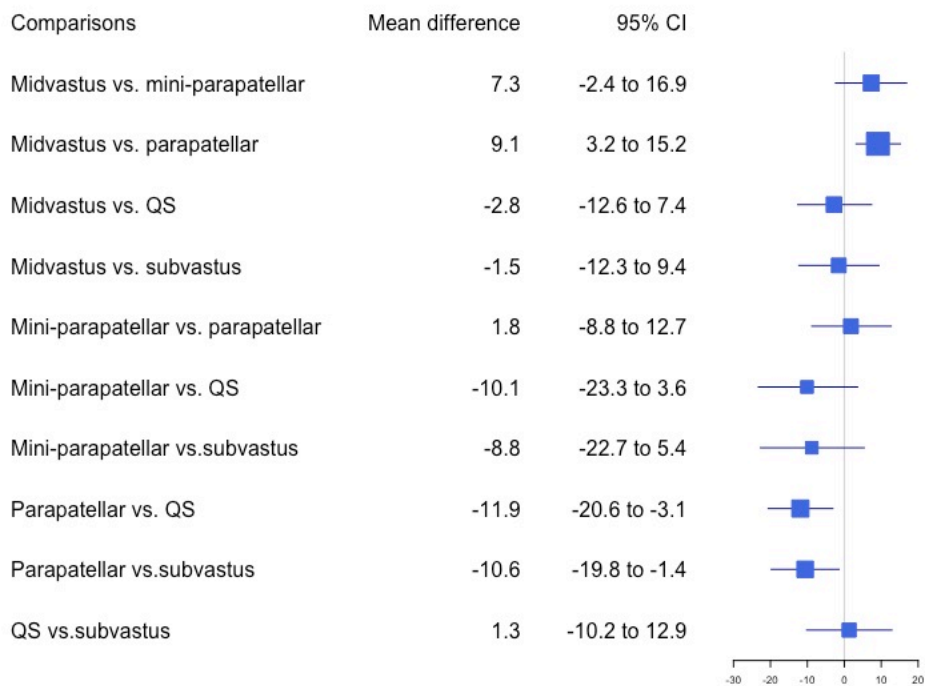
VAS at Day 1 PO network



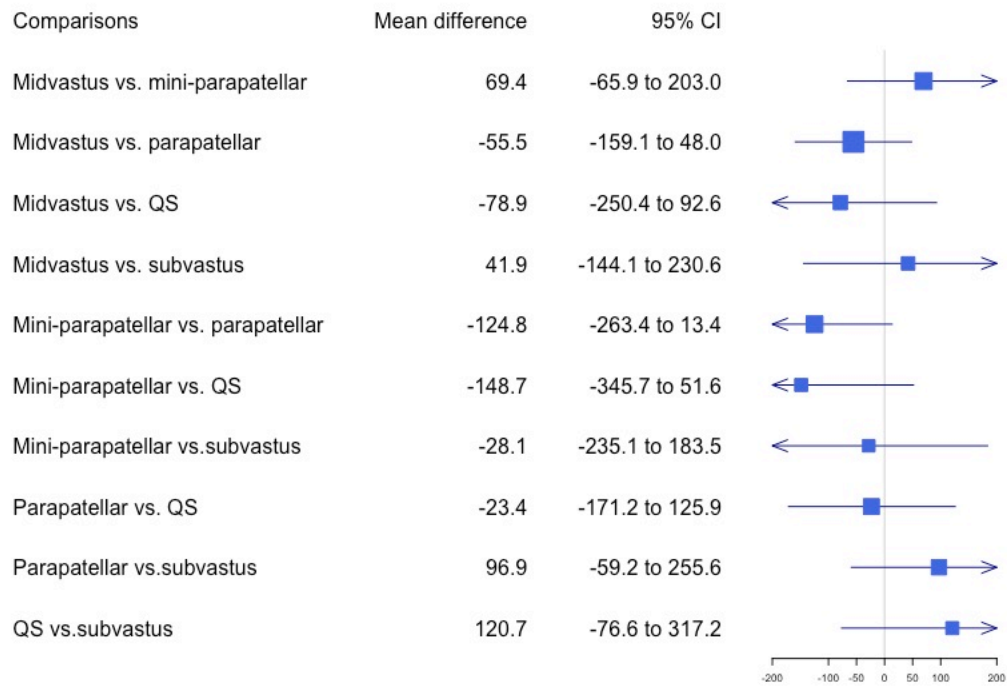


B. Relative effects

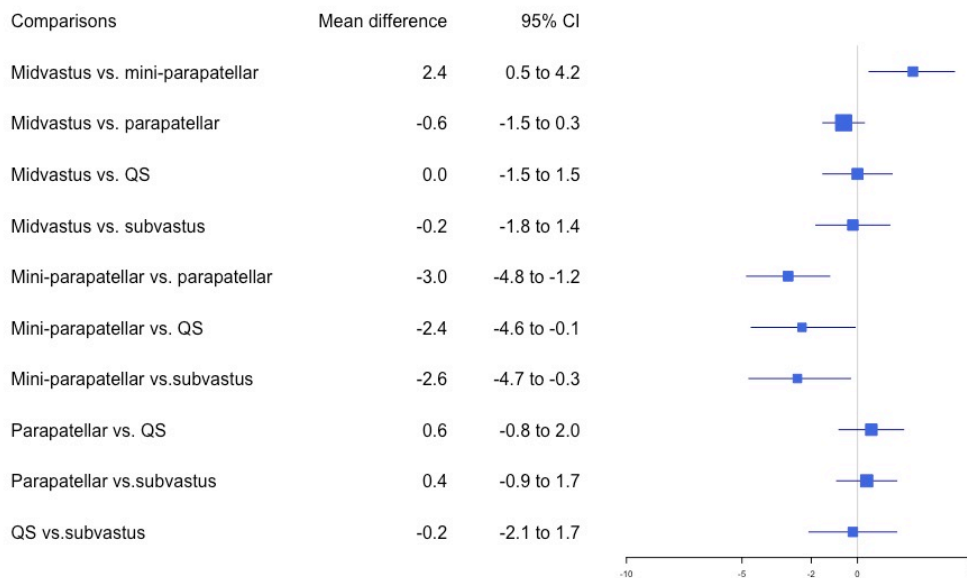
Forest plots in the network meta-analysis of the mean difference of surgical time. QS= quadriceps sparring.



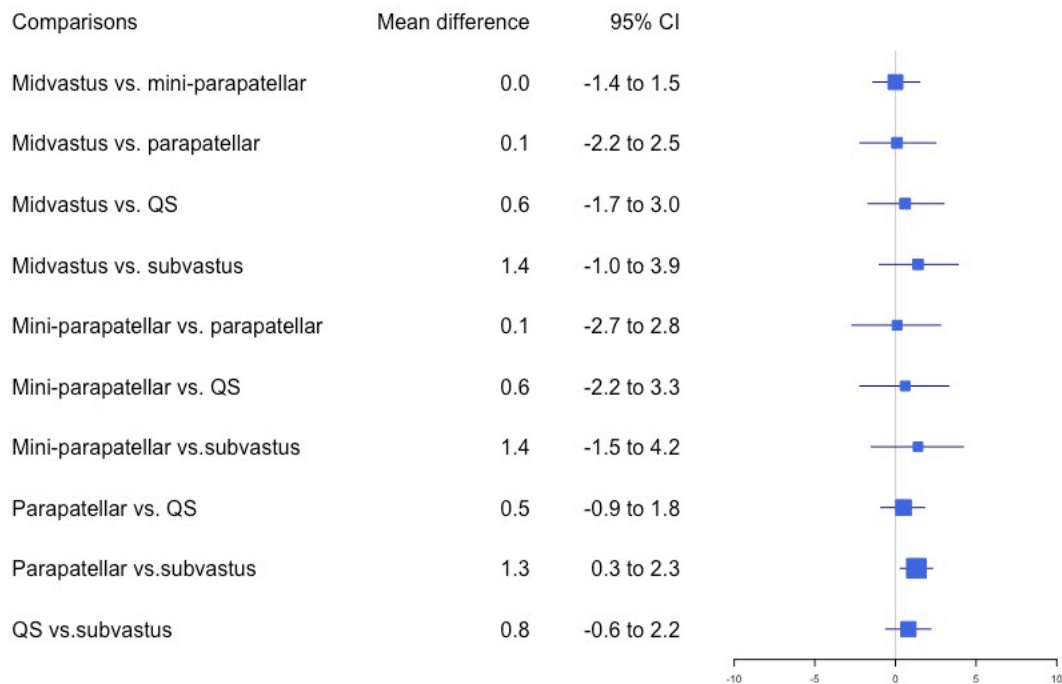
Forest plots in the network meta-analysis of the mean difference of blood loss. QS= quadriceps sparring.



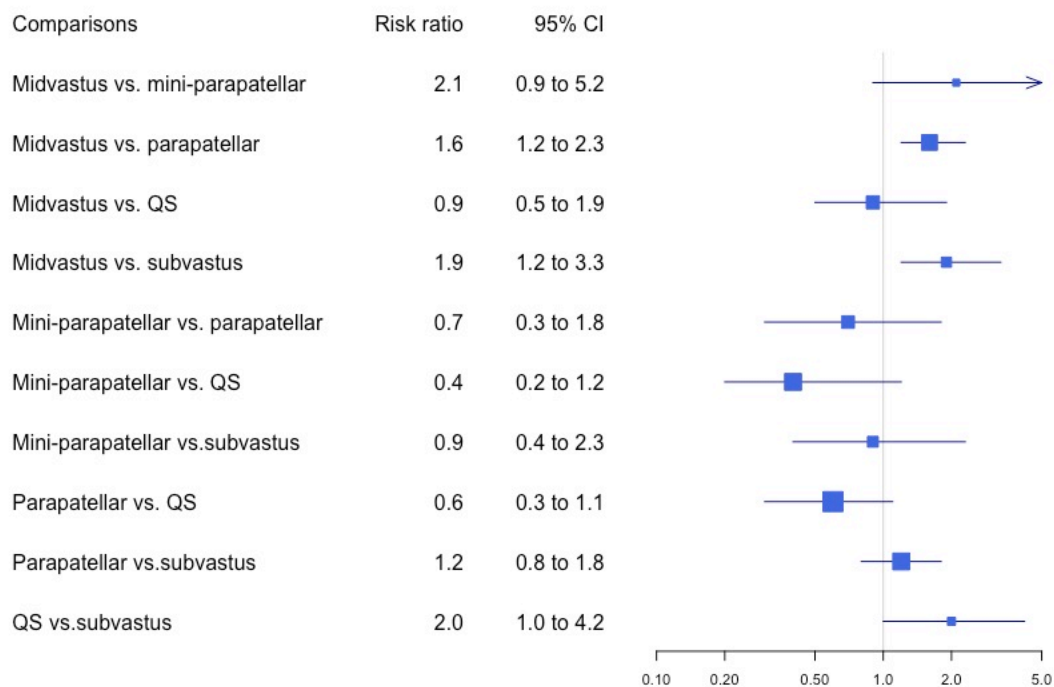
Forest plots in the network meta-analysis of the mean difference of length of stay. QS= quadriceps sparring.



Forest plots in the network meta-analysis of the mean difference of VAS at day one post-surgical. QS= quadriceps sparring.



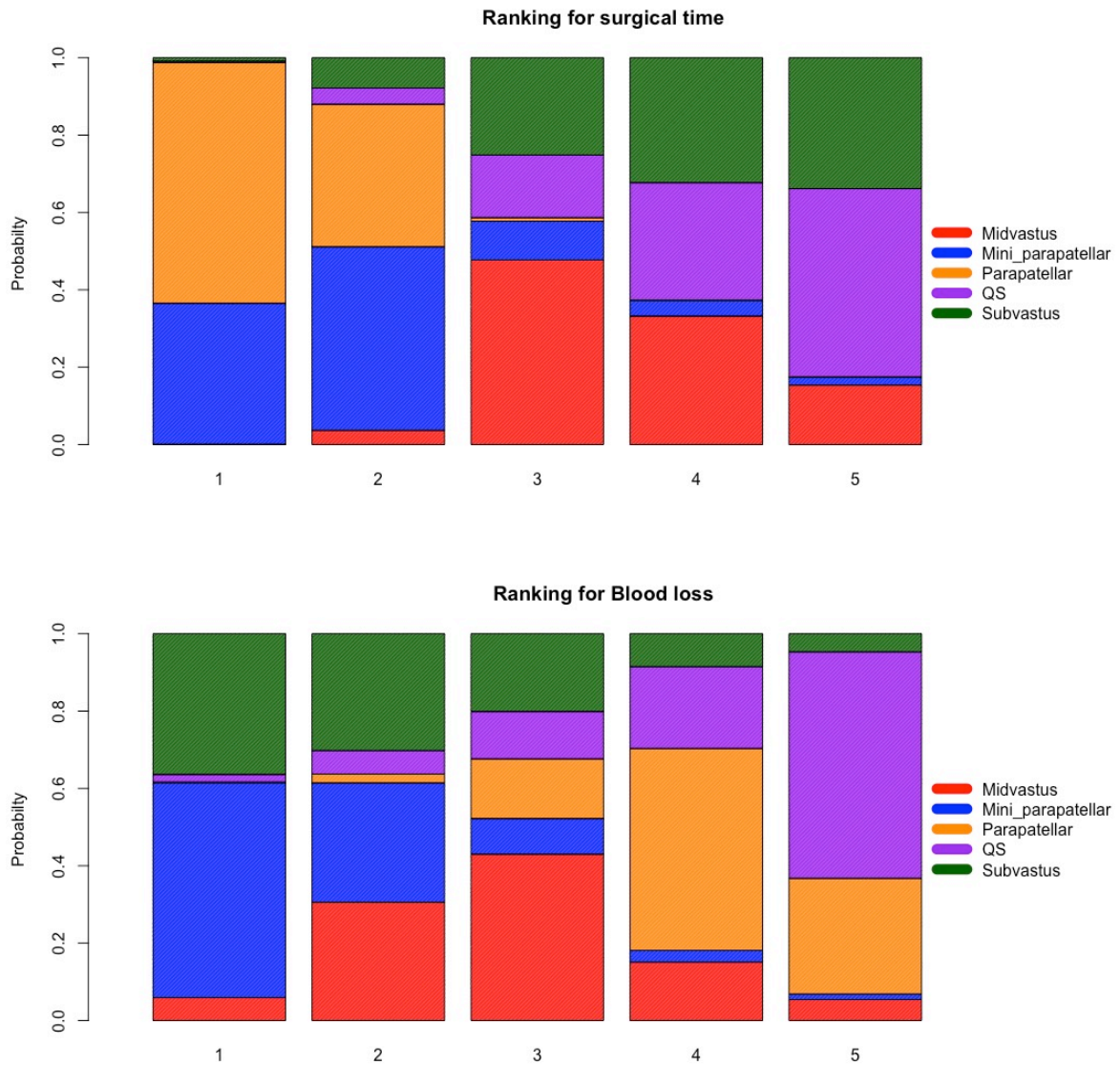
Forest plots in the network meta-analysis of the relative risks of complications between the different surgical approaches. QS= quadriceps sparing.



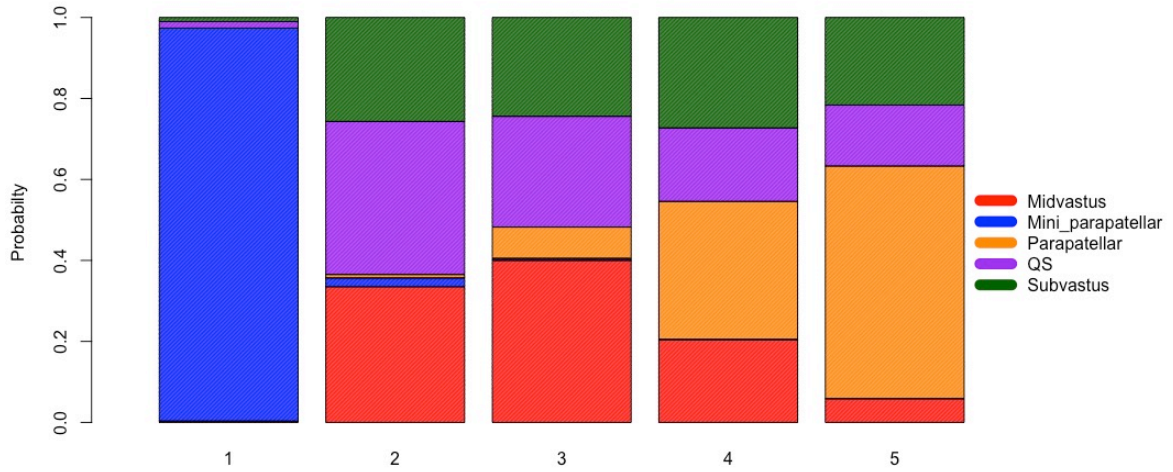
Ranking

Mean rank

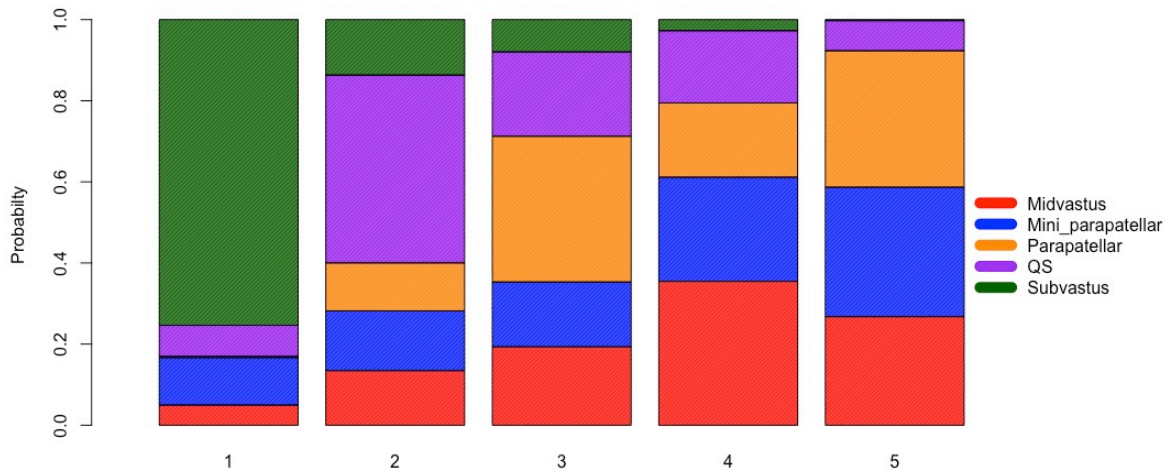
These images show rank probabilities of secondary radiological outcomes for each cutting guide. The x axis represents the ranking. The y axis represents the probability of rank: 0% to 100%. Each color represents one surgical approach: red for midvastus, blue for mini-parapatellar, orange for parapatellar, purple for quadriceps sparing, and green for the subvastus.



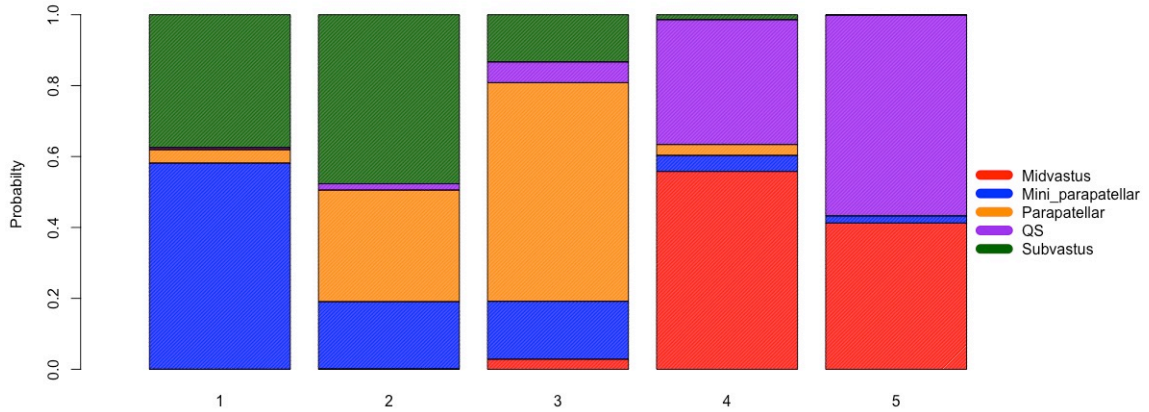
Ranking for length of stay



Ranking for VAS at day one PO



Ranking for complications



Appendix 10. Complications

Types of complications	Midvastus	Mini-parapatellar	Parapatellar	Quadriceps sparring	Subvastus
Total	7.7% (75/973)	0.6% (3/523)	7.6% (146/1909)	9.0% (29/321)	10.4% (68/653)
Infections	0.7% (5/763)	0.0% (0/253)	1.8% (21/1179)	3.5% (6/170)	2.6% (10/390)
Others	2.4% (13/543)	0.0% (0/179)	2.9% (27/926)	1.7% (2/120)	3.8% (13/342)
Mechanical	0.9% (6/667)	0.0% (0/149)	1.4% (14/966)	5.3% (8/150)	1.4% (5/350)
Skin	4.4% (27/620)	0.0% (0/179)	1.7% (20/1204)	3.4% (6/176)	3.2% (13/406)
Stiffness	1.3% (8/618)	0.0% (0/149)	1.9% (21/1091)	0.7% (1/150)	1.7% (7/406)
Vascular	3.0% (5/393)	1.8% (3/171)	3.7% (36/969)	1.9% (4/210)	5.7% (20/350)
Fractures	0.5% (2/422)	0.0% (0/171)	0.7% (6/813)	1.7% (2/120)	0.0% (0/306)

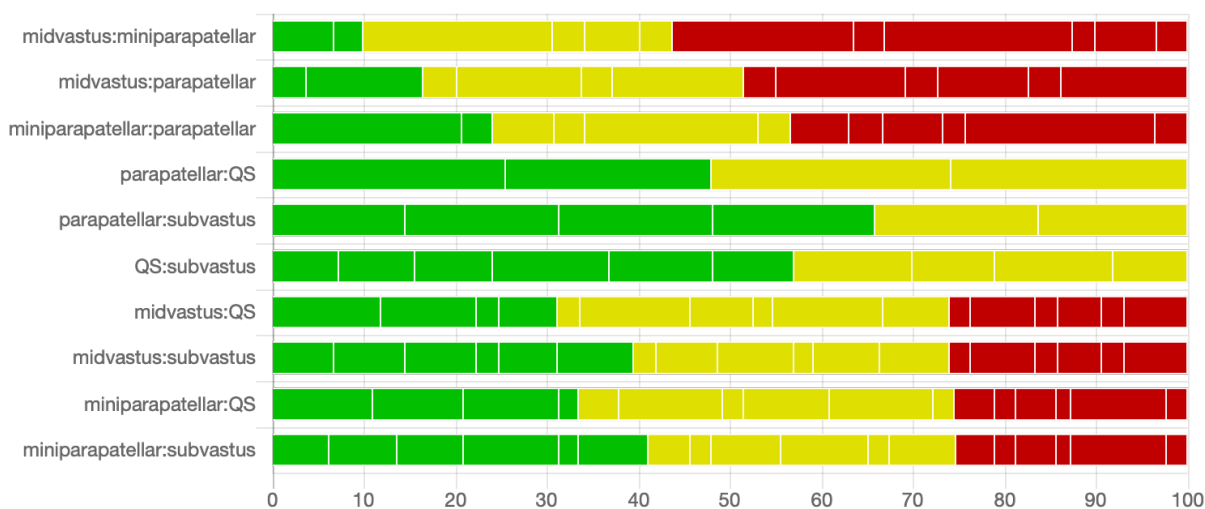
Appendix 11. Grade evaluation

The quality of evidence of the network meta-analysis has been assessed using the GRADE framework [1]. The online tool CiNeMA [2] was used to evaluate study limitations. It combines judgments about direct comparison with their contribution in the network meta-analysis results. Below, our GRADE evaluation for KSS at 6 months postoperative, ROM at 6 months postoperative and WOMAC at 6 months postoperative are presented. *QS* = *Quadriceps sparring*.

Grade evaluation for KSS within first 6 months postoperative

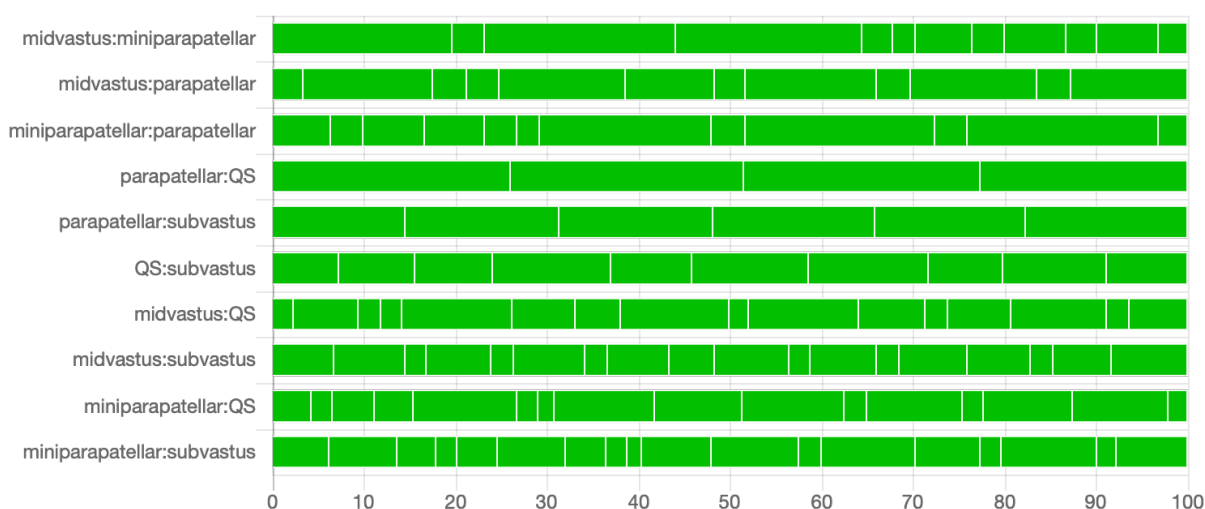
Study limitations

QS:quadriceps sparring.



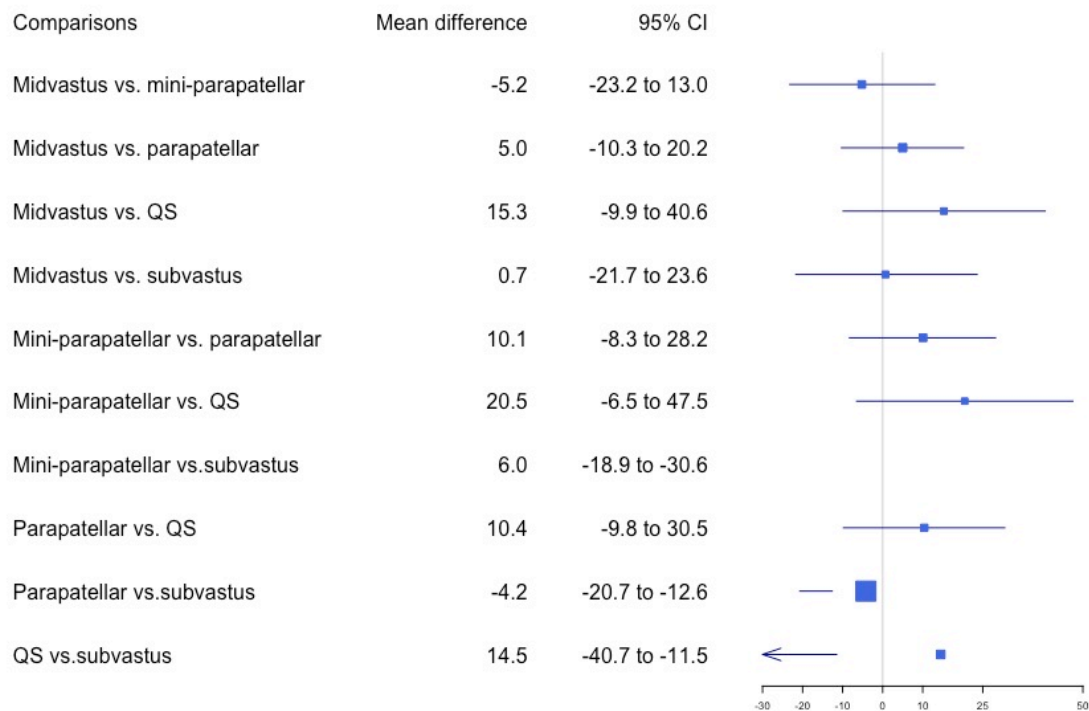
The bar-graph shows for the network estimate of every pairwise comparison how much information comes from low (green), moderate (yellow) or high (red) risk of bias trials.

Indirectness



Imprecision

We have some concerns for parapatellar/QS, midvastus/QS, mini-parapatellar/QS, QS/subvastus and mini-parapatellar/subvastus because the credible intervals was large.



Heterogeneity/inconsistency

We have major concerns for the heterogeneity of parapatellar/mini-parapatellar comparison ($I^2 = 99.1\%$), QS/parapatellar ($I^2 = 99.1\%$) and subvastus/parapatellar ($I^2 = 83.8\%$).

We did not find inconsistency for this network.

Publication bias

We searched in ClinicalTrials.gov and in The World Health Organization International Clinical Trials Registry Platform (ICTRP) to locate potential unpublished (or gray) literature. So, we do not have serious concerns about the presence of publications bias.

Final assessments

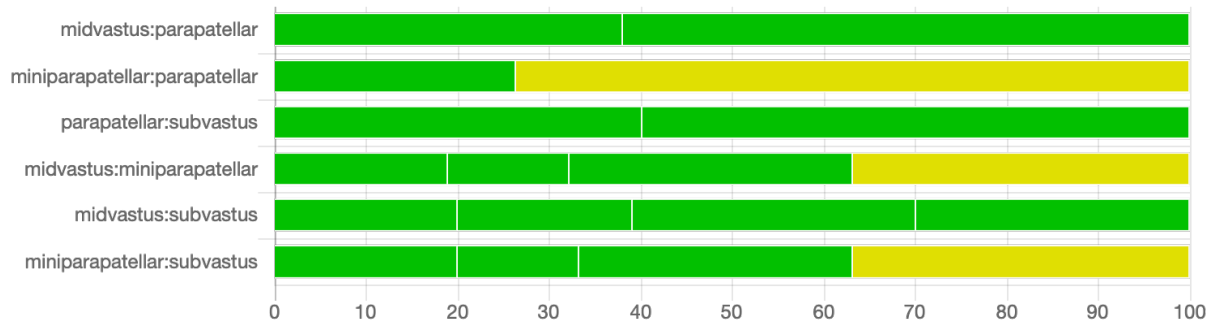
Comparison	Number of studies	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
parapatellar:QS	4	Some concerns	Undetected	No concerns	Some concerns	Major concerns	No concerns	Moderate
midvastus:mini-parapatellar	3	Some concerns	Undetected	No concerns	No concerns	No concerns	No concerns	Low
midvastus:parapatellar	6	Some concerns	Undetected	No concerns	No concerns	No concerns	No concerns	Low
Mini-parapatellar:parapatellar	3	Some concerns	Undetected	No concerns	No concerns	Major concerns	No concerns	Low
parapatellar:subvastus	6	No concerns	Undetected	No concerns	No concerns	Major concerns	No concerns	Low

- Salanti G, Giovane CD, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. *PLoS One*. 2014;9:e99682.
- Confidence in Network Meta-Analysis [Software]. Institute of Social and Preventive Medicine, University of Bern, 2017. Available from cinema.ispm.unibe.ch. Accessed June 7, 2018.

Grade evaluation for WOMAC at 6 months postoperative

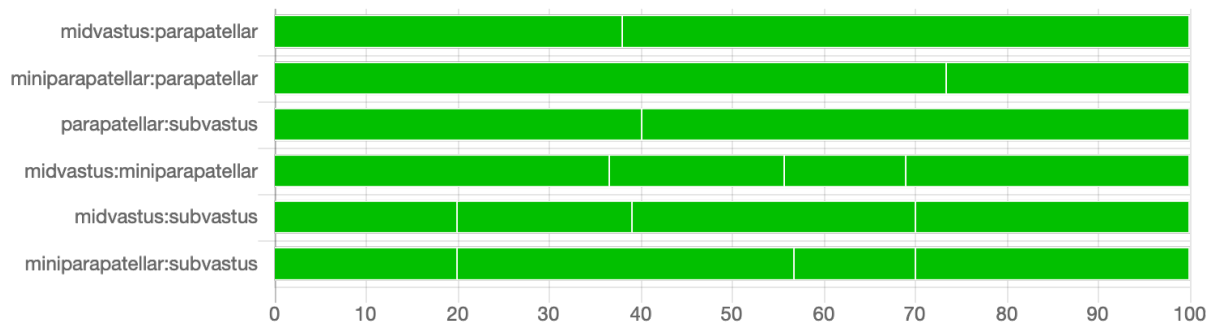
Study limitations

QS: quadriceps sparing.



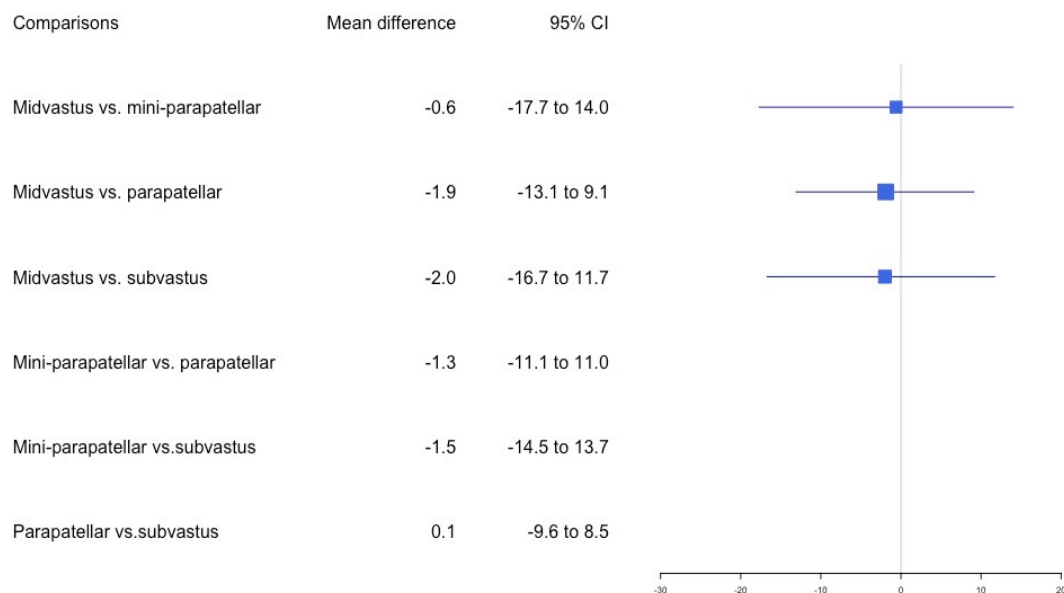
The bar-graph shows for the network estimate of every pairwise comparison how much information comes from low (green), moderate (yellow) or high (red) risk of bias trials.

Indirectness



Imprecision

We have no concerns for any comparisons because the credible intervals were small.



Heterogeneity/inconsistency

We have some major concerns for the heterogeneity of papatellar/mini-parapatellar comparison ($I^2 = 78.7\%$) and subvastus/parapatellar ($I^2 = 77.7\%$).

We did not find inconsistency for this network.

Publication bias

We searched in ClinicalTrials.gov and in The World Health Organization International Clinical Trials Registry Platform (ICTRP) to locate potential unpublished (or gray) literature. So, we don't have serious concerns about the presence of publications bias.

Final assessments

Comparison	Number of studies	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
midvastus:parapatellar	2	No concerns	Undetected	No concerns	Some concerns	No concerns	No concerns	Moderate
miniparapatellar:parapatellar	2	Some concerns	Undetected	No concerns	No concerns	Major concerns	No concerns	Low
parapatellar:subvastus	2	No concerns	Undetected	No concerns	No concerns	Major concerns	No concerns	Low

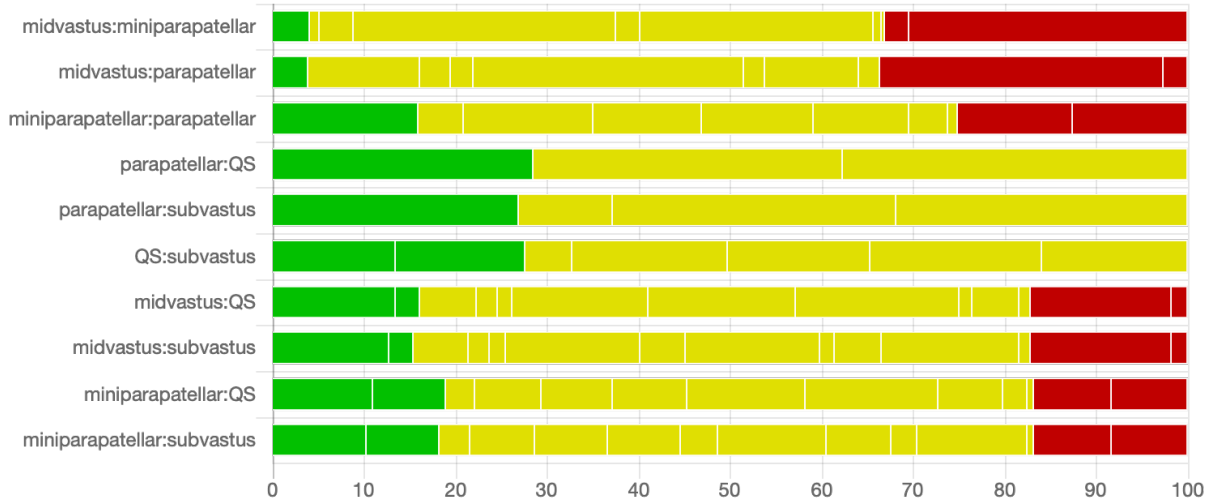
1. Salanti G, Giovane CD, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. *PLoS One*. 2014;9:e99682.

2. Confidence in Network Meta-Analysis [Software]. Institute of Social and Preventive Medicine, University of Bern, 2017. Available from cinema.ispm.unibe.ch. Accessed June 7, 2018.

Grade evaluation for ROM at the first 6 months post-surgery

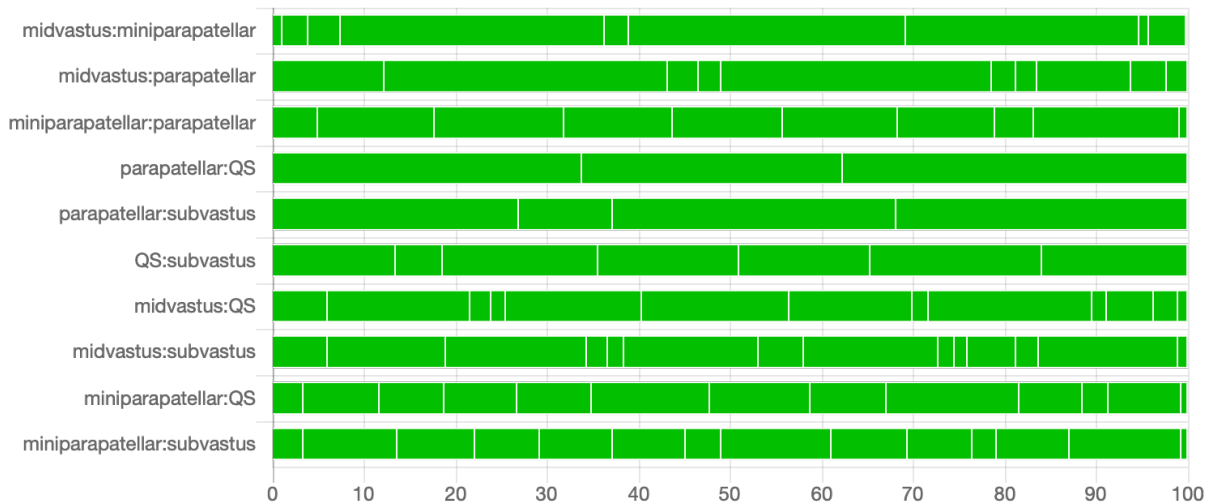
Study limitations

QS: quadriceps sparing.



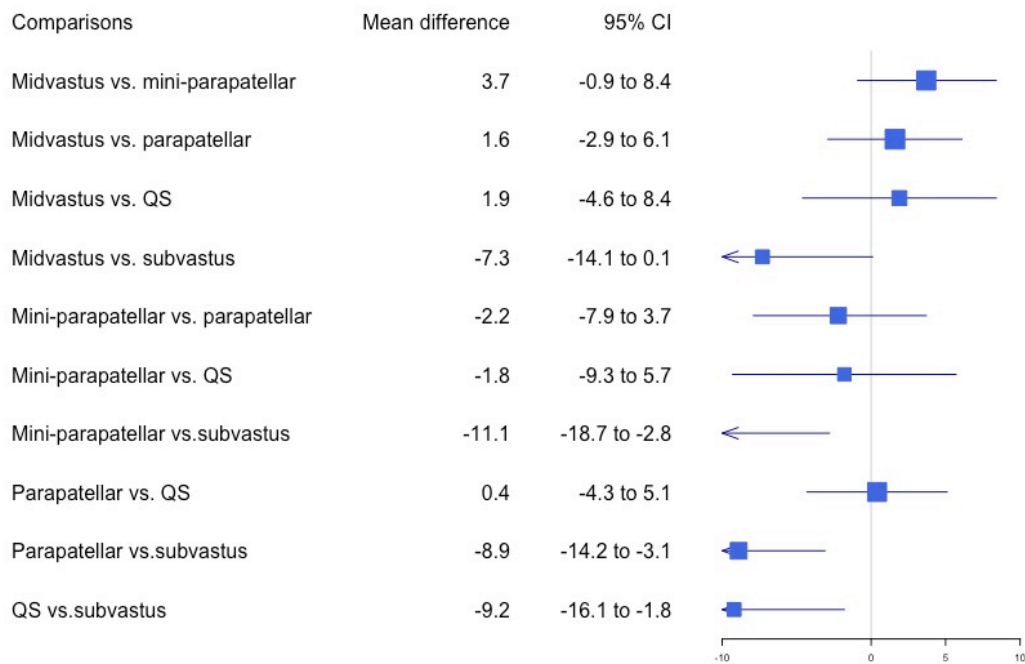
The bar-graph shows for the network estimate of every pairwise comparison how much information comes from low (green), moderate (yellow) or high (red) risk of bias trials.

Indirectness



Imprecision

We have some concerns for midvastus/mini-parapatellar, midvastus/parapatellar, mini-parapatellar/parapatellar because the credible intervals was large.



Heterogeneity/inconsistency

We have major concerns for the heterogeneity of parapatellar/mini-parapatellar comparison ($I^2 = 78.6\%$) and some concerns for subvastus/parapatellar ($I^2 = 65.5\%$).

We did not find inconsistency for this network.

Publication bias

We searched in ClinicalTrials.gov and in The World Health Organization International Clinical Trials Registry Platform (ICTRP) to locate potential unpublished (or gray) literature. So, we don't have serious concerns about the presence of publications bias.

Final assessments

Comparison	Number of studies	Within-study bias	Reporting bias	Indirectness	Imprecision	Heterogeneity	Incoherence	Confidence rating
parapatellar:QS	3	Some concerns	Undetected	No concerns	No concerns	No concerns	No concerns	Moderate
midvastus:mini-parapatellar	3	Some concerns	Undetected	No concerns	Some concerns	No concerns	No concerns	Moderate
midvastus:parapatellar	5	Some concerns	Undetected	No concerns	Some concerns	No concerns	No concerns	Moderate
mini-parapatellar:parapatellar	2	Some concerns	Undetected	No concerns	Some concerns	Major concerns	No concerns	Low
parapatellar:subvastus	4	Some concerns	Undetected	No concerns	No concerns	Some concerns	No concerns	Moderate

1. Salanti G, Giovane CD, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. *PLoS One*. 2014;9:e99682.
2. Confidence in Network Meta-Analysis [Software]. Institute of Social and Preventive Medicine, University of Bern, 2017. Available from cinema.ispm.unibe.ch. Accessed June 7, 2018.

Supplementary appendix for:

Article 1: Are There Differences in Accuracy or Outcomes Scores Among Navigated, Robotic, Patient-specific Instruments or Standard Cutting Guides in TKA? A Network Meta-analysis

Appendix 1. Search terms

- PubMed

((((((((((arthroplasty, replacement, knee[MeSH Terms]) OR knee prosthesis[MeSH Terms]) OR ((knee[Title/Abstract]) AND replacement[Title/Abstract]) AND arthroplasty[Title/Abstract])) OR ((total[Title/Abstract]) AND knee[Title/Abstract]) AND arthroplasty[Title/Abstract])) OR ((total[Title/Abstract]) AND knee[Title/Abstract]) AND replacement[Title/Abstract])) OR TKA[Title/Abstract]) OR TKR[Title/Abstract]) OR ((knee[Title/Abstract]) AND prosthesis[Title/Abstract])) AND (((((((((((navigation[Title/Abstract]) OR navigated[Title/Abstract]) OR ((imageless[Title/Abstract]) AND computer[Title/Abstract]) AND assisted[Title/Abstract])) OR ((computer[Title/Abstract]) AND assisted[Title/Abstract])) OR ((computer[Title/Abstract]) AND navigation[Title/Abstract])) OR ((computer[Title/Abstract]) AND assisted[Title/Abstract]) AND techniques[Title/Abstract])) OR ((image[Title/Abstract]) AND free[Title/Abstract])) OR (((Accelerometric[Title/Abstract]) OR iAssist[Title/Abstract]) OR accelerometers[Title/Abstract]) OR accelerometer-based[Title/Abstract])) OR (((iblock[Title/Abstract]) OR mako[Title/Abstract]) OR (((robot assisted[Title/Abstract]) OR robotic guidance[Title/Abstract]) OR ((RIO[Title/Abstract]) OR NAVIO[Title/Abstract]) OR ROBODOC[Title/Abstract])) OR (((((((((PSI[Title/Abstract]) OR (((patient[Title/Abstract]) AND specific[Title/Abstract]) AND cutting[Title/Abstract]) AND blocks[Title/Abstract])) OR ((patient[Title/Abstract]) AND specific[Title/Abstract]) AND instrumentation[Title/Abstract])) OR ((patient[Title/Abstract]) AND specific[Title/Abstract]) AND instruments[Title/Abstract])) OR ((patient[Title/Abstract]) AND specific[Title/Abstract]) AND guides[Title/Abstract])) OR ((custom[Title/Abstract]) AND fit[Title/Abstract]) AND instrumentation[Title/Abstract])) OR ((custom[Title/Abstract]) AND made[Title/Abstract]) AND guides[Title/Abstract])) OR ((patient[Title/Abstract]) AND matched[Title/Abstract]) AND instrumentation[Title/Abstract])) OR ((customized[Title/Abstract]) AND instrumentation[Title/Abstract])) OR (((patient[Title/Abstract]) AND specific[Title/Abstract]) AND templating[Title/Abstract])) AND (((((((((((randomized controlled trial[Publication Type]) OR controlled clinical trial[Publication Type]) OR randomized[Title/Abstract]) OR placebo[Title/Abstract]) OR drug therapy[MeSH Subheading]) OR randomly[Title/Abstract]) OR trial[Title/Abstract]) OR groups[Title/Abstract])) NOT ((animals[MeSH Terms]) NOT humans[MeSH Terms]))

- Embase

#43 #9 AND #42
#42 #17 OR #25 OR #30 OR #41
#41 #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40
#40 'patient':ab,ti AND 'specific':ab,ti AND 'templating':ab,ti AND [randomized controlled trial]/lim
#39 'customized':ab,ti AND 'instrumentation':ab,ti AND [randomized controlled trial]/lim
#38 'patient':ab,ti AND 'matched':ab,ti AND 'instrumentation':ab,ti AND [randomized controlled trial]/lim

#37 'custom':ab,ti AND 'made':ab,ti AND 'guides':ab,ti AND [randomized controlled trial]/lim

#36 'custom':ab,ti AND 'fit':ab,ti AND 'instrumentation':ab,ti AND [randomized controlled trial]/lim

#35 'patient':ab,ti AND 'specific':ab,ti AND 'guides':ab,ti AND [randomized controlled trial]/lim

#34 'patient':ab,ti AND 'specific':ab,ti AND 'instruments':ab,ti AND [randomized controlled trial]/lim

#33 'patient':ab,ti AND 'specific':ab,ti AND 'instrumentation':ab,ti AND [randomized controlled trial]/lim

#32 'patient':ab,ti AND 'specific':ab,ti AND 'cutting':ab,ti AND 'blocks':ab,ti AND [randomized controlled trial]/lim

#31 'psi':ab,ti AND [randomized controlled trial]/lim

#30 #26 OR #27 OR #28 OR #29

#29 'accelerometer':ab,ti AND 'based':ab,ti AND [randomized controlled trial]/lim

#28 'accelerometers':ab,ti AND [randomized controlled trial]/lim

#27 'iassist':ab,ti AND [randomized controlled trial]/lim

#26 'accelerometric':ab,ti AND [randomized controlled trial]/lim

#25 #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24

#24 'iblock':ab,ti AND [randomized controlled trial]/lim

#23 'mako':ab,ti AND [randomized controlled trial]/lim

#22 'rio':ab,ti AND [randomized controlled trial]/lim

#21 'navio':ab,ti AND [randomized controlled trial]/lim

#20 'robodoc':ab,ti AND [randomized controlled trial]/lim

#19 'robotic':ab,ti AND 'guidance':ab,ti AND [randomized controlled trial]/lim

#18 'robot':ab,ti AND 'assisted':ab,ti AND [randomized controlled trial]/lim

#17 #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16

#16 'imageless':ab,ti AND 'computer':ab,ti AND 'assisted':ab,ti AND [randomized controlled trial]/lim

#15 'image':ab,ti AND 'free':ab,ti AND [randomized controlled trial]/lim

#14 'navigated':ab,ti AND [randomized controlled trial]/lim

#13 'navigation':ab,ti AND [randomized controlled trial]/lim

#12 'computer':ab,ti AND 'assisted':ab,ti AND 'techniques':ab,ti AND [randomized controlled trial]/lim

#11 'computer':ab,ti AND 'navigation':ab,ti AND [randomized controlled trial]/lim

#10 'computer':ab,ti AND 'assisted':ab,ti AND [randomized controlled trial]/lim

#9 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8

#8 'knee':ab,ti AND 'arthroplasty':ab,ti AND 'replacement':ab,ti

#7 'knee':ti,ab,kw AND 'prosthesis':ti,ab,kw

#6 'tkr':ti,ab,kw

#5 'tka':ti,ab,kw

#4 'total':ti,ab,kw AND 'knee':ti,ab,kw AND 'replacement':ti,ab,kw

#3 'total':ti,ab,kw AND 'knee':ti,ab,kw AND 'arthroplasty':ti,ab,kw

#2 'total knee arthroplasty'/exp

#1 'knee prosthesis'/exp

- CENTRAL

- #1 MeSH descriptor: [Arthroplasty, Replacement, Knee] explode all trees
- #2 MeSH descriptor: [Knee Prosthesis] explode all trees
- #3 total:ti,ab,kw and knee and arthroplasty (Word variations have been searched)
- #4 total:ti,ab,kw and knee and replacement (Word variations have been searched)
- #5 tka:ti,ab,kw (Word variations have been searched)
- #6 tkr:ti,ab,kw (Word variations have been searched)
- #7 knee:ti,ab,kw and prosthesis:ti,ab,kw (Word variations have been searched)
- #8 knee:ti,ab,kw and arthroplasty:ti,ab,kw and replacement:ti,ab,kw (Word variations have been searched)
- #9 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 in Trials
- #10 robot:ti,ab,kw and assisted:ti,ab,kw (Word variations have been searched)
- #11 robotic:ti,ab,kw and guidance:ti,ab,kw (Word variations have been searched)
- #12 rio:ti,ab,kw (Word variations have been searched)
- #13 robodoc:ti,ab,kw (Word variations have been searched)
- #14 navio:ti,ab,kw (Word variations have been searched)
- #15 MAKO:ti,ab,kw (Word variations have been searched)
- #16 iBlock:ti,ab,kw (Word variations have been searched)
- #17 #10 or #11 or #12 or #13 or #14 or #15 or #16 in Trials
- #18 accelerometric:ti,ab,kw (Word variations have been searched)
- #19 iAssist:ti,ab,kw (Word variations have been searched)
- #20 accelerometers:ti,ab,kw (Word variations have been searched)
- #21 accelerometer:ti,ab,kw and based:ti,ab,kw (Word variations have been searched)
- #22 #18 or #19 or #20 or #21 in Trials
- #23 navigation:ti,ab,kw (Word variations have been searched)
- #24 navigated:ti,ab,kw (Word variations have been searched)
- #25 imageless:ti,ab,kw and computer:ti,ab,kw and assisted:ti,ab,kw (Word variations have been searched)
- #26 computer:ti,ab,kw and assisted:ti,ab,kw (Word variations have been searched)
- #27 computer:ti,ab,kw and navigation:ti,ab,kw (Word variations have been searched)
- #28 computer:ti,ab,kw and assisted:ti,ab,kw and techniques:ti,ab,kw (Word variations have been searched)
- #29 image:ti,ab,kw and free:ti,ab,kw (Word variations have been searched)
- #30 #23 or #24 or #25 or #26 or #27 or #28 or #29 in Trials
- #31 PSI:ti,ab,kw (Word variations have been searched)
- #32 patient:ti,ab,kw and specific:ti,ab,kw and cutting:ti,ab,kw and blocks:ti,ab,kw (Word variations have been searched)
- #33 patient:ti,ab,kw and specific:ti,ab,kw and instrumentation:ti,ab,kw (Word variations have been searched)
- #34 patient:ti,ab,kw and specific:ti,ab,kw and instruments:ti,ab,kw (Word variations have been searched)
- #35 patient:ti,ab,kw and specific:ti,ab,kw and guides:ti,ab,kw (Word variations have been searched)
- #36 custom:ti,ab,kw and fit:ti,ab,kw and instrumentation:ti,ab,kw (Word variations have been searched)
- #37 custom:ti,ab,kw and made:ti,ab,kw and guides:ti,ab,kw (Word variations have been searched)
- #38 patient:ti,ab,kw and matched:ti,ab,kw and instrumentation:ti,ab,kw (Word variations have been searched)
- #39 customized:ti,ab,kw and instrumentation:ti,ab,kw (Word variations have been searched)

- #40 patient:ti,ab,kw and specific:ti,ab,kw and templating:ti,ab,kw (Word variations have been searched)
- #41 #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 in Trials
- #42 #17 or #22 or #30 or #41 in Trials
- #43 #9 and #42

Appendix 2. Outcomes definition

Radiological outcomes:

Seven radiological data were extracted:

- The proportion at 6 months post-surgery of malalignments in the hip-knee-ankle (HKA) angle, which measures the mechanical axis of the lower limb. Malalignments were defined by an HKA angle higher or lower than $180^{\circ} \pm 3^{\circ}$.
- The proportion at 6 months post-surgery of malalignments for tibial/femoral implant in the frontal plane, defined by an angle higher or lower than $90 \pm 3^{\circ}$.
- The proportion at 6 months post-surgery of malalignments for tibial/femoral implant in the sagittal plane, defined by an angle higher or lower than $90 \pm 3^{\circ}$.
- The proportion at 6 months post-surgery of malalignments for femoral and tibial implant rotation in the axial plane, defined by more or less 3° from neutral rotation.

The HKA was the angle formed by the lines uniting at the center of the femoral head (hip) to the middle of the knee (knee) and then from the center of the knee to the middle of the ankle (ankle) [1]. HKA was measured using an x-ray of the lower limb.

Tibial and femoral alignments are defined in the frontal plane [2, 3]. They are evaluated using standard AP lower limb radiographs or CT scan.

Proximal tibial alignment is the angle between the mechanical axis of the tibia and the tangent to the tibial plateau in the frontal plane.

Distal femoral alignment is the angle between the mechanical axis of the femur and the tangent to the femoral condyles, in the frontal plane.

Alignment of the femoral and tibial components in the sagittal plane were evaluated with a classical lateral radiograph or using CT-scan. For the tibial component, the critical feature was the tibial slope, defined by the angle between the tibial shaft axis and the tangent to the tibial plateau. Regarding the femoral component, the main parameter in the sagittal plane was the angle between the femoral shaft axis and the tangent to the femoral condyles. Tibial and femoral implant rotations were analyzed on a postoperative CT-scan.

Femoral component rotation was evaluated using the angle between the surgical transepicondylar axis and the tangent to the implant's posterior condyles. Implant tibial rotation

was calculated relative to the junction between the central and medial third of the anterior tibial tuberosity and the best fit with the anterior tibial cortex.

References

1. Luo C-F. Reference axes for reconstruction of the knee. *Knee*. 2004;11:251–257.
2. Abdel MP, Oussedik S, Parratte S, Lustig S, Haddad FS. Coronal alignment in total knee replacement: historical review, contemporary analysis, and future direction. *Bone Joint J*. 2014;96:857–862.
3. Fang DM, Ritter MA, Davis KE. Coronal alignment in total knee arthroplasty: just how important is it? *J Arthroplasty*. 2009;24(6 Suppl):39–43.

B. KSS score

The Knee Society Score (KSS) is composed of a clinical part and a functional part (1). In the clinical part, seven items are evaluated by the clinician during the last 4 weeks (about 100 points): pain, flexion contracture, extension lag, total range of flexion, alignment (varus and valgus), AP stability, and mediolateral stability. In the functional part, three items are evaluated by the patient: walking, stairs, and walking aids used (about 100 points). The total score is calculated by summing the results of clinical part and functional part: a high score is related to little knee joint damage.

Knee Society Score																															
Part 1- Knee score :																															
Pain : <ul style="list-style-type: none"> • None • Mild/occasional • Mild (stairs only) • Mild (walking and stairs) • Moderate-occasional • Moderate-Continual • Severe 	Flexion contracture (if present) : <ul style="list-style-type: none"> • 5-10° • 10-15° • 16-20° • > 20° 																														
	Extension lag : <ul style="list-style-type: none"> • < 10° • 10-20° • > 20° 																														
Total range of flexion : <table style="width: 100%; border: none;"> <tr> <td style="width: 33%;">• 0-5°</td> <td style="width: 33%;">• 51-55°</td> <td style="width: 33%;">• 101-105°</td> </tr> <tr> <td>• 6-10°</td> <td>• 56-60°</td> <td>• 106-110°</td> </tr> <tr> <td>• 11-15°</td> <td>• 61-65°</td> <td>• 111-115°</td> </tr> <tr> <td>• 16-20°</td> <td>• 66-70°</td> <td>• 116-120°</td> </tr> <tr> <td>• 21-25°</td> <td>• 71-75°</td> <td>• 121-125°</td> </tr> <tr> <td>• 26-30°</td> <td>• 76-80°</td> <td></td> </tr> <tr> <td>• 31-35°</td> <td>• 81-85°</td> <td></td> </tr> <tr> <td>• 36-40°</td> <td>• 86-90°</td> <td></td> </tr> <tr> <td>• 41-45°</td> <td>• 91-95°</td> <td></td> </tr> <tr> <td>• 46-50°</td> <td>• 96-100°</td> <td></td> </tr> </table>		• 0-5°	• 51-55°	• 101-105°	• 6-10°	• 56-60°	• 106-110°	• 11-15°	• 61-65°	• 111-115°	• 16-20°	• 66-70°	• 116-120°	• 21-25°	• 71-75°	• 121-125°	• 26-30°	• 76-80°		• 31-35°	• 81-85°		• 36-40°	• 86-90°		• 41-45°	• 91-95°		• 46-50°	• 96-100°	
• 0-5°	• 51-55°	• 101-105°																													
• 6-10°	• 56-60°	• 106-110°																													
• 11-15°	• 61-65°	• 111-115°																													
• 16-20°	• 66-70°	• 116-120°																													
• 21-25°	• 71-75°	• 121-125°																													
• 26-30°	• 76-80°																														
• 31-35°	• 81-85°																														
• 36-40°	• 86-90°																														
• 41-45°	• 91-95°																														
• 46-50°	• 96-100°																														
Alignement (Varus-Valgus) : <table style="width: 100%; border: none;"> <tr> <td style="width: 50%;">• 0</td> <td style="width: 50%;">• 11</td> </tr> <tr> <td>• 1</td> <td>• 12</td> </tr> <tr> <td>• 2</td> <td>• 13</td> </tr> <tr> <td>• 3</td> <td>• 14</td> </tr> <tr> <td>• 4</td> <td>• 15</td> </tr> <tr> <td>• 5-10</td> <td>• Over 15°</td> </tr> </table>		• 0	• 11	• 1	• 12	• 2	• 13	• 3	• 14	• 4	• 15	• 5-10	• Over 15°																		
• 0	• 11																														
• 1	• 12																														
• 2	• 13																														
• 3	• 14																														
• 4	• 15																														
• 5-10	• Over 15°																														
Stability : <table style="width: 100%; border: none;"> <tr> <td style="width: 50%;"> • AP: <ul style="list-style-type: none"> ○ < 5mm ○ 5-10 mm ○ 10+ mm </td> <td style="width: 50%;"> • Mediolateral : <ul style="list-style-type: none"> ○ < 5° ○ 6-9° ○ 10-14° </td> </tr> </table>		• AP: <ul style="list-style-type: none"> ○ < 5mm ○ 5-10 mm ○ 10+ mm 	• Mediolateral : <ul style="list-style-type: none"> ○ < 5° ○ 6-9° ○ 10-14° 																												
• AP: <ul style="list-style-type: none"> ○ < 5mm ○ 5-10 mm ○ 10+ mm 	• Mediolateral : <ul style="list-style-type: none"> ○ < 5° ○ 6-9° ○ 10-14° 																														

○ 15°

Part 2- Function score :

Walking : <ul style="list-style-type: none">○ Unlimited○ > 10 blocks○ 5-10 blocks○ < 5 blocks○ Housebound○ Unable	Stairs : <ul style="list-style-type: none">○ Normal up and down○ Normal up down with rail○ Up and down with rail○ Up with rail, down unable○ Unable	Walking aids used : <ul style="list-style-type: none">○ None used○ Use of cane/walking stick○ Two canes/sticks○ Crutches frame
---	--	--

1. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res.* 1989;248:13-4.

C. WOMAC score

The WOMAC is a self-questionnaire completed by the patient [1]. It evaluates pain, stiffness and function in patients suffering from lower limb osteoarthritis. There are 24 questions rated from 0 to 4: five about the pain, two about the stiffness and 17 about the function. The total score is calculated by summing the results of each question: 0 is a knee or a hip with normal function, and the maximum score of 96 corresponds to a nonfunctional and very painful joint.

Instructions: please rate the activities in each category according to the following scale of difficulty: 0 = none, 1 = slight, 2 = moderate, 3 = very, 4 = extremely.

Circle one number for each activity.

Pain	1. Walking	0	1	2	3	4
	2. Stair climbing	0	1	2	3	4
	3. Nocturnal	0	1	2	3	4
	4. Rest	0	1	2	3	4
	5. Weight bearing	0	1	2	3	4
Stiffness	1. Morning stiffness	0	1	2	3	4
	2. Stiffness occurring later in the day	0	1	2	3	4
Physical function	1. Descending stairs	0	1	2	3	4
	2. Ascending stairs	0	1	2	3	4
	3. Rising from sitting	0	1	2	3	4
	4. Standing	0	1	2	3	4
	5. Bending to floor	0	1	2	3	4
	6. Walking on flat surface	0	1	2	3	4
	7. Getting in/out of car	0	1	2	3	4
	8. Going shopping	0	1	2	3	4
	9. Putting on socks	0	1	2	3	4
	10. Lying in bed	0	1	2	3	4
	11. Taking off socks	0	1	2	3	4
	12. Rising from bed	0	1	2	3	4
	13. Getting in/out of bath	0	1	2	3	4
	14. Sitting	0	1	2	3	4
	15. Getting on/off toilet	0	1	2	3	4
	16. Heavy domestic duties	0	1	2	3	4
	17. Light domestic duties	0	1	2	3	4
Total Score	/96 = %					

1. McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. *Arthritis Rheum.* 2001;45:453–61.

Appendix 3. Characteristics and references of studies

Study	Country	Funding	Cutting guide	Age (years) (mean)	BM I (kg/m ²) (mean)	Sex (women)	Numbers of patients	Numbers of knees	Prosthesis	System
Abane [1]	France	No specific funding	Standard PSI	69.3	28.6	88	140	140	Genesis II (Smith & Nephew, London, UK)	Visionaire (Smith & Nephew, London, UK)
Abane [2]	France	Private	Standard PSI	69.6	27.3	107	140	140	Global Medacta (Medacta, Swiss, Castel San Pietro, Switzerland)	MyKnee system (Medacta, Swiss)
Bathis [3]	Germany	No specific funding	Standard Navigation	69.8		112	160	160	PFC [®] implants (DePuy Synthes, Raynham, MA, USA)	VectorVision Knee navigation (BrainLab, Munich, Germany)
Blakeney [4]	Australia	Private	Standard Navigation	70.2		58	107	107	Genesis II (Smith & Nephew)	BrainLab navigation system (BrainLab)
Blyth [5]	Scotland	Private	Standard Navigation	65.5		116	200	200	NexGen [®] LPS Flex (Zimmer Biomet, Warsaw, IN, USA)	iNav Electromagnetic navigation system (Medtronic, Minneapolis, MN, USA)
Bohling [6]	Germany	No specific funding	Standard Navigation	70.5		75	100	100		Surgetics System (DePuy Synthes)

Boonen [7]	Netherlands	No specific funding	Standard PSI	67	29.9	106	180	180	Vanguard® system (Zimmer Biomet)	Signature™ (Zimmer Biomet)
Boonen [8]	Netherlands	No specific funding	Standard PSI	67	29.9	106	180	180	Vanguard system (Zimmer Biomet)	Signature (Zimmer Biomet)
Chareancholvanich [9]	Thailand	No specific funding	Standard PSI	69.9	27.9	70	80	80	NexGen LPS Flex (Zimmer Biomet)	Patient Specific instruments System (Zimmer Biomet)
Chauhan [10]	Australia	Private	Standard Navigation				70	70	Duracon total knee prosthesis (Stryker, Kalamazoo, MI, USA)	Stryker navigation system (Stryker)
Chen [11]	Singapore	No specific funding	Standard Navigation	67	28	67	100	100	PFC implants (DePuy Synthes)	BrainLab navigation system 2.6 (BrainLab, Germany, Munich)
Chin [12]	Singapore	No specific funding	Standard Navigation	66.6		67	90	90	PFC implants (DePuy Synthes)	VectorVision Knee navigation (BrainLab)
Choong [13]	Australia	No specific funding	Standard Navigation	69.5	29.5	67	111	111	PFC implants (DePuy Synthes)	Ci navigation system (DePuy Synthes)
Chotanaphuti [14]	Thailand	No specific funding	Standard Navigation	67.5		163	180	180	PFC implants (DePuy Synthes)	BrainLab navigation system (BrainLab)
Chotanaphuti [15]	Thailand	No specific	Standard PSI	69.5	25	70	80	80	PFC implants	TruMatch® Personalized

		funding							(DePuy Synthes)	solutions (DePuy Synthes)
Confalonieri [16]	Italy	No specific funding	Standard Navigation	73.1	28	39	74	74	Genesis II (Smith & Nephew)	VectorVision Knee navigation (BrainLab)
Decking [17]	Germany	Private	Standard Navigation	66	29	35	52	52	Search Evolution (Aesculap, Tuttlingen, Germany)	OrthoPilot system (Aesculap)
Decking [18]	Germany	No specific funding	Standard Navigation	66	29	35	52	52	Search Evolution (Aesculap)	OrthoPilot system (Aesculap)
Dutton [19]	Singapore	No specific funding	Standard Navigation	67.5	27.3	88	108	108	PFC implants (DePuy Synthes)	Ci navigation system (BrainLab)
Ensini [20]	Italy	No specific funding	Standard Navigation	70	29.3	70	107	120	Scorpio (Stryker) Optetrak (Exatech, Gainesville, FL, USA)	Stryker navigation system (Stryker)
Gan [21]	China	No specific funding	Standard PSI	68.2		51	70	70	Scorpio (Stryker)	Patient specific navigational template (LightSpeed VCT; GE Healthcare, Milwaukee, WI, USA)
Gharaibeh [22]	Australia	No specific funding	Standard Accelerometer	69.1	29.4	105	190	190	Legion (Smith & Nephew) Score (Amplitude Valence, France)	KneeAlign (OrthoAlign, San Jose, CA, USA)
Gøthesen [23]	Norway	Public	Standard	68		115	192	192	CR Profix	VectorVision Knee

			Navigation						(Smith & Nephew)	navigation (BrainLab)
Hamilton [24]	USA	No specific funding	Standard PSI	67.9	31	31	52	52	PFC implants (DePuy Synthes)	TruMatch Personalized solutions (DePuy Synthes)
Han [25]	South Korea	No specific funding	Standard Navigation	70.5	27.5	54	55	55	PFC implants (DePuy Synthes)	OrthoPilot system (Aesculap)
Hasegawa [26]	Japan	No specific funding	Standard Navigation	73.5	26.7	80	100	100	Columbus Aesculap (Aesculap)	Orthopilot system (Aesculap)
Hernandez-Vaquero [27]	Spain	Public	Standard Navigation				79	80		Stryker navigation system (Stryker)
Hiscox [28]	Canada	No specific funding	Standard Navigation	67		79	120	141	Duracon total knee prosthesis (Stryker)	Stryker navigation system (Stryker)
Hsu [29]	Taiwan	No specific funding	Standard Navigation	70	27	41	60	120	PFC implants (DePuy Synthes)	BrainLab navigation system (BrainLab)
Huang [30]	Taiwan	Public	Standard Navigation	70	27.7	90	113	226		BrainLab navigation system (BrainLab)
Huijbregts [31]	Australia	Public	Standard PSI	67.9		68	107	133	Genesis II/Legion (Smith & Nephew)	Visionaire (Smith & Nephew)
Ikawa [32]	Japan	No specific funding	Standard Accelerometer	74	26.4	207	241	241	Vanguard system (Zimmer Biomet)	KneeAlign2 [®] (OrthAlign, Aliso Viejo, CA, USA)

Khuangsi rikul [33]	Thailand	No specific funding	Standard PSI				80	80	PFC implants (DePuy Synthes)	TruMatch Personalized solutions (DePuy Synthes)
Kim [34]	South Korea	No specific funding	Standard Navigation	67.6	27.1	15	100	200	PFC implants (DePuy Synthes)	VectorVision Knee navigation (BrainLab)
Kim [35]	South Korea	No specific funding	Standard Navigation	68	27.8	452	520	1040	NexGen LPS Flex (Zimmer Biomet) /PFC implants (DePuy Synthes)	VectorVision Knee navigation (BrainLab)
Kinney [36]	USA	Private	Standard Accelerometer	65.7	30.8	29	47	50	Persona components (Zimmer Biomet)	iAssist (Zimmer Biomet)
Kosse [37]	Netherlands	Private	Standard PSI	63.1	27.9	22	42	42	Genesis II (Smith & Nephew)	Visionaire (Smith & Nephew)
Kotela [38]	Poland	No specific funding	Standard PSI	67.3	29.8	66	95	95	Vanguard system (Zimmer Biomet)	Signature (Zimmer Biomet)
Kotela [39]	Poland	No specific funding	Standard PSI	67.3	29.8	66	95	95	Vanguard system (Zimmer Biomet)	Signature (Zimmer Biomet)
Lee [40]	South Korea	No specific funding	Standard Navigation	66.5		111	116	116	e.motion (B.Braun, Germany, Melsungen)	Orthopilot system (B.Braun, Germany, Melsungen)
Lee [41]	Singapore	No specific funding	Standard Navigation			67	90	90	PFC implants (DePuy Synthes)	VectorVision Knee navigation (BrainLab)

Lin [42]	Taiwan	No specific funding	Standard Navigation	68.7	26	86	100	100	NexGen Legacy (Zimmer Biomet)	BrainLab navigation system (BrainLab)
Lionberger [43]	USA	No specific funding	Navigation PSI				60	60	Vanguard system (Zimmer Biomet)	Signature (Zimmer Biomet)
Liow [44]	Singapore	No specific funding	Standard Robot	67.9	27.4		60	60	NexGen LPS (Zimmer Biomet)	Robodoc (THINK Surgical Inc., Fremont, CA)
Liow [45]	Singapore	No specific funding	Standard Robot	67.9	27.4		60	60	NexGen LPS (Zimmer Biomet)	Robodoc (Think Surgical Inc, Fremont, CA, USA)
Luring [46]	Germany	No specific funding	Standard Navigation	69.3	31.3		90	90	PFC implants (DePuy Synthes)	Colibri (DePuy Synthes)
Lützner [47]	Germany	Private	Standard Navigation	69	29.9	51	80	80	Scorpio (Stryker)	Stryker navigation system (Stryker)
Lützner [48]	Germany	Private	Standard Navigation	68.5	29.9	49	73	73	Scorpio (Stryker)	Stryker navigation system (Stryker)
Macule-Beneyto [49]	Spain	No specific funding	Standard Navigation	71.9	30.9		186	186		Stryker navigation system (Stryker)/Navitrack/OrthoPilot (Aesculap)
Maderbacher [50]	Germany	No specific funding	Standard Navigation	86.8	31.1	53	80	80	PFC implants (DePuy Synthes)	BrainLab navigation system 2.6 (BrainLab)

Matos [51]	Brazil	No specific funding	Standard Navigation				39	42	Search Evolution (Aesculap)	OrthoPilot system (Aesculap)
Matziolis [52]	Germany	No specific funding	Standard Navigation	70.5	31.1	40	60	60	TCPlus™ (Smith & Nephew)	PiGalileo (Plus Orthopedics Aarau, Switzerland)
Maus [53]	Germany	Private	Standard PSI	69.9	31.2	76	125	125	e.motion/Vega system/Columbus Knee system (Aesculap Implant Systems)	Imprint (Aesculap)
Mitsuhashi [54]	Japan	No specific funding	Standard Navigation	72.1	27.1	61	72	72	Scorpio (Stryker)	
Molicnik [55]	Slovenia	No specific funding	Standard PSI	67	32.6	31	38	38	Vanguard system (Zimmer Biomet)	Signature (Zimmer Biomet)
Mombert [56]	Belgium	No specific funding	Standard Navigation				42	42	Scorpio (Stryker)	VectorVision Knee navigation (BrainLab)
Mullaji [57]	India	No specific funding	Standard Navigation	66		360	467	467	PFC implants (DePuy Synthes)	Ci navigation system
Nam [58]	USA	Private	Standard Accelerometer	66.1	31.2	49	94	94		KneeAlign (OrthAlign)
Noble [59]	USA	Private	Standard PSI	66.6	30.5	15	29	29	Legion (Smith & Nephew)	Visionaire (Smith & Nephew)

Oberst [60]	Germany	No specific funding	Standard Navigation				62	69	PFC implants (DePuy Synthes)	VectorVision Knee navigation (BrainLab)
Ollivier [61]	France	No specific funding	Navigation PSI	73.7	26.9	56	80	80	Legion (Smith & Nephew)	Visionaire (Smith & Nephew)
Pang [62]	Singapore	No specific funding	Standard Navigation	69	28.9	118	140	140	PFC implants (DePuy Synthes)	Ci Mi
Parratte [63]	France	No specific funding	Standard PSI				40	40	NexGen LPS Flex (Zimmer Biomet)	
Perlick [64]	Germany	No specific funding	Standard Navigation	69		82	100	100	PFC implants (DePuy Synthes)	VectorVision Knee navigation (BrainLab)
Pietsch [65]	Austria	No specific funding	Standard PSI	70.3	29.9	48	80	80	NexGen LPS Flex (Zimmer Biomet)	Patient Specific instruments System (Zimmer Biomet)
Qiu [66]	China	Public	Standard PSI	66.3		20	26	26		
Roh [67]	South Korea	No specific funding	Standard PSI	70	27	82	90	90	Vanguard system (Zimmer Biomet)	Signature (Zimmer Biomet)
Saragaglia [68]	France	No specific funding	Standard Navigation	69.5		15	50	50	Search Evolution (Aesculap)	
Schmitt [69]	Germany	No specific funding	Standard Navigation	69.7	31.1	61	89	90	NexGen (Zimmer Biomet)/Sc	KneeTrac (Stryker)

		fun ding							orpio (Stryker)	
Seon [70]	South Korea	No spec ific fun ding	Standar d Navigat ion	66	26. 3	91	89	102	e.motion (Aesculap)	OrthoPilot system (Aesculap)
Seon [71]	South Korea	No spec ific fun ding	Standar d Navigat ion	67. 4		79	85	85	e.motion (Aesculap)	OrthoPilot system (Aesculap)
Shinozaki [72]	Japan	No spec ific fun ding	Standar d Navigat ion	74. 4		62	75	75	TCPlus (Smith & Nephew)	Galileo (Smith & Nephew)
Sparmann [73]	Germ any	Priv ate	Standar d Navigat ion	66. 8		167	240	240	Duracon total knee prosthesis (Stryker)	Stryker navigation system (Stryker)
Spencer [74]	Austr alia	No spec ific fun ding	Standar d Navigat ion				60	60	Duracon total knee prosthesis (Stryker)	Stryker navigation system (Stryker)
Stolarczy k [75]	Polan d	No spec ific fun ding	Standar d PSI	69. 9	31	40	58	58	Genesis II (Smith & Nephew)	Visionaire (Smith & Nephew)
Tammach ote [76]	Thaila nd	Pub lic and priv ate	Standar d PSI	72	25. 5	82	102	102	Genesis II (Smith & Nephew)	Visionaire (Smith & Nephew)
Thiengwit tayaporn [77]	Thaila nd	No spec ific fun ding	Standar d Navigat ion	67. 5	27. 1	141	151	160	NexGen HiFlex (Zimmer Biomet)	Zimmer computer- assisted solutions electromagnetic
Thiengwit tayaporn [78]	Thaila nd	Pub lic	Standar d Acceler ometer	67	26. 4	66	80	80	NexGen HiFlex (Zimmer Biomet)	iAssist (Zimmer Biomet)

Todesca [79]	Italy	No specific funding	Standard Navigation	75.3	30.2	165	225	240	Amplitude (Amplitude)	Amplivision® (Amplitude)
VanLeeuwen [80]	Norway	No specific funding	Standard PSI	65.4	29.9	62	94	94	Vanguard system (Zimmer Biomet)	Signature (Zimmer Biomet)
Victor [81]	Belgium	Private	Standard Navigation	71		74	100	100	Genesis II (Smith & Nephew)	Hardware ION
Victor [82]	Belgium	No specific funding	Standard PSI	66.5		86	128	128	Multiple	Multiple
Vide [83]	Portugal	No specific funding	Standard PSI	68.6	30.6	65	95	95	TCPlus (Smith & Nephew)	Visionaire (Smith & Nephew)
Woolson [84]	USA	Private	Standard PSI				63	64	PFC implants (DePuy Synthes)	TruMatch Personalized solutions (DePuy Synthes)
Yan [85]	China	No specific funding	Standard PSI Navigation	68.1		67	81	90	NexGen Legacy (Zimmer Biomet)	Imageless Ci
Zhang [86]	China	No specific funding	Standard Navigation	63		25	32	64	Gemini MK II (Link, Hamburg, Germany)	VectorVision Knee navigation (BrainLab)
Zhang [87]	China	No specific funding	Standard Navigation	65.2	25.5	42	82	82	PFC implants (DePuy Synthes)	Stryker navigation system (Stryker)
Zhang [88]	China	Public	Standard PSI	62.7		14	40	40	Triathlon prosthesis	

									(Stryker)	
Zhu [89]	Singapore	No specific funding	Standard Navigation	66.7	28.1	63	100	100	PFC implants (DePuy Synthes)	VectorVision Knee navigation (BrainLab)

References

1. Abane L, Anract P, Boisgard S, Descamps S, Courpied J, Hamadouche M. A comparison of patient-specific and conventional instrumentation for total knee arthroplasty: a multicentre randomised controlled trial. *Bone Joint J.* 2015;97:56-63.
2. Abane L, Zaoui A, Anract P, Lefevre N, Herman S, Hamadouche M. Can a single-use and patient-specific instrumentation be reliably used in primary total knee arthroplasty? A multicenter controlled study. *J Arthroplasty.* 2018;33:2111-2118.
3. B athis H, Perlick L, Tingart M, L uring C, Zurakowski D, Grifka J. Alignment in total knee arthroplasty. A comparison of computer-assisted surgery with the conventional technique. *J Bone Joint Surg Br.* 2004;86:682–687.
4. Blakeney WG, Khan RJK, Wall SJ. Computer-assisted techniques versus conventional guides for component alignment in total knee arthroplasty: A randomized controlled trial. *J Bone Joint Surg Am.* 2011;93:1377–1384.
5. Blyth MJG, Smith JR, Anthony IC, Strict NE, Rowe PJ, Jones BG. Electromagnetic navigation in total knee arthroplasty-A single center, randomized, single-blind study comparing the results with conventional techniques. *J Arthroplasty.* 2015;30:199–205.
6. B ohling U, Schamberger H, Grittner U, Scholz J. Computerised and technical navigation in total knee-arthroplasty. *J Orthop Traumatol.* 2005;6:69–75.
7. Boonen B, Schotanus MGM, Kerens B, van der Weegen W, van Drumpt R a. M, Kort NP. Intra-operative results and radiological outcome of conventional and patient-specific surgery in total knee arthroplasty: a multicentre, randomised controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2013;21:2206–2212.
8. Boonen B, Schotanus MGM, Kerens B, van der Weegen W, Hoekstra HJ, Kort NP. No difference in clinical outcome between patient-matched positioning guides and conventional instrumented total knee arthroplasty two years post-operatively: a multicentre, double-blind, randomised controlled trial. *Bone Joint J.* 2016;98:939–944.
9. Chareancholvanich K, Narkbunnam R, Pornrattanamaneewong C. A prospective randomised controlled study of patient-specific cutting guides compared with conventional instrumentation in total knee replacement. *Bone Joint J.* 2013;95:354–359.
10. Chauhan SK, Clark GW, Lloyd S, Scott RG, Breidahl W, Sikorski JM. Computer-assisted total knee replacement. A controlled cadaver study using a multi-parameter quantitative CT assessment of alignment (the Perth CT Protocol). *J Bone Joint Surg Br.* 2004;86:818–823.

11. Chen JY, Chin PL, Li Z, Yew AK, Tay DK, Chia S-L, Lo NN, Yeo SJ. Radiological outcomes of pinless navigation in total knee arthroplasty: a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2015;23:3556–3562.
12. Chin PL, Yang KY, Seo SJ, Lo NN. Randomized control trial comparing radiographic total knee arthroplasty implant placement using computer navigation versus conventional technique. *J Arthroplasty.* 2005;20:618–626.
13. Choong PF, Dowsey MM, Stoney JD. Does Accurate Anatomical Alignment Result in Better Function and Quality of Life? Comparing Conventional and Computer-Assisted Total Knee Arthroplasty. *J Arthroplasty.* 2009;24:560–569.
14. Chotanaphuti T, Panichcharoen W, Laoruengthana A. Comparative study of anatomical landmark technique and self-aligned tibial component rotation determined by computer-assisted TKA. *J Med Assoc Thai.* 2012;95(Suppl 10):S37-41.
15. Chotanaphuti T, Wangwittayakul V, Khuangsirikul S, Foojareonyos T. The accuracy of component alignment in custom cutting blocks compared with conventional total knee arthroplasty instrumentation: prospective control trial. *Knee.* 2014;21:185–188.
16. Confalonieri N, Manzotti A, Pullen C, Ragone V. Mini-incision versus mini-incision and computer-assisted surgery in total knee replacement: A radiological prospective randomised study. *Knee.* 2007;14:443–437.
17. Decking R, Markmann Y, Mattes T, Puhl W, Scharf HP. On the outcome of computer-assisted total knee replacement. *Acta Chir Orthop Traumatol Cech.* 2007;74:171–174.
18. Decking R, Markmann Y, Fuchs J, Puhl W, Scharf H-P. Leg axis after computer-navigated total knee arthroplasty: a prospective randomized trial comparing computer-navigated and manual implantation. *J Arthroplasty.* 2005;20:282–288.
19. Dutton AQ, Yeo S. Computer-assisted Minimally Invasive Total Knee Arthroplasty Compared with Standard Total Knee Arthroplasty: Surgical Technique. *J Bone Joint Surg Am.* 2009;91(Suppl 2 Pt 1):116–130.
20. Ensini A, Catani F, Leardini A, Romagnoli M, Giannini S. Alignments and clinical results in conventional and navigated total knee arthroplasty. *Clin Orthop Relat Res.* 2007;457:156–162.
21. Gan Y, Ding J, Xu Y, Hou C. Accuracy and efficacy of osteotomy in total knee arthroplasty with patient-specific navigational template. *Int J Clin Exp Med.* 2015;8:12192–12201.
22. Gharaibeh MA, Solayar GN, Harris IA, Chen DB, MacDessi SJ. Accelerometer-Based, Portable Navigation (KneeAlign) vs Conventional Instrumentation for Total Knee Arthroplasty: A Prospective Randomized Comparative Trial. *J Arthroplasty.* 2017;32:777–782.

23. Gøthesen Ø, Espehaug B, Havelin LI, Petursson G, Hallan G, Strøm E, Dyrhovden G, Furnes O. Functional outcome and alignment in computer-assisted and conventionally operated total knee replacements: A multicentre parallel-group randomised controlled trial. *Bone Joint J.* 2014;96:609–618.
24. Hamilton WG, Parks NL, Saxena A. Patient-specific instrumentation does not shorten surgical time: A prospective, randomized trial. *J Arthroplasty.* 2013;28(8 Suppl):96–100.
25. Han HS, Seong SC, Lee S, Lee MC. Rotational alignment of femoral components in total knee arthroplasty: Nonimage-based navigation system versus conventional technique. *Orthopedics.* 2006;29(10 Suppl):S148–51.
26. Hasegawa M, Yoshida K, Wakabayashi H, Sudo A. Minimally invasive total knee arthroplasty: comparison of jig-based technique versus computer navigation for clinical and alignment outcome. *Knee Surg Sports Traumatol Arthrosc.* 2011;19:904–910.
27. Hernandez-Vaquero D, Noriega-Fernandez A, Fernandez-Carreira JM, Fernandez-Simon JM, Llorens de los Rios J. Computer-assisted surgery improves rotational positioning of the femoral component but not the tibial component in total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc.* 2014;22:3127–3134.
28. Hiscox CM, Bohm ER, Turgeon TR, Hedden DR, Burnell CD. Randomized trial of computer-assisted knee arthroplasty: impact on clinical and radiographic outcomes. *J Arthroplasty.* 2011;26:1259–1264.
29. Hsu W-H, Hsu RWW, Weng Y-J. Effect of preoperative deformity on postoperative leg axis in total knee arthroplasty: a prospective randomized study. *Knee Surg Sports Traumatol Arthrosc.* 2010;18:1323–1327.
30. Huang T-W, Chuang P-Y, Lee C-Y, Lin S-J, Huang K-C, Shen S-H, Tsai YH, Lee MS, Hsu RW. Total knee arthroplasty in patients with Ranawat type-II valgus arthritic knee with a marked coronal femoral bowing deformity: comparison between computer-assisted surgery and intra-articular resection. *J Orthop Surg Res.* 2016;11:88.
31. Huijbregts HJTAM, Khan RJK, Fick DP, Hall MJ, Punwar SA, Sorensen E, Reid MJ, Vedove SD, Haebich S. Component alignment and clinical outcome following total knee arthroplasty: a randomised controlled trial comparing an intramedullary alignment system with patient-specific instrumentation. *Bone Joint J.* 2016;98:1043–1049.
32. Ikawa T, Takemura S, Kim M, Takaoka K, Minoda Y, Kadoya Y. Usefulness of an accelerometer-based portable navigation system in total knee arthroplasty. *Bone Joint J.* 2017;99:1047–1052.
33. Khuangsirikul S, Lertcharoenchoke T, Chotanaphuti T. Rotational alignment of femoral component between custom cutting block and conventional technique in total knee arthroplasty. *J Med Assoc Thai.* 2014;97(Suppl 2):S47-51.
34. Kim Y-H, Kim J-S, Yoon S-H. Alignment and orientation of the components in total knee replacement with and without navigation support. *J Bone Joint Surg Br.* 2007;89:471-476.

35. Kim Y, Park J, Kim J. Computer-navigated Versus Conventional Total Knee Arthroplasty: A Prospective Randomized Trial. *J Bone Joint Surg Am.* 2012;94:2017–2024.
36. Kinney MC, Cidambi KR, Severns DL, Gonzales FB. Comparison of the iAssist Handheld Guidance System to Conventional Instruments for Mechanical Axis Restoration in Total Knee Arthroplasty. *J Arthroplasty.* 2018;33:61–66.
37. Kosse NM, Heesterbeek PJC, Schimmel JJP, van Hellemond GG, Wymenga AB, Defoort KC. Stability and alignment do not improve by using patient-specific instrumentation in total knee arthroplasty: a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2018;26:1792–1799.
38. Kotela A, Lorkowski J, Kucharzewski M, Wilk-Frańczuk M, Śliwiński Z, Frańczuk B, Legosz P, Kotela I. Patient-Specific CT-Based Instrumentation versus Conventional Instrumentation in Total Knee Arthroplasty: a Prospective Randomized Controlled Study on Clinical Outcomes and In-Hospital Data. *Biomed Res Int.* 2015;2015:165908.
39. Kotela A, Kotela I. Patient-specific computed tomography based instrumentation in total knee arthroplasty: a prospective randomized controlled study. *Int Orthop.* 2014;38:2099-2107.
40. Lee D-H, Park J-H, Song D-I, Padhy D, Jeong W-K, Han S-B. Accuracy of soft tissue balancing in TKA: Comparison between navigation-assisted gap balancing and conventional measured resection. *Knee Surg Sports Traumatol Arthroscopy.* 2010;18:381–387.
41. Lee WT, Chin PL, Lo NN, Yeo SJ. Short-term outcome after computer-assisted versus conventional total knee arthroplasty: a randomised controlled trial. *J Orthop Surg (Hong Kong).* 2015;23:71–75.
42. Lin S-Y, Chen C-H, Fu Y-C, Huang P-J, Lu C-C, Su J-Y, et al. Comparison of the clinical and radiological outcomes of three minimally invasive techniques for total knee replacement at two years. *Bone Joint J.* 2013;95:906–910.
43. Lionberger DR. Patient specific instrumentation. *J Arthroplasty.* 2014;29:1699–1704.
44. Liow MHL, Xia Z, Wong MK, Tay KJ, Yeo SJ, Chin PL. Robot-assisted total knee arthroplasty accurately restores the joint line and mechanical axis: A prospective randomised study. *J Arthroplasty.* 2014;29:2373–2377.
45. Liow MHL, Goh GS-H, Wong MK, Chin PL, Tay DK-J, Yeo S-J. Robotic-assisted total knee arthroplasty may lead to improvement in quality-of-life measures: a 2-year follow-up of a prospective randomized trial. *Knee Surg Sports Traumatol Arthrosc.* 2017;25:2942–2951.
46. Luring C, Beckmann J, Haibock P, Perlick L, Grifka J, Tingart M. Minimal invasive and computer assisted total knee replacement compared with the conventional technique: a prospective, randomised trial. *Knee Surg Sports Traumatol Arthrosc.* 2008;16:928–934.
47. Lützner J, Günther K-P, Kirschner S. Functional outcome after computer-assisted versus conventional total knee arthroplasty: A randomized controlled study. *Knee Surg Sports Traumatol Arthroscopy.* 2010;18:1339–1344.

48. Lützner J, Krummenauer F, Wolf C, Günther K-P, Kirschner S. Computer-assisted and conventional total knee replacement: A comparative, prospective, randomised study with radiological and CT evaluation. *J Bone Joint Surg Br.* 2008;90:1039–1044.
49. Maculé-Beneyto F, Hernández-Vaquero D, Segur-Vilalta JM, Colomina-Rodríguez R, Hinarejos-Gomez P, García-Forcada I, Seral Garcia B. Navigation in total knee arthroplasty. A multicenter study. *Int Orthop.* 2006;30:536–540.
50. Maderbacher G, Schaumburger J, Keshmiri A, Barthel M, Springorum H-R, Craiovan B, Grifka J Baier C. Pinless navigation in total knee arthroplasty: navigation reduced by the maximum? *Int Orthop.* 2015;39:455–460.
51. Matos LFC, Alves ALQ, Sobreiro AL, Giordano MN, Albuquerque RSP de, Carvalho ACP. Navigation in total knee arthroplasty: is there any advantage? *Acta Ortopédica Brasileira.* 2011;19:184–188.
52. Matziolis G, Krockner D, Weiss U, Tohtz S, Perka C. A prospective, randomized study of computer-assisted and conventional total knee arthroplasty. Three-dimensional evaluation of implant alignment and rotation. *J Bone Joint Surg Am.* 2007;89:236–243.
53. Maus U, Marques CJ, Scheunemann D, Lampe F, Lazovic D, Hommel H, Vogel D, Haunschild M, Pfitzner T. No improvement in reducing outliers in coronal axis alignment with patient-specific instrumentation. *Knee Surg Sports Traumatol Arthrosc.* 2017;26:2788-2796.
54. Mitsuhashi S, Akamatsu Y, Kobayashi H, Kusayama Y, Kumagai K, Saito T. Combined CT-based and image-free navigation systems in TKA reduces postoperative outliers of rotational alignment of the tibial component. *Arch Orthop Trauma Surg.* 2018;138:259–266.
55. Molicnik A, Naranda J, Dolinar D. Patient-matched instruments versus standard instrumentation in total knee arthroplasty: a prospective randomized study. *Wien Klin Wochenschr.* 2015;127(Suppl 5):S235-240.
56. Mombert M, Van Den Daelen L, Gunst P, Missinne L. Navigated total knee arthroplasty: A radiological analysis of 42 randomised cases. *Acta Orthop Belg.* 2007;73:49–54.
57. Mullaji A, Kanna R, Marawar S, Kohli A, Sharma A. Comparison of limb and component alignment using computer-assisted navigation versus image intensifier-guided conventional total knee arthroplasty: a prospective, randomized, single-surgeon study of 467 knees. *J Arthroplasty.* 2007;22:953–959.
58. Nam D, Cody EA, Nguyen JT, Figgie MP, Mayman DJ. Extramedullary guides versus portable, accelerometer-based navigation for tibial alignment in total knee arthroplasty: A randomized, controlled trial: Winner of the 2013 HAP PAUL award. *J Arthroplasty.* 2014;29:288–294.
59. Noble JWJ, Moore CA, Liu N. The value of patient-matched instrumentation in total knee arthroplasty. *J Arthroplasty.* 2012;27:153–155.

60. Oberst M, Bertsch C, Konrad G, Lahm A, Holz U. CT analysis after navigated versus conventional implantation of TKA. *Arch Orthop Trauma Surg.* 2008;128:561–566.
61. Ollivier M, Tribot-Laspierre Q, Amzallag J, Boisrenoult P, Pujol N, Beaufils P. Abnormal rate of intraoperative and postoperative implant positioning outliers using “MRI-based patient-specific” compared to “computer assisted” instrumentation in total knee replacement. *Knee Surg Sports Traumatol Arthrosc.* 2016;24:3441–3447.
62. Pang H-N, Yeo S-J, Chong H-C, Chin P-L, Ong J, Lo N-N. Computer-assisted gap balancing technique improves outcome in total knee arthroplasty, compared with conventional measured resection technique. *Knee Surg Sports Traumatol Arthrosc.* 2011;19:1496–1503.
63. Parratte S, Blanc G, Boussemart T, Ollivier M, Le Corroller T, Argenson J-N. Rotation in total knee arthroplasty: No difference between patient-specific and conventional instrumentation. *Knee Surg Sports Traumatol Arthroscopy.* 2013;21:2213–2219.
64. Perlick L, Bathis H, Tingart M, Perlick C, Grifka J. Navigation in total-knee arthroplasty: CT-based implantation compared with the conventional technique. *Acta Orthop Scand.* 2004;75:464–470.
65. Pietsch M, Djahani O, Zweiger C, Plattner F, Radl R, Tschauner C, Hofmann S. Custom-fit minimally invasive total knee arthroplasty: effect on blood loss and early clinical outcomes. *Knee Surg Sports Traumatol Arthrosc.* 2013;21:2234-2240.
66. Qiu B, Liu F, Tang B, Deng B, Liu F, Zhu W, Zhen D, Xue M, Zhang M. Clinical Study of 3D Imaging and 3D Printing Technique for Patient-Specific Instrumentation in Total Knee Arthroplasty. *J Knee Surg.* 2017;30:822–828.
67. Roh YW, Kim TW, Lee S, Seong SC, Lee MC. Is TKA using patient-specific instruments comparable to conventional TKA? A randomized controlled study of one system knee. *Clin Orthop Relat Res.* 2013;471:3988–3995.
68. Saragaglia D, Picard F, Chaussard C, Montbarbon E, Leitner F, Cinquin P. Computer-assisted knee arthroplasty: comparison with a conventional procedure. Results of 50 cases in a prospective randomized study [in French]. *Rev Chir Orthop Reparatrice Appar Mot.* 2001;87:18–28.
69. Schmitt J, Hauk C, Kienapfel H, Pfeiffer M, Efe T, Fuchs-Winkelmann S, Heyse TJ. Navigation of total knee arthroplasty: rotation of components and clinical results in a prospectively randomized study. *BMC Musculoskelet Disord.* 2011;12:16.
70. Seon JK, Song EK. Navigation-Assisted Less Invasive Total Knee Arthroplasty Compared With Conventional Total Knee Arthroplasty. A Randomized Prospective Trial. *J Arthroplasty.* 2006;21:777–782.
71. Seon JK, Park SJ, Lee KB, Li G, Kozanek M, Song EK. Functional comparison of total knee arthroplasty performed with and without a navigation system. *Int Orthop.* 2009;33:987–990.

72. Shinozaki T, Gotoh M, Mitsui Y, Hirai Y, Okawa T, Higuchi F, Nagata K. Computer-assisted total knee arthroplasty: comparisons with the conventional technique. *Kurume Med J*. 2011;58:21–26.
73. Sparmann M, Wolke B, Czupalla H, Banzer D, Zink A. Positioning of total knee arthroplasty with and without navigation support. *J Bone Joint Surg Br*. 2003;85:830–835.
74. Spencer JM, Chauhan SK, Sloan K, Taylor A, Beaver RJ. Computer navigation versus conventional total knee replacement: no difference in functional results at two years. *J Bone Joint Surg Br*. 2007;89:477–480.
75. Stolarczyk A, Nagraba L, Mitek T, Stolarczyk M, Deszczynski JM, Jakucinski M. Does Patient-Specific Instrumentation Improve Femoral and Tibial Component Alignment in Total Knee Arthroplasty? A Prospective Randomized Study. *Adv Exp Med Biol*. 2018;1096:11-17.
76. Tammachote N, Panichkul P, Kanitnate S. Comparison of Customized Cutting Block and Conventional Cutting Instrument in Total Knee Arthroplasty: A Randomized Controlled Trial. *J Arthroplasty*. 2018;33:746-751.
77. Thiengwittayaporn S, Fusakul Y, Kangkano N, Jarupongprapa C, Charoenphandhu N. Hand-held navigation may improve accuracy in minimally invasive total knee arthroplasty: a prospective randomized controlled trial. *Int Orthop*. 2016;40:51-57.
78. Thiengwittayaporn S, Kanjanapiboonwong A, Junsee D. Midterm outcomes of electromagnetic computer-assisted navigation in minimally invasive total knee arthroplasty. *J Orthop Surg Res*. 2013;8:37
79. Todesca A, Garro L, Penna M, Bejui-Hugues J. Conventional versus computer-navigated TKA: a prospective randomized study. *Knee Surg Sports Traumatol Arthrosc*. 2017;25:1778–1783.
80. Van Leeuwen JAMJ, Snorrason F, Röhrli SM. No radiological and clinical advantages with patient-specific positioning guides in total knee replacement. *Acta Orthop*. 2018;89:89–94.
81. Victor J, Dujardin J, Vandenneucker H, Arnout N, Bellemans J. Patient-specific guides do not improve accuracy in total knee arthroplasty: A prospective randomized controlled trial knee. *Clin Orthop Relat Res*. 2014;472:263–271.
82. Victor J, Hoste D. Image-based computer-assisted total knee arthroplasty leads to lower variability in coronal alignment. *Clin Orthop Relat Res*. 2004;428:131–139.
83. Vide J, Freitas TP, Ramos A, Cruz H, Sousa JP. Patient-specific instrumentation in total knee arthroplasty: simpler, faster and more accurate than standard instrumentation-a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc*. 2017;25:2616–2621.
84. Woolson ST, Harris AHS, Wagner DW, Giori NJ. Component alignment during total knee arthroplasty with use of standard or custom instrumentation: A randomized clinical trial using computed tomography for postoperative alignment measurement. *J Bone Joint Surg Am*. 2014;96:366–372.

85. Yan CH, Chiu KY, Ng FY, Chan PK, Fang CX. Comparison between patient-specific instruments and conventional instruments and computer navigation in total knee arthroplasty: a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2015;23:3637–3645.
86. Zhang G, Chen J, Chai W, Liu M, Wang Y. Comparison between computer-assisted-navigation and conventional total knee arthroplasties in patients undergoing simultaneous bilateral procedures: a randomized clinical trial. *J Bone Joint Surg Am.* 2011;93:1190–1196.
87. Zhang YZ, Lu S, Zhang HQ, Jin ZM, Zhao JM, Huang J, Zhang ZF. Alignment of the lower extremity mechanical axis by computer-aided design and application in total knee arthroplasty. *Int J Comput Assist Radiol Surg.* 2016;11:1881–1890.
88. Zhang Z, Gu B, Zhu W, Zhu L, Li Q, Du Y. Minimal invasive and computer-assisted total knee replacement compared with the minimal invasive technique: a prospective, randomized trial with short-term outcomes. *Arch Orthop Trauma Surg.* 2014;134:65-71.
89. Zhu M, Chen JY, Chong HC, Pang HN, Tay DKJ, Chia SL, Lo NN, Yeo SJ. No Difference in Functional Outcomes after Total Knee Arthroplasty with or without Pinless Navigation. *J Knee Surg.* 2018;31:649-653.

Appendix 4. Characteristic risk of bias per study

Study	Sequence generation	Allocation concealment	Blinding of patients	Blinding of care providers	Blinding of outcome assessment	Incomplete radiological outcome	Incomplete clinical outcome	Incomplete per operative outcome	Selective reporting of radiological outcomes	Selective reporting of clinical outcomes	Selective reporting of per operative outcomes	Other risk of bias
Abane [1]	Low	Low	High	High	Low	High	High	Low	Low	Low	Low	Low
Abane [2]	Low	Low	High	High	Low	Low	Low	Low	Low	Low	Low	Low
Bathis [3]	High	High	Unclear	High	Unclear	Low		Low	Low			Low
Blakeney [4]	Low	Low	Low	High	Low	Low		Low	Low		Low	Low
Blyth [5]	Low	Unclear	Unclear	High	Low	Low	Low		Low	Low		Low
Bohling [6]	High	High	Unclear	High	Unclear	Low		Low	Low		Low	High
Boonen [7]	Low	Unclear	Low	Low	Low	Low		Low	Low		Low	Low
Boonen [8]	Low	Unclear	Low	Low	Low		Low			Low		Low
Chareancholvanich [9]	Unclear	Unclear	High	High	Low	Low		Low	Low		Low	Low
Chauhan [10]	Low	Low	Low	High	Low	Unclear		Low	Low		Low	Low
Chen [11]	Low	Unclear	Low	High	Low	Low		Low	Low		Low	Low
Chin [12]	Low	Low	Unclear	High	Low	Low		Low	Low		Low	Low
Choong [13]	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Low
Chotanaphuti [14]	Unclear	Unclear	Unclear	High	Low	High			Low	High		Low

Chotanap huti [15]	Uncl ear	Uncle ar	Unc lear	High	Low	Low		Low	Low		Low	Lo w
Confaloni eri [16]	Low	Low	Unc lear	High	Low	Low		Low	Low		Low	Lo w
Decking [17]	Low	Low	Hig h	High	Low	Low	Low	Low	Low	Low	Low	Lo w
Decking [18]	Low	Low	Hig h	High	Low		Low			Low		Lo w
Dutton [19]	Low	Uncle ar	Unc lear	High	Uncle ar	Uncle ar	Uncle ar	Low	Low	Low	Low	Unc lear
Ensini [20]	Uncl ear	Uncle ar	Unc lear	High	Low	Low			Low			Lo w
Gan [21]	Low	Uncle ar	Unc lear	High	Low	Low		Low	Low		Low	Lo w
Gharaibe h [22]	Low	Uncle ar	Low	High	Uncle ar	Low		Low	Low		Low	Lo w
Gøthesen [23]	Low	Low	Low	High	Low	Low	Low	Low	Low	Low	Low	Lo w
Hamilton [24]	Uncl ear	Uncle ar	Hig h	High	Uncle ar	Low		Low	Low		Low	Lo w
Han [25]	Low	Uncle ar	Unc lear	High	Uncle ar	Low			Low			Lo w
Hasegawa [26]	High	High	Unc lear	High	Uncle ar	Low	Low	Low	Low	Low	Low	Lo w
Hernande z- Vaquero [27]	High	Low	Unc lear	High	Uncle ar	Low		Low	Low		Low	Lo w
Hiscox [28]	Low	High	Low	High	Uncle ar	Low	Low		Low	Low		Lo w
Hsu [29]	High	High	Unc lear	High	Uncle ar	Low			Low			Lo w
Huang [30]	High	High	Unc lear	High	Low	Low		Low	Low		Low	Lo w
Huijbregt s [31]	Low	Low	Hig h	High	Low	Low		Low	Low		Low	Lo w
Ikawa [32]	Low	Low	Unc lear	High	Low	Low			Low			Lo w
Khuangsi rikul [33]	Uncl ear	Uncle ar	Unc lear	High	Uncle ar	Uncle ar			Low			Lo w

Kim [34]	Low	Low	Unclear	High	Low	Low		Low	Low		Low	Low
Kim [35]	Low	Unclear	Unclear	High	Low	Low		Low	Low		Low	Low
Kinney [36]	Low	Low	Low	High	Low	Low			Low			Low
Kosse [37]	Low	Low	Low	High	Low		Low	Low		Low	Low	Low
Kotela [38]	Unclear	Unclear	Unclear	High	Low	High			Low			Low
Kotela [39]	Unclear	Unclear	Unclear	High	Unclear		High			Low		Low
Lee [40]	Low	Low	Unclear	High	Low	Low	Low		Low	Low		Low
Lee [41]	Unclear	Unclear	Unclear	High	Unclear		High			Low		Low
Lin [42]	Low	Low	Unclear	High	Unclear	Low	Low		Low	Low		Low
Lionberger [43]	Unclear	Unclear	Unclear	High	Unclear		High			Low		High
Liow [44]	Low	High	High	High	Low	Low	Low	Low	Low	Low	Low	Low
Liow [45]	Low	High	High	High	Unclear		Low			Low		Low
Luring [46]	Low	Unclear	Unclear	High	Unclear	Low	Low	Low	Low	Low	Low	Low
Lützner [47]	Low	Low	Unclear	High	Unclear	Low		Low	Low		Low	Low
Lützner [48]	Low	Low	Unclear	High	Unclear		Low			Low		Low
Macule-Beneyto [49]	Low	Low	Unclear	High	Unclear	Low		Low	Low		Low	Low
Maderbacher [50]	Low	Low	Unclear	High	Unclear	Low		Low	Low		Low	Low
Matos [51]	High	Low	Unclear	High	Unclear	Low		Low	Low		Low	Low
Matziolis [52]	Low	Low	Unclear	High	Unclear	Low	Low	Low	Low	Low	Low	Low
Maus [53]	Low	Low	High	High	High	High	High	H	Low	Low	Low	Low

Mitsuhas hi [54]	Uncl ear	Uncle ar	Unc lear	High	Uncle ar	Low			Low		Low	Lo w
Molicnik [55]	Uncl ear	Uncle ar	Hig h	High	Uncle ar	Low			Low			Lo w
Mombert [56]	Low	Low	Unc lear	High	Low	Low			Low			Lo w
Mullaji [57]	High	High	Unc lear	High	Low	Low			Low			Lo w
Nam [58]	Low	Low	Low	High	Low	Low			Low			Lo w
Noble [59]	Uncl ear	Uncle ar	Hig h	High	Uncle ar			Low			Low	Lo w
Oberst [60]	High	High	Unc lear	High	High	Low		Low	Low		Low	Lo w
Ollivier [61]	Low	Low	Unc lear	High	Uncle ar	Low			Low			Lo w
Pang [62]	Low	Low	Low	High	Low		Low	Low		Low	Low	Lo w
Parratte [63]	Low	Low	Hig h	High	Low	Low			Low			Lo w
Perlick [64]	High	High	Unc lear	High	Low	Low		Low	Low		Low	Lo w
Pietsch [65]	Uncl ear	Low	Unc lear	High	Low		Low	Low		Low	Low	Lo w
Qiu [66]	Uncl ear	Uncle ar	Unc lear	High	Uncle ar	Low			Low			Lo w
Roh [87]	Low	Low	Unc lear	High	Uncle ar	High		Low	Low		Low	Lo w
Saragagli a [68]	Uncl ear	Uncle ar	Unc lear	High	Uncle ar	Low		Low	Low		Low	Lo w
Schmitt [69]	Low	Low	Unc lear	High	Low	Low	Low	Low	Low	Low	Low	Lo w
Seon [70]	High	High	Unc lear	High	Low	Low		Low	Low		Low	Lo w
Seon [71]	Low	Low	Unc lear	High	Low	Uncle ar	Uncle ar		Low	Low		Lo w
Shinozaki [72]	High	High	Unc lear	High	Uncle ar	Low		Low	Low		Low	Lo w
Sparmann [73]	High	High	Unc lear	High	Uncle ar	Low			Low			Lo w

Spencer [74]	Low	Low	Low	High	Low		Low			Low		Low
Stolarczyk [75]	Low	Low	Unclear	High	Unclear		Low	Low		Low	Low	Low
Tammachote [76]	Low	Low	Unclear	High	Low	Low	Low	Low	Low	Low	Low	Low
Thiengwitayaporn [77]	Unclear	Unclear	Unclear	High	Low	High	High	Low	Low	Low	Low	Low
Thiengwitayaporn [78]	Low	Low	Low	High	Low	Low		Low	Low		Low	Low
Todesca [79]	Low	Low	Unclear	High	Unclear	Low	Low		Low	Low		Low
VanLeeuwen [80]	Low	Low	Low	High	Unclear	Low		Low	Low		Low	Low
Victor [81]	Low	Low	High	High	High	Low	Low	Low	Low	Low	Low	Low
Victor [82]	Low	Low	High	High	Low	Low			Low			Low
Vide [83]	Low	Unclear	High	High	Low	Low		Low	Low		Low	Low
Woolson [84]	Low	Low	High	High	Low	High		High	Low		Low	Low
Yan [85]	Low	Low	High	High	Low	Low	Low	Low	Low	Low	Low	Low
Zhang [86]	Unclear	Unclear	Low	High	Low	Low		Low	Low		Low	Low
Zhang [87]	Unclear	Unclear	Unclear	High	Unclear		Low			Low		Low
Zhang [88]	Unclear	Unclear	High	High	Unclear			Low			Low	Low
Zhu [89]	Low	Low	Low	High	Low		Low	Low		Low	Low	Low

References

1. Abane L, Anract P, Boisgard S, Descamps S, Courpied J, Hamadouche M. A comparison of patient-specific and conventional instrumentation for total knee arthroplasty: a multicentre randomised controlled trial. *Bone Joint J.* 2015;97:56-63.
2. Abane L, Zaoui A, Anract P, Lefevre N, Herman S, Hamadouche M. Can a single-use and patient-specific instrumentation be reliably used in primary total knee arthroplasty? A multicenter controlled study. *J Arthroplasty.* 2018;33:2111-2118.
3. Bächli H, Perlick L, Tingart M, Lüring C, Zurakowski D, Grifka J. Alignment in total knee arthroplasty. A comparison of computer-assisted surgery with the conventional technique. *J Bone Joint Surg Br.* 2004;86:682–687.
4. Blakeney WG, Khan RJK, Wall SJ. Computer-assisted techniques versus conventional guides for component alignment in total knee arthroplasty: A randomized controlled trial. *J Bone Joint Surg Am.* 2011;93:1377–1384.
5. Blyth MJG, Smith JR, Anthony IC, Strict NE, Rowe PJ, Jones BG. Electromagnetic navigation in total knee arthroplasty-A single center, randomized, single-blind study comparing the results with conventional techniques. *J Arthroplasty.* 2015;30:199–205.
6. Böhling U, Schamberger H, Grittner U, Scholz J. Computerised and technical navigation in total knee-arthroplasty. *J Orthop Traumatol.* 2005;6:69–75.
7. Boonen B, Schotanus MGM, Kerens B, van der Weegen W, van Drumpt R a. M, Kort NP. Intra-operative results and radiological outcome of conventional and patient-specific surgery in total knee arthroplasty: a multicentre, randomised controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2013;21:2206–2212.
8. Boonen B, Schotanus MGM, Kerens B, van der Weegen W, Hoekstra HJ, Kort NP. No difference in clinical outcome between patient-matched positioning guides and conventional instrumented total knee arthroplasty two years post-operatively: a multicentre, double-blind, randomised controlled trial. *Bone Joint J.* 2016;98:939–944.
9. Chareancholvanich K, Narkbunnam R, Pornrattanamaneewong C. A prospective randomised controlled study of patient-specific cutting guides compared with conventional instrumentation in total knee replacement. *Bone Joint J.* 2013;95:354–359.
10. Chauhan SK, Clark GW, Lloyd S, Scott RG, Breidahl W, Sikorski JM. Computer-assisted total knee replacement. A controlled cadaver study using a multi-parameter quantitative CT assessment of alignment (the Perth CT Protocol). *J Bone Joint Surg Br.* 2004;86:818–823.
11. Chen JY, Chin PL, Li Z, Yew AK, Tay DK, Chia S-L, Lo NN, Yeo SJ. Radiological outcomes of pinless navigation in total knee arthroplasty: a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2015;23:3556–3562.
12. Chin PL, Yang KY, Seo SJ, Lo NN. Randomized control trial comparing radiographic total knee arthroplasty implant placement using computer navigation versus conventional technique. *J Arthroplasty.* 2005;20:618–626.

13. Choong PF, Dowsey MM, Stoney JD. Does Accurate Anatomical Alignment Result in Better Function and Quality of Life? Comparing Conventional and Computer-Assisted Total Knee Arthroplasty. *J Arthroplasty*. 2009;24:560–569.
14. Chotanaphuti T, Panichcharoen W, Laoruengthana A. Comparative study of anatomical landmark technique and self-aligned tibial component rotation determined by computer-assisted TKA. *J Med Assoc Thai*. 2012;95(Suppl 10):S37-41.
15. Chotanaphuti T, Wangwittayakul V, Khuangsirikul S, Foojareonyos T. The accuracy of component alignment in custom cutting blocks compared with conventional total knee arthroplasty instrumentation: prospective control trial. *Knee*. 2014;21:185–188.
16. Confalonieri N, Manzotti A, Pullen C, Ragone V. Mini-incision versus mini-incision and computer-assisted surgery in total knee replacement: A radiological prospective randomised study. *Knee*. 2007;14:443–437.
17. Decking R, Markmann Y, Mattes T, Puhl W, Scharf HP. On the outcome of computer-assisted total knee replacement. *Acta Chir Orthop Traumatol Cech*. 2007;74:171–174.
18. Decking R, Markmann Y, Fuchs J, Puhl W, Scharf H-P. Leg axis after computer-navigated total knee arthroplasty: a prospective randomized trial comparing computer-navigated and manual implantation. *J Arthroplasty*. 2005;20:282–288.
19. Dutton AQ, Yeo S. Computer-assisted Minimally Invasive Total Knee Arthroplasty Compared with Standard Total Knee Arthroplasty: Surgical Technique. *J Bone Joint Surg Am*. 2009;91(Suppl 2 Pt 1):116–130.
20. Ensini A, Catani F, Leardini A, Romagnoli M, Giannini S. Alignments and clinical results in conventional and navigated total knee arthroplasty. *Clin Orthop Relat Res*. 2007;457:156–162.
21. Gan Y, Ding J, Xu Y, Hou C. Accuracy and efficacy of osteotomy in total knee arthroplasty with patient-specific navigational template. *Int J Clin Exp Med*. 2015;8:12192–12201.
22. Gharaibeh MA, Solayar GN, Harris IA, Chen DB, MacDessi SJ. Accelerometer-Based, Portable Navigation (KneeAlign) vs Conventional Instrumentation for Total Knee Arthroplasty: A Prospective Randomized Comparative Trial. *J Arthroplasty*. 2017;32:777–782.
23. Gøthesen Ø, Espehaug B, Havelin LI, Petursson G, Hallan G, Strøm E, Dyrhovden G, Furnes O. Functional outcome and alignment in computer-assisted and conventionally operated total knee replacements: A multicentre parallel-group randomised controlled trial. *Bone Joint J*. 2014;96:609–618.
24. Hamilton WG, Parks NL, Saxena A. Patient-specific instrumentation does not shorten surgical time: A prospective, randomized trial. *J Arthroplasty*. 2013;28(8 Suple):96–100.

25. Han HS, Seong SC, Lee S, Lee MC. Rotational alignment of femoral components in total knee arthroplasty: Nonimage-based navigation system versus conventional technique. *Orthopedics*. 2006;29(10 Suppl):S148–51.
26. Hasegawa M, Yoshida K, Wakabayashi H, Sudo A. Minimally invasive total knee arthroplasty: comparison of jig-based technique versus computer navigation for clinical and alignment outcome. *Knee Surg Sports Traumatol Arthrosc*. 2011;19:904–910.
27. Hernandez-Vaquero D, Noriega-Fernandez A, Fernandez-Carreira JM, Fernandez-Simon JM, Llorens de los Rios J. Computer-assisted surgery improves rotational positioning of the femoral component but not the tibial component in total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc*. 2014;22:3127–3134.
28. Hiscox CM, Bohm ER, Turgeon TR, Hedden DR, Burnell CD. Randomized trial of computer-assisted knee arthroplasty: impact on clinical and radiographic outcomes. *J Arthroplasty*. 2011;26:1259–1264.
29. Hsu W-H, Hsu RWW, Weng Y-J. Effect of preoperative deformity on postoperative leg axis in total knee arthroplasty: a prospective randomized study. *Knee Surg Sports Traumatol Arthrosc*. 2010;18:1323–1327.
30. Huang T-W, Chuang P-Y, Lee C-Y, Lin S-J, Huang K-C, Shen S-H, Tsai YH, Lee MS, Hsu RW. Total knee arthroplasty in patients with Ranawat type-II valgus arthritic knee with a marked coronal femoral bowing deformity: comparison between computer-assisted surgery and intra-articular resection. *J Orthop Surg Res*. 2016;11:88.
31. Huijbregts HJTAM, Khan RJK, Fick DP, Hall MJ, Punwar SA, Sorensen E, Reid MJ, Vedove SD, Haebich S. Component alignment and clinical outcome following total knee arthroplasty: a randomised controlled trial comparing an intramedullary alignment system with patient-specific instrumentation. *Bone Joint J*. 2016;98:1043–1049.
32. Ikawa T, Takemura S, Kim M, Takaoka K, Minoda Y, Kadoya Y. Usefulness of an accelerometer-based portable navigation system in total knee arthroplasty. *Bone Joint J*. 2017;99:1047–1052.
33. Khuangsirikul S, Lertcharoenchoke T, Chotanaphuti T. Rotational alignment of femoral component between custom cutting block and conventional technique in total knee arthroplasty. *J Med Assoc Thai*. 2014;97(Suppl 2):S47-51.
34. Kim Y-H, Kim J-S, Yoon S-H. Alignment and orientation of the components in total knee replacement with and without navigation support. *J Bone Joint Surg Br*. 2007;89:471-476.
35. Kim Y, Park J, Kim J. Computer-navigated Versus Conventional Total Knee Arthroplasty: A Prospective Randomized Trial. *J Bone Joint Surg Am*. 2012;94:2017–2024.
36. Kinney MC, Cidambi KR, Severns DL, Gonzales FB. Comparison of the iAssist Handheld Guidance System to Conventional Instruments for Mechanical Axis Restoration in Total Knee Arthroplasty. *J Arthroplasty*. 2018;33:61–66.

37. Kosse NM, Heesterbeek PJC, Schimmel JJP, van Hellemond GG, Wymenga AB, Defoort KC. Stability and alignment do not improve by using patient-specific instrumentation in total knee arthroplasty: a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2018;26:1792–1799.
38. Kotela A, Lorkowski J, Kucharzewski M, Wilk-Frańczuk M, Śliwiński Z, Frańczuk B, Legosz P, Kotela I. Patient-Specific CT-Based Instrumentation versus Conventional Instrumentation in Total Knee Arthroplasty: a Prospective Randomized Controlled Study on Clinical Outcomes and In-Hospital Data. *Biomed Res Int.* 2015;2015:165908.
39. Kotela A, Kotela I. Patient-specific computed tomography based instrumentation in total knee arthroplasty: a prospective randomized controlled study. *Int Orthop.* 2014;38:2099-2107.
40. Lee D-H, Park J-H, Song D-I, Padhy D, Jeong W-K, Han S-B. Accuracy of soft tissue balancing in TKA: Comparison between navigation-assisted gap balancing and conventional measured resection. *Knee Surg Sports Traumatol Arthroscopy.* 2010;18:381–387.
41. Lee WT, Chin PL, Lo NN, Yeo SJ. Short-term outcome after computer-assisted versus conventional total knee arthroplasty: a randomised controlled trial. *J Orthop Surg (Hong Kong).* 2015;23:71–75.
42. Lin S-Y, Chen C-H, Fu Y-C, Huang P-J, Lu C-C, Su J-Y, et al. Comparison of the clinical and radiological outcomes of three minimally invasive techniques for total knee replacement at two years. *Bone Joint J.* 2013;95:906–910.
43. Lionberger DR. Patient specific instrumentation. *J Arthroplasty.* 2014;29:1699–1704.
44. Liow MHL, Xia Z, Wong MK, Tay KJ, Yeo SJ, Chin PL. Robot-assisted total knee arthroplasty accurately restores the joint line and mechanical axis: A prospective randomised study. *J Arthroplasty.* 2014;29:2373–2377.
45. Liow MHL, Goh GS-H, Wong MK, Chin PL, Tay DK-J, Yeo S-J. Robotic-assisted total knee arthroplasty may lead to improvement in quality-of-life measures: a 2-year follow-up of a prospective randomized trial. *Knee Surg Sports Traumatol Arthrosc.* 2017;25:2942–2951.
46. Luring C, Beckmann J, Haibock P, Perlick L, Grifka J, Tingart M. Minimal invasive and computer assisted total knee replacement compared with the conventional technique: a prospective, randomised trial. *Knee Surg Sports Traumatol Arthrosc.* 2008;16:928–934.
47. Lützner J, Günther K-P, Kirschner S. Functional outcome after computer-assisted versus conventional total knee arthroplasty: A randomized controlled study. *Knee Surg Sports Traumatol Arthroscopy.* 2010;18:1339–1344.
48. Lützner J, Krummenauer F, Wolf C, Günther K-P, Kirschner S. Computer-assisted and conventional total knee replacement: A comparative, prospective, randomised study with radiological and CT evaluation. *J Bone Joint Surg Br.* 2008;90:1039–1044.

49. Maculé-Beneyto F, Hernández-Vaquero D, Segur-Vilalta JM, Colomina-Rodríguez R, Hinarejos-Gomez P, García-Forcada I, Seral Garcia B. Navigation in total knee arthroplasty. A multicenter study. *Int Orthop*. 2006;30:536–540.
50. Maderbacher G, Schaumburger J, Keshmiri A, Barthel M, Springorum H-R, Craiovan B, Grifka J Baier C. Pinless navigation in total knee arthroplasty: navigation reduced by the maximum? *Int Orthop*. 2015;39:455–460.
51. Matos LFC, Alves ALQ, Sobreiro AL, Giordano MN, Albuquerque RSP de, Carvalho ACP. Navigation in total knee arthroplasty: is there any advantage? *Acta Ortopédica Brasileira*. 2011;19:184–188.
52. Matziolis G, Kroker D, Weiss U, Tohtz S, Perka C. A prospective, randomized study of computer-assisted and conventional total knee arthroplasty. Three-dimensional evaluation of implant alignment and rotation. *J Bone Joint Surg Am*. 2007;89:236–243.
53. Maus U, Marques CJ, Scheunemann D, Lampe F, Lazovic D, Hommel H, Vogel D, Haunschild M, Pfitzner T. No improvement in reducing outliers in coronal axis alignment with patient-specific instrumentation. *Knee Surg Sports Traumatol Arthrosc*. 2017;26:2788-2796.
54. Mitsuhashi S, Akamatsu Y, Kobayashi H, Kusayama Y, Kumagai K, Saito T. Combined CT-based and image-free navigation systems in TKA reduces postoperative outliers of rotational alignment of the tibial component. *Arch Orthop Trauma Surg*. 2018;138:259–266.
55. Molicnik A, Naranda J, Dolinar D. Patient-matched instruments versus standard instrumentation in total knee arthroplasty: a prospective randomized study. *Wien Klin Wochenschr*. 2015;127(Suppl 5):S235-240.
56. Mombert M, Van Den Daelen L, Gunst P, Missinne L. Navigated total knee arthroplasty: A radiological analysis of 42 randomised cases. *Acta Orthop Belg*. 2007;73:49–54.
57. Mullaji A, Kanna R, Marawar S, Kohli A, Sharma A. Comparison of limb and component alignment using computer-assisted navigation versus image intensifier-guided conventional total knee arthroplasty: a prospective, randomized, single-surgeon study of 467 knees. *J Arthroplasty*. 2007;22:953–959.
58. Nam D, Cody EA, Nguyen JT, Figgie MP, Mayman DJ. Extramedullary guides versus portable, accelerometer-based navigation for tibial alignment in total knee arthroplasty: A randomized, controlled trial: Winner of the 2013 HAP PAUL award. *J Arthroplasty*. 2014;29:288–294.
59. Noble JWJ, Moore CA, Liu N. The value of patient-matched instrumentation in total knee arthroplasty. *J Arthroplasty*. 2012;27:153–155.
60. Oberst M, Bertsch C, Konrad G, Lahm A, Holz U. CT analysis after navigated versus conventional implantation of TKA. *Arch Orthop Trauma Surg*. 2008;128:561–566.
61. Ollivier M, Tribot-Laspierre Q, Amzallag J, Boisrenoult P, Pujol N, Beaufils P. Abnormal rate of intraoperative and postoperative implant positioning outliers using “MRI-based

patient-specific” compared to “computer assisted” instrumentation in total knee replacement. *Knee Surg Sports Traumatol Arthrosc.* 2016;24:3441–3447.

62. Pang H-N, Yeo S-J, Chong H-C, Chin P-L, Ong J, Lo N-N. Computer-assisted gap balancing technique improves outcome in total knee arthroplasty, compared with conventional measured resection technique. *Knee Surg Sports Traumatol Arthrosc.* 2011;19:1496–1503.

63. Parratte S, Blanc G, Boussemart T, Ollivier M, Le Corroller T, Argenson J-N. Rotation in total knee arthroplasty: No difference between patient-specific and conventional instrumentation. *Knee Surg Sports Traumatol Arthroscopy.* 2013;21:2213–2219.

64. Perlick L, Bathis H, Tingart M, Perlick C, Grifka J. Navigation in total-knee arthroplasty: CT-based implantation compared with the conventional technique. *Acta Orthop Scand.* 2004;75:464–470.

65. Pietsch M, Djahani O, Zweiger C, Plattner F, Radl R, Tschauer C, Hofmann S. Custom-fit minimally invasive total knee arthroplasty: effect on blood loss and early clinical outcomes. *Knee Surg Sports Traumatol Arthrosc.* 2013;21:2234-2240.

66. Qiu B, Liu F, Tang B, Deng B, Liu F, Zhu W, Zhen D, Xue M, Zhang M. Clinical Study of 3D Imaging and 3D Printing Technique for Patient-Specific Instrumentation in Total Knee Arthroplasty. *J Knee Surg.* 2017;30:822–828.

67. Roh YW, Kim TW, Lee S, Seong SC, Lee MC. Is TKA using patient-specific instruments comparable to conventional TKA? A randomized controlled study of one system knee. *Clin Orthop Relat Res.* 2013;471:3988–3995.

68. Saragaglia D, Picard F, Chaussard C, Montbarbon E, Leitner F, Cinquin P. Computer-assisted knee arthroplasty: comparison with a conventional procedure. Results of 50 cases in a prospective randomized study [in French]. *Rev Chir Orthop Reparatrice Appar Mot.* 2001;87:18–28.

69. Schmitt J, Hauk C, Kienapfel H, Pfeiffer M, Efe T, Fuchs-Winkelmann S, Heyse TJ. Navigation of total knee arthroplasty: rotation of components and clinical results in a prospectively randomized study. *BMC Musculoskelet Disord.* 2011;12:16.

70. Seon JK, Song EK. Navigation-Assisted Less Invasive Total Knee Arthroplasty Compared With Conventional Total Knee Arthroplasty. A Randomized Prospective Trial. *J Arthroplasty.* 2006;21:777–782.

71. Seon JK, Park SJ, Lee KB, Li G, Kozanek M, Song EK. Functional comparison of total knee arthroplasty performed with and without a navigation system. *Int Orthop.* 2009;33:987–990.

72. Shinozaki T, Gotoh M, Mitsui Y, Hirai Y, Okawa T, Higuchi F, Nagata K. Computer-assisted total knee arthroplasty: comparisons with the conventional technique. *Kurume Med J.* 2011;58:21–26.

73. Sparmann M, Wolke B, Czupalla H, Banzer D, Zink A. Positioning of total knee arthroplasty with and without navigation support. *J Bone Joint Surg Br.* 2003;85:830–835.
74. Spencer JM, Chauhan SK, Sloan K, Taylor A, Beaver RJ. Computer navigation versus conventional total knee replacement: no difference in functional results at two years. *J Bone Joint Surg Br.* 2007;89:477–480.
75. Stolarczyk A, Nagraba L, Mitek T, Stolarczyk M, Deszczynski JM, Jakucinski M. Does Patient-Specific Instrumentation Improve Femoral and Tibial Component Alignment in Total Knee Arthroplasty? A Prospective Randomized Study. *Adv Exp Med Biol.* 2018;1096:11-17.
76. Tammachote N, Panichkul P, Kanitnate S. Comparison of Customized Cutting Block and Conventional Cutting Instrument in Total Knee Arthroplasty: A Randomized Controlled Trial. *J Arthroplasty.* 2018;33:746-751.
77. Thiengwittayaporn S, Fusakul Y, Kangkano N, Jarupongprapa C, Charoenphandhu N. Hand-held navigation may improve accuracy in minimally invasive total knee arthroplasty: a prospective randomized controlled trial. *Int Orthop.* 2016;40:51-57.
78. Thiengwittayaporn S, Kanjanapiboonwong A, Junsee D. Midterm outcomes of electromagnetic computer-assisted navigation in minimally invasive total knee arthroplasty. *J Orthop Surg Res.* 2013;8:37
79. Todesca A, Garro L, Penna M, Bejui-Hugues J. Conventional versus computer-navigated TKA: a prospective randomized study. *Knee Surg Sports Traumatol Arthrosc.* 2017;25:1778–1783.
80. Van Leeuwen JAMJ, Snorrason F, Röhrli SM. No radiological and clinical advantages with patient-specific positioning guides in total knee replacement. *Acta Orthop.* 2018;89:89–94.
81. Victor J, Dujardin J, Vandenneucker H, Arnout N, Bellemans J. Patient-specific guides do not improve accuracy in total knee arthroplasty: A prospective randomized controlled trial knee. *Clin Orthop Relat Res.* 2014;472:263–271.
82. Victor J, Hoste D. Image-based computer-assisted total knee arthroplasty leads to lower variability in coronal alignment. *Clin Orthop Relat Res.* 2004;428:131–139.
83. Vide J, Freitas TP, Ramos A, Cruz H, Sousa JP. Patient-specific instrumentation in total knee arthroplasty: simpler, faster and more accurate than standard instrumentation-a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2017;25:2616–2621.
84. Woolson ST, Harris AHS, Wagner DW, Giori NJ. Component alignment during total knee arthroplasty with use of standard or custom instrumentation: A randomized clinical trial using computed tomography for postoperative alignment measurement. *J Bone Joint Surg Am.* 2014;96:366–372.
85. Yan CH, Chiu KY, Ng FY, Chan PK, Fang CX. Comparison between patient-specific instruments and conventional instruments and computer navigation in total knee arthroplasty: a randomized controlled trial. *Knee Surg Sports Traumatol Arthrosc.* 2015;23:3637–3645.

86. Zhang G, Chen J, Chai W, Liu M, Wang Y. Comparison between computer-assisted-navigation and conventional total knee arthroplasties in patients undergoing simultaneous bilateral procedures: a randomized clinical trial. *J Bone Joint Surg Am*. 2011;93:1190–1196.
87. Zhang YZ, Lu S, Zhang HQ, Jin ZM, Zhao JM, Huang J, Zhang ZF. Alignment of the lower extremity mechanical axis by computer-aided design and application in total knee arthroplasty. *Int J Comput Assist Radiol Surg*. 2016;11:1881–1890.
88. Zhang Z, Gu B, Zhu W, Zhu L, Li Q, Du Y. Minimal invasive and computer-assisted total knee replacement compared with the minimal invasive technique: a prospective, randomized trial with short-term outcomes. *Arch Orthop Trauma Surg*. 2014;134:65-71.
89. Zhu M, Chen JY, Chong HC, Pang HN, Tay DKJ, Chia SL, Lo NN, Yeo SJ. No Difference in Functional Outcomes after Total Knee Arthroplasty with or without Pinless Navigation. *J Knee Surg*. 2018;31:649-653.

Appendix 5. Formulas used

When we had only the first quartile and the third quartile, we used the estimator of the Cochrane Handbook to obtain a estimation of the SD:

$$SD \approx \frac{q3-q1}{1.35}$$

When we had only the range, we used this formula to obtain a estimation of the SD [1]:

Sample size	$n \leq 15$	$15 < n \leq 70$	$70 < n$
Estimate SD	$s^2 \approx \frac{1}{12} \left(\frac{(a-2m+b)^2}{4} + (b-a)^2 \right)$	Range/4	Range/6

When we had only the median, we used this formula to obtain a estimation of the mean [1]:

Sample size	$15 < n \leq 25$	$25 < n$
Estimate mean	$\bar{x} = \frac{a+2m+b}{4}$	Median

With:

m = Median

a = The smallest value (minimum)

b = The largest value (maximum)

n = The size of the sample.

To estimate the SD of the difference between baseline score and post-surgery score, we used this formula:

$$sd_t = \sqrt{sd_1^2 + sd_2^2}$$

With:

- sd_t : SD of the difference between standard deviation of the preoperative score and SD of the postoperative score.
- sd_1 : SD of the preoperative score.
- sd_2 : SD of the post-surgery score.

1. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol.* 2005;5:13.

Appendix 6. Pairwise comparisons; standard meta-analyses

HKA outliers +/- 3°: Pairwise

Treatment comparisons	Navigation vs PSI		Navigation vs standard		PSI vs standard		Robot vs standard		Accelerometer vs standard	
Trials included/total available trials	1 of 3		38 of 53		20 of 28		1 of 2		3 of 5	
Patient included in meta-analysis	60		5001		1776		60		353	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RR (95% CI)	1.00 (0.27 to 3.45)	1.00 (0.27 to 3.45)	0.43 (0.36 to 0.53)	0.50 (0.43 to 0.56)	1.00 (0.77 to 1.27)	1.04 (0.91 to 1.23)	0.14 (0.01 to 2.04)	0.14 (0.01 to 2.04)	0.56 (0.29 to 1.08)	0.59 (0.37 to 0.91)
I ²			51.2%	46.5%	37.0%	35.5%			0.0%	0.0%

PSI = patient-specific instrumentation.

Frontal femoral outliers +/-3°: Pairwise

Treatment comparisons	Navigation vs PSI		Navigation vs standard		PSI vs standard		Accelerometer vs standard	
Trials included/total available trials	2 of 3		19 of 53		15 of 28		4 of 5	
Patient included in meta-analysis	133		2312		1397		550	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis	1.89 (0.32 to 12.3)	1.64 (0.38 to 7.14)	0.43 (0.29 to 0.67)	0.48 (0.37 to 0.59)	0.83 (0.53 to 1.28)	0.91 (0.71 to 1.20)	0.18 (0.06 to 0.50)	0.18 (0.07 to 0.37)

RR (95% CI)								
I ²	57.7%	53.2%	47.1%	45.3%	35.8%	31,2%	0.0%	0.0%

PSI = patient-specific instrumentation.

Frontal tibial outliers +/-3°: Pairwise

Treatment comparisons	Navigation vs PSI		Navigation vs standard		PSI vs standard		Accelerometer vs standard	
Trials included/total available trials	2 of 3		23 of 53		15 of 28		3 of 5	
Patient included in meta-analysis	133		2698		1397		309	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RR (95% CI)	1.00 (0.15 to 6.67)	1.03 (0.22 to 5.26)	0.53 (0.33 to 0.77)	0.56 (0.42 to 0.71)	1.47 (0.91 to 2.56)	1.49 (1.09 to 2.04)	0.67 (0.16 to 2.63)	0.63 (0.21 to 1.69)
I ²	0.0%	0.0%	38.5%	38.7%	44.3%	44,0%	34.3%	33.2%

PSI = patient-specific instrumentation.

Sagittal frontal outliers +/-3°: Pairwise

Treatment comparisons	Navigation vs PSI	Navigation vs standard	PSI vs standard	Accelerometer vs standard
Trials included/total available trials	2 of 3	11 of 53	9 of 28	1 of 5
Patient included in meta-analysis	133	1336	841	179

Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RR (95% CI)	0.63 (0.11 to 3.45)	0.67 (0.13 to 3.23)	0.56 (0.38 to 0.77)	0.59 (0.48 to 0.71)	1.09 (0.77 to 1.59)	1.01 (0.83 to 1.19)	0.38 (0.02 to 7.14)	0.38 (0.02 to 7.14)
I ²	0.0%	0.0%	47.5%	47%	51.7%	48.6%		

PSI = patient-specific instrumentation.

Sagittal tibial outliers +/-3°: Pairwise

Treatment comparisons	Navigation vs PSI		Navigation vs standard		PSI vs standard		Accelerometer vs standard	
Trials included/total available trials	2 of 3		13 of 53		10 of 28		2 of 5	
Patient included in meta-analysis	133		1554		1069		259	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RR (95% CI)	0.37 (0.07 to 1.64)	0.34 (0.10 to 1.09)	0.56 (0.36 to 0.83)	0.59 (0.50 to 0.71)	1.67 (1.04 to 2.70)	1.47 (1.20 to 1.81)	0.59 (0.13 to 2.63)	0.50 (0.16 to 1.45)
I ²	0.0%	0.0%	66.4%	65.9%	69.5%	67.3%	72.9%	72.5%

PSI = patient-specific instrumentation.

Femoral rotation outliers +/-3°: Pairwise

Treatment comparisons	Navigation vs PSI	Navigation vs standard	PSI vs standard
Trials included/total available trials	1 of 3	7 of 53	8 of 28

Patient included in meta-analysis	73		774		686	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RR (95% CI)	1.12 (0.24 to 5.26)	1.12 (0.24 to 5.26)	1.18 (0.67 to 2.22)	0.53 (0.77 to 1.30)	0.71 (0.37 to 1.32)	0.77 (0.53 to 1.11)
I ²			62.6%	59.5%	32.3%	29.2%

PSI = patient-specific instrumentation.

Tibial rotation outliers +/-3°: Pairwise

Treatment comparisons	Navigation vs standard	
Trials included/total available trials	5 of 53	
Patient included in meta-analysis	651	
Fixed or random effect	Random effect	Fixed effect
Pairwise meta-analysis RR (95% CI)	0.83 (0.34 to 1.33)	0.91 (0.77 to 1.04)
I ²	49.0%	47.8%

WOMAC at the first 6 months post-surgery: Pairwise

Treatment comparisons	Navigation vs standard		PSI vs standard	
Trials included/total available trials	3 of 53		2 of 28	
Patient included in meta-analysis	388		265	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RR (95% CI)	-1.0 (-5.5 to 5.1)	-1.7 (-3.2 to 0.29)	4.5 (-2.7 to 12)	4.5 (-0.84 to 9.8)
I ²	33.2%	4.2%	0.0%	0.0%

PSI = patient-specific instrumentation.

WOMAC within 1 and 2 years post-surgery: Pairwise

Treatment comparisons	Navigation vs standard		PSI vs standard	
Trials included/total available trials	2 of 53		3 of 28	
Patient included in meta-analysis	145		360	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RR (95% CI)	-4.2 (-17 to 7.3)	-2.6 (-8.4 to 3.2)	-1.6 (-11 to 7.9)	-1.4 (-5.9 to 3.3)
I ²	61.7%	56.8%	71.2%	71.213%

PSI = patient-specific instrumentation.

KSS at the first 6 months post-surgery: Pairwise

Treatment comparisons	Navigation vs PSI		Navigation vs standard		PSI vs standard		Robot vs standard	
Trials included/total available trials	2 of 3		13 of 53		7 of 28		1 of 2	
Patient included in meta-analysis	120		1425		744		60	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RMD (95% CI)	-2.6 (-20.0 to 15.0)	-2.4 (-16.0 to 11.0)	3.2 (-1.9 to 8.5)	5.4 (4.2 to 6.7)	1.5 (-5.9 to 9.1)	0.79 (-3.3 to 4.9)	-8.5 (-26.0 to 9.0)	-8.5 (-26.0 to 9.0)
I ²	0.0%	0.0%	89.6%	88.3%	66.8%	66.6%		

PSI = patient-specific instrumentation.

KSS within 1 and 2 years post-surgery: Pairwise

Treatment comparisons	Navigation vs standard		PSI vs standard		Robot vs standard	
Trials included/total available trials	9 of 53		3 of 28		1 of 2	
Patient included in meta-analysis	863		300		60	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
Pairwise meta-analysis RMD (95% CI)	8.1 (-2.1 to 18)	11 (9.5 to 13)	1.9 (-16 to 20)	1.2 (-4.4 to 6.8)	-7.8 (-25.0 to 9.1)	-7.8 (-25.0 to 9.1)
I ²	97.0%	96.9%	0.00%	0.00%		

PSI = patient-specific instrumentation.

Surgical time: Pairwise

Treatment comparisons	Navigation vs PSI		Navigation vs standard		PSI vs standard		Robot vs standard		Accelerometer vs standard	
Trials included/total available trials	2 of 3		29 of 53		19 of 28		1 of 2		2 of 5	
Patient included in meta-analysis	120		3778		1624		60		259	
Fixed or random effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect

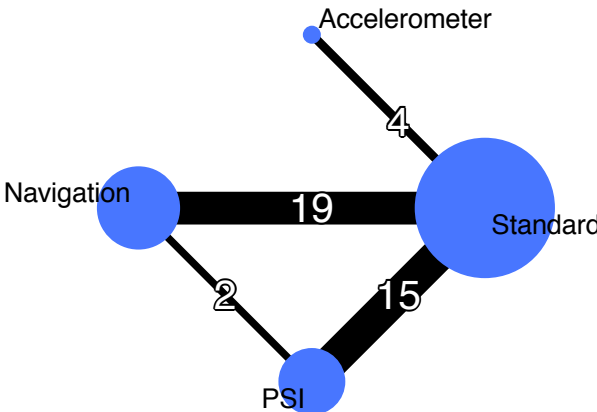
Pairwise meta-analysis RR (95% CI)	12 (-3.6 to 28.0)	7.8 (4.0 to 12.0)	18 (14.0 to 22.0)	18 (18.0 to 18.1)	-3.9 (-8.9 to 1.1)	-4.3 (-5.0 to -3.6)	2.0 (-8.5 to 4.5)	2.0 (-8.5 to 4.5)	2.6 (-13.0 to 18.0)	3.0 (0.0 to 6 to 5.9)
I ²	89.8%	79.7%	97.3%	97.2%	90.7%	90.6%			0.0%	0.0%

PSI = patient-specific instrumentation.

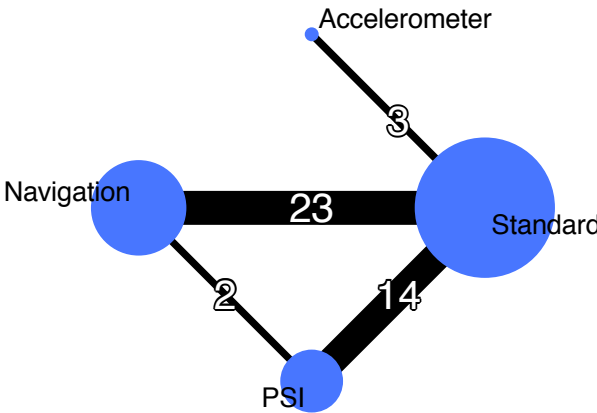
Appendix 7. Secondary outcomes: radiological outcomes

A. Network diagrams

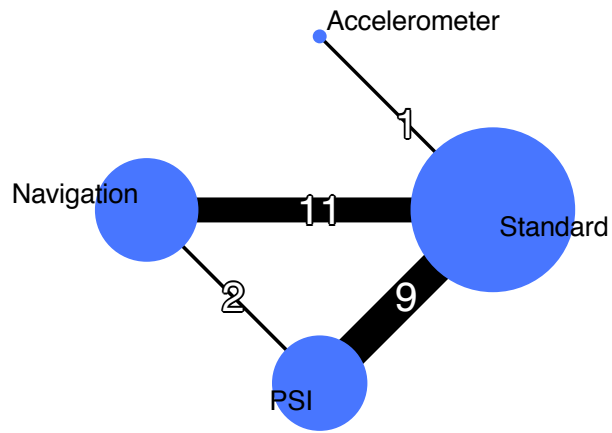
Frontal femoral misalignment network



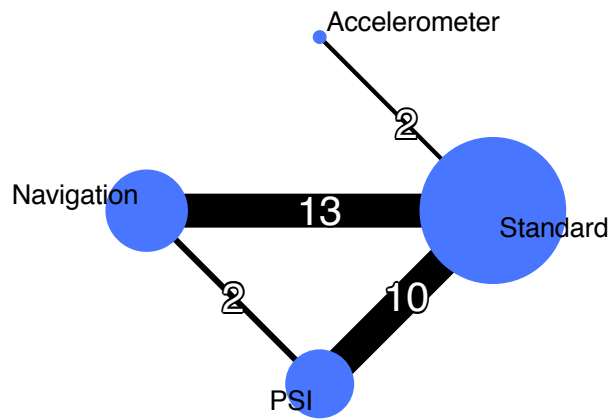
Frontal tibial misalignment network



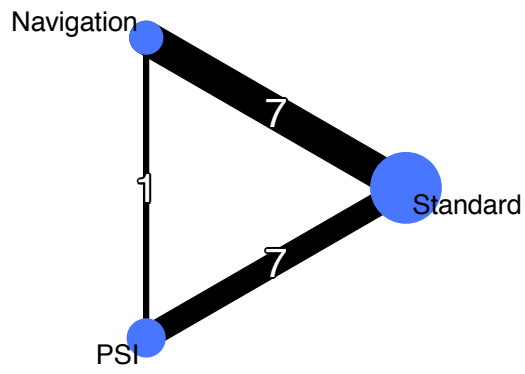
Sagittal femoral misalignment network



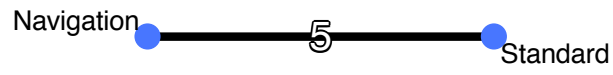
Sagittal tibial misalignment network



Femoral rotation misalignment network

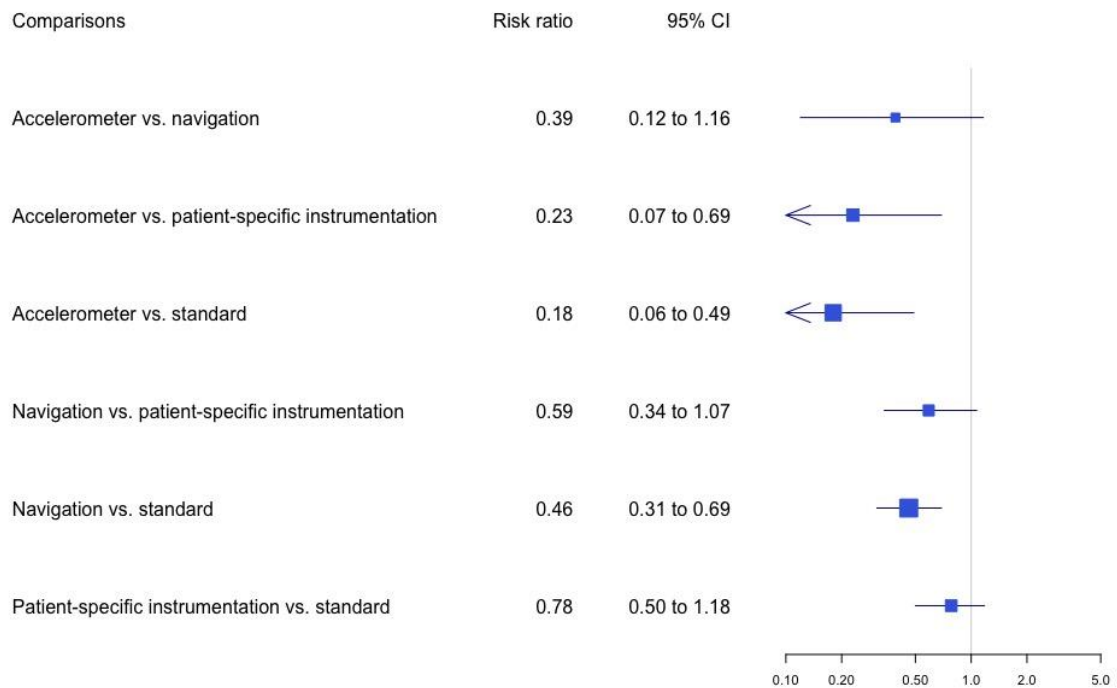


Tibial rotation misalignment network

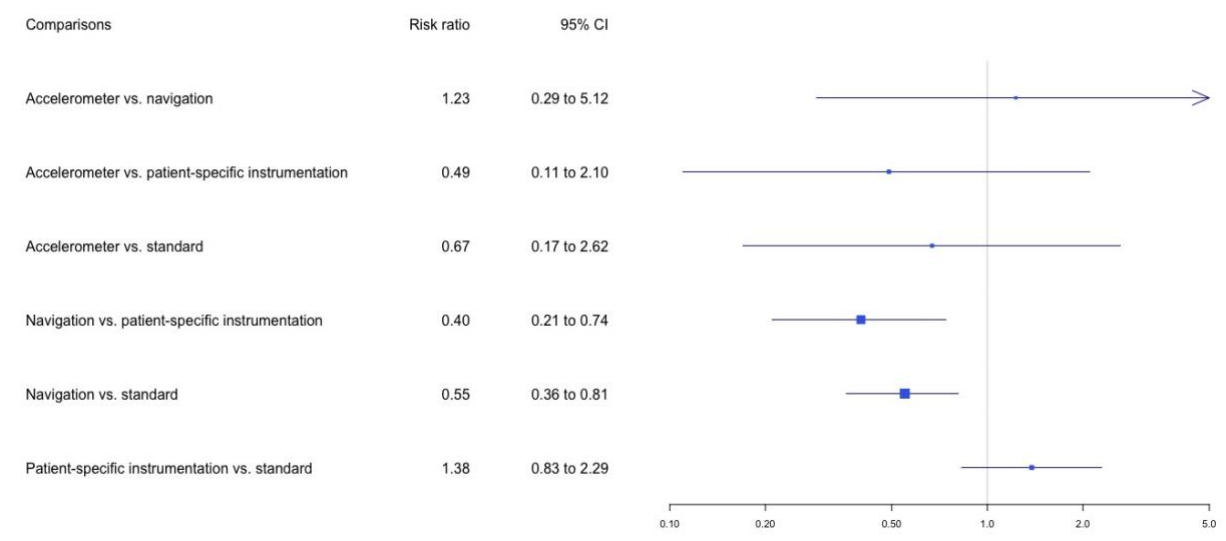


B. Relative effects

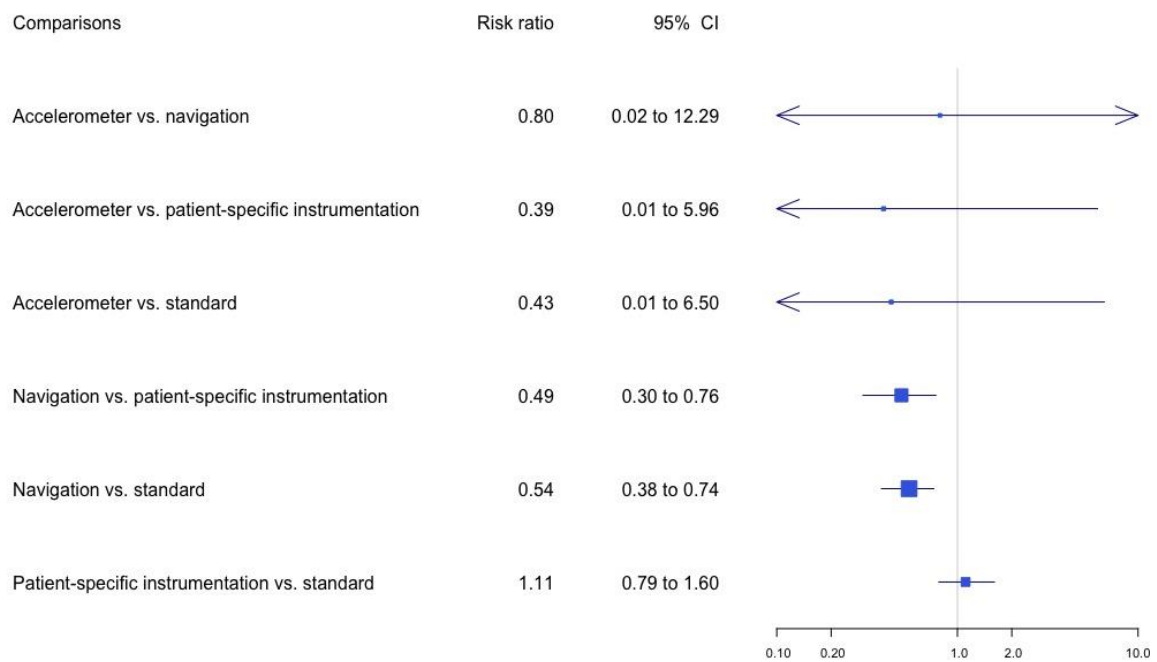
Forest plots in the network meta-analysis of the relative risks of frontal femoral misalignments between the different cutting guides



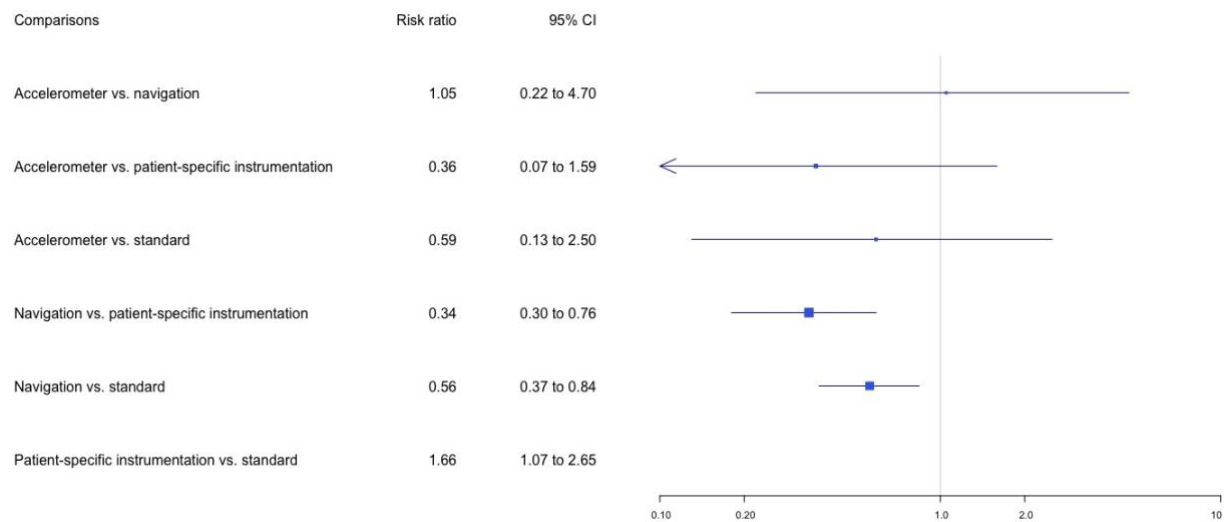
Forest plots in the network meta-analysis of the relative risks of frontal tibial misalignments between the different cutting guides



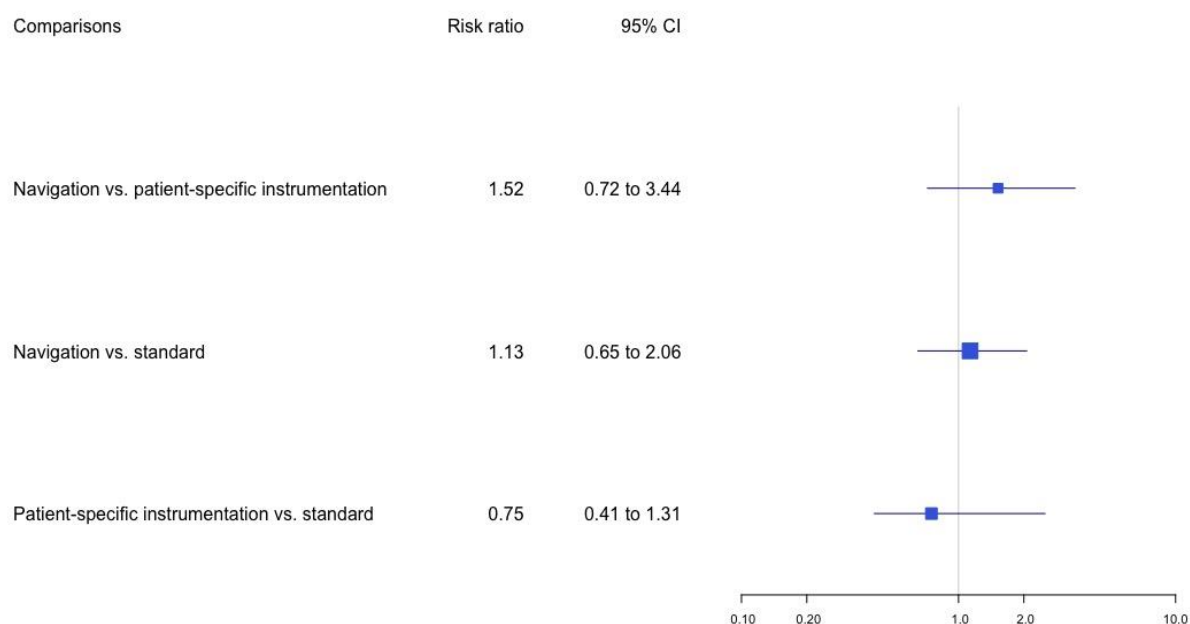
Forest plots in the network meta-analysis of the relative risks of sagittal femoral misalignments between the different cutting guides



Forest plots in the network meta-analysis of the relative risks of sagittal tibial misalignments between the different cutting guides



Forest plots in the network meta-analysis of the relative risks of axial femoral misalignments between the different cutting guides

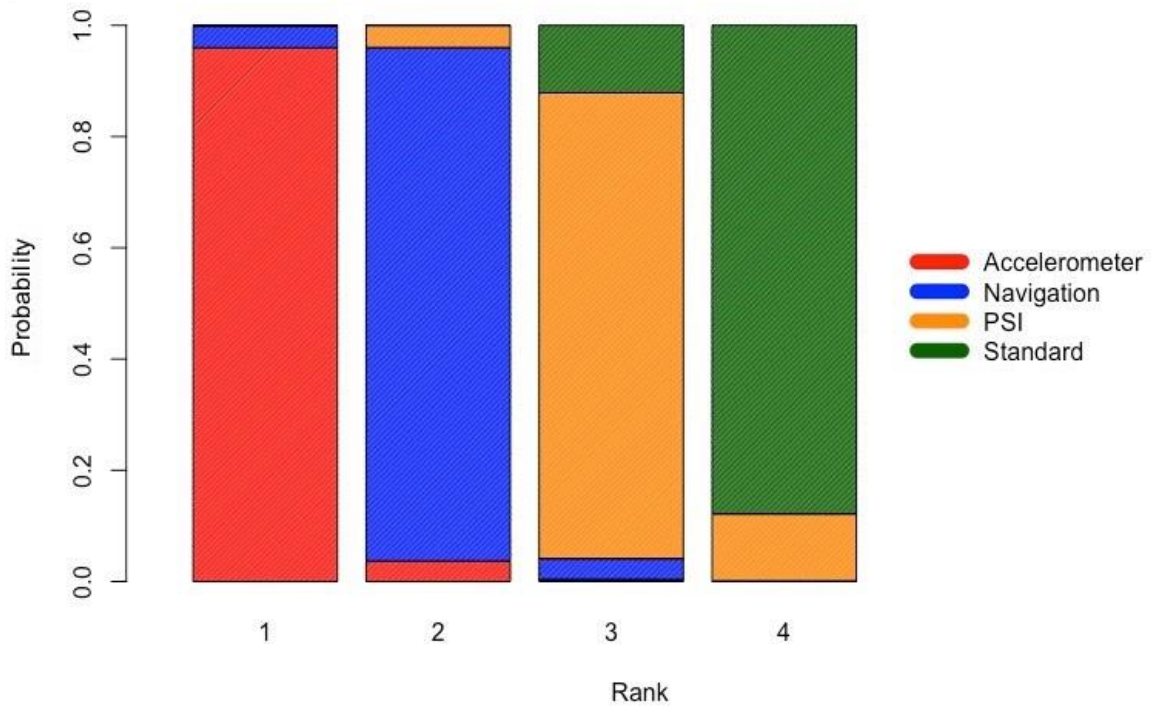


C. Ranking

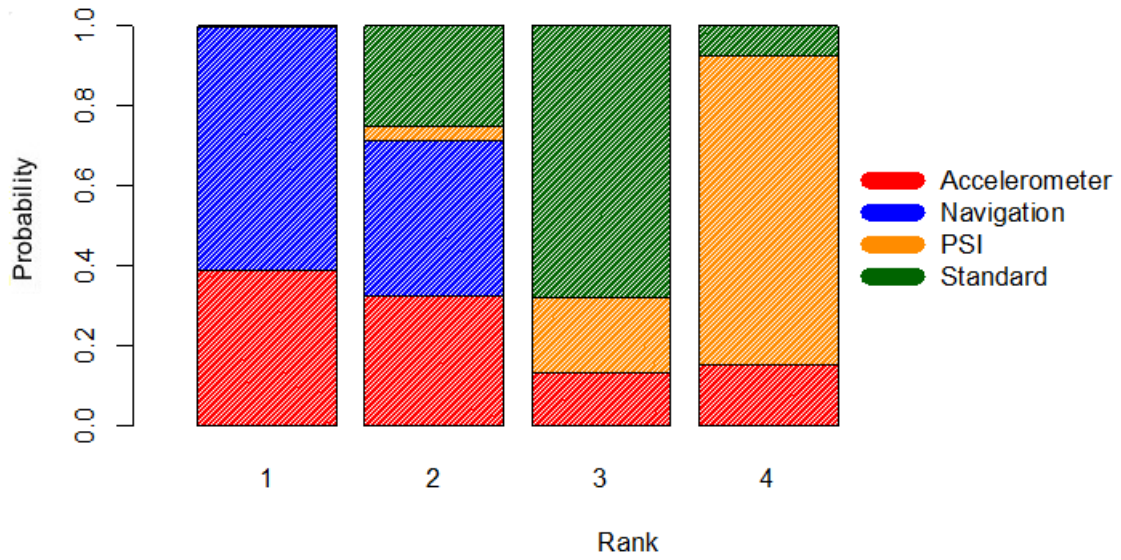
Mean rank

These images show rank probabilities of secondary radiological outcomes for each cutting guide. The x axis represents the ranking. The y axis represents the probability of rank: 0% to 100%. Each color represents one cutting guide: red for accelerometer-based navigation, blue for navigation, yellow for patient-specific instrumentation, purple for robot navigation, and green for the standard cutting guide; PSI = patient-specific instrumentation.

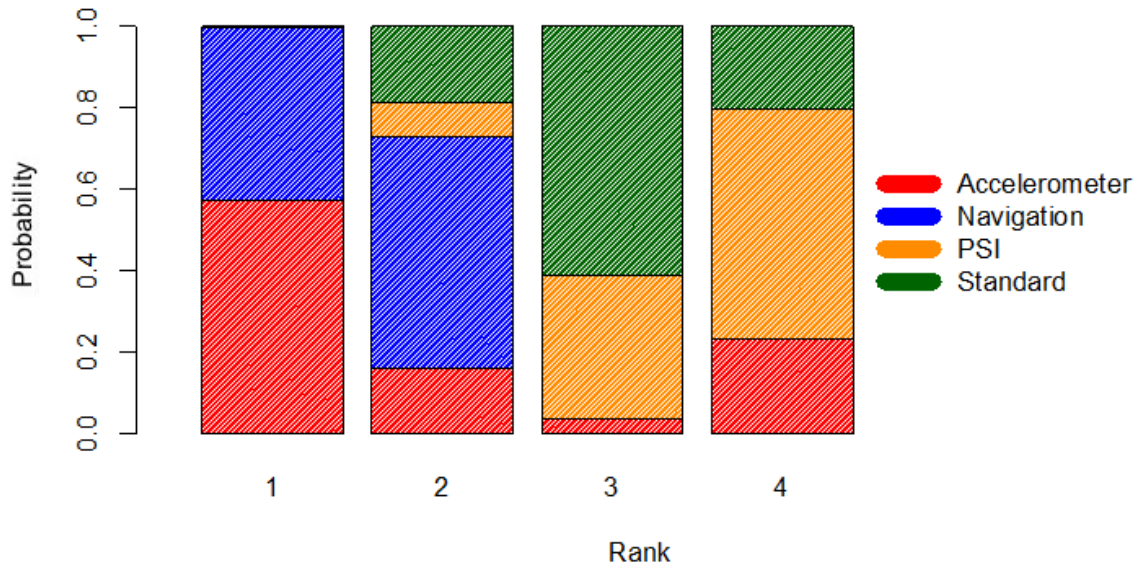
Ranking for frontal femoral misalignments



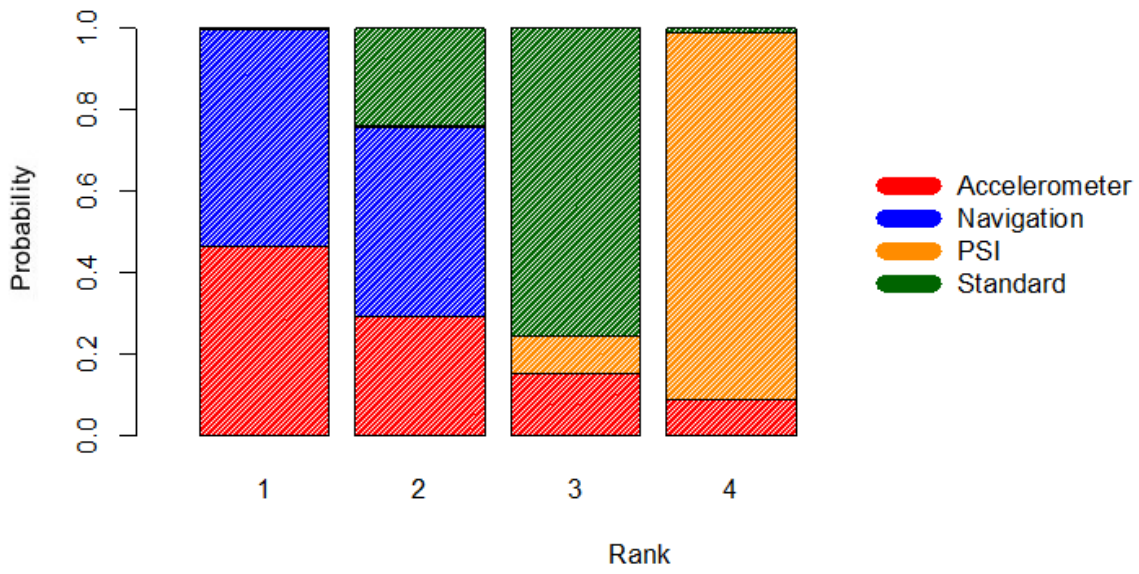
Ranking for frontal tibial misalignments

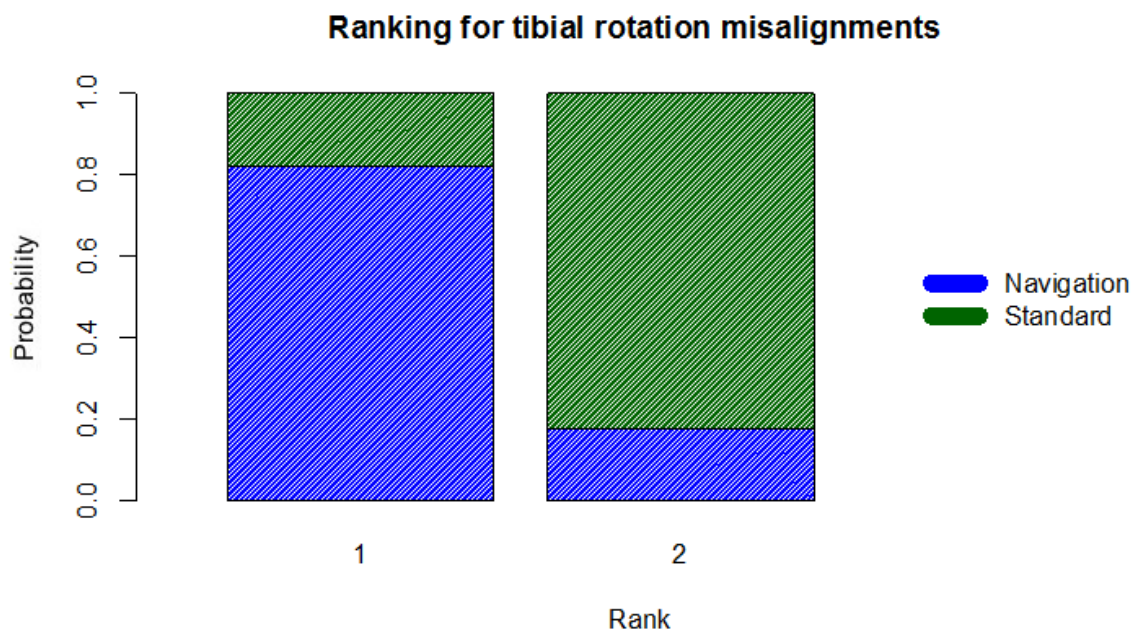
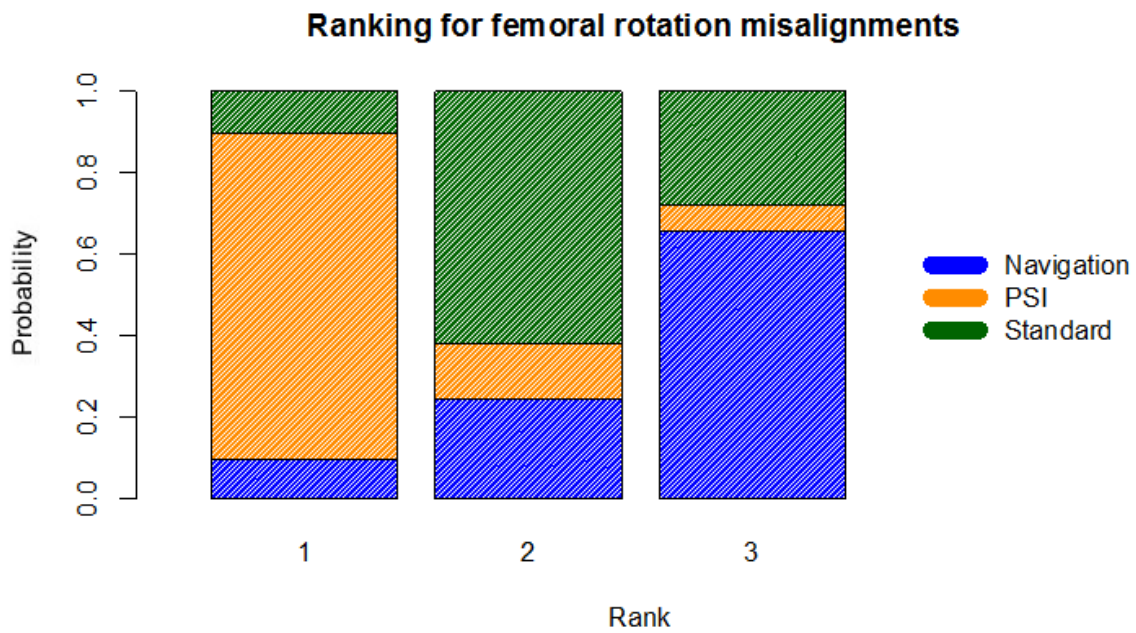


Ranking for femoral sagittal misalignments



Ranking for tibial sagittal misalignments

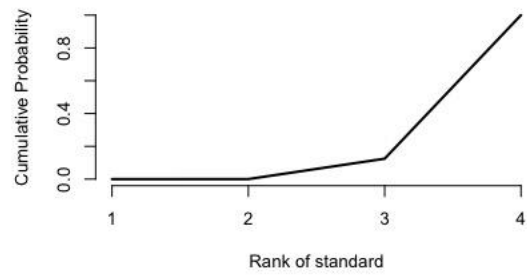
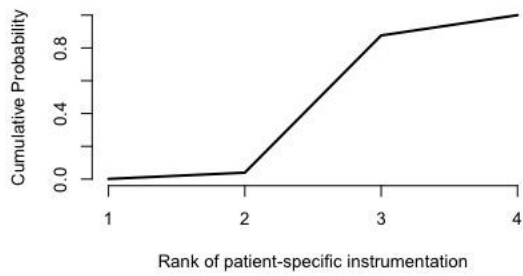
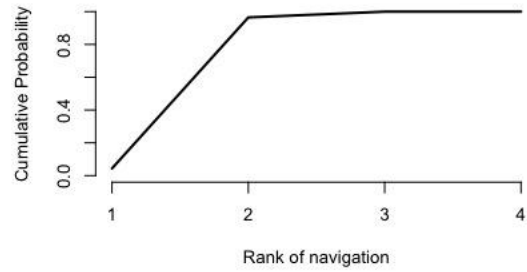
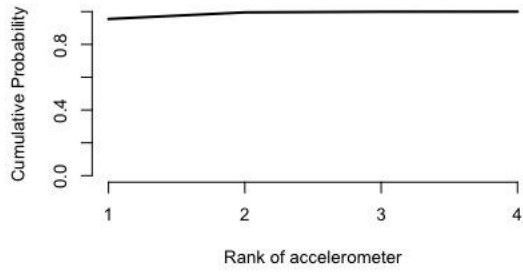




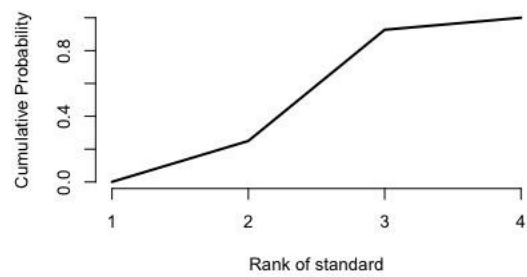
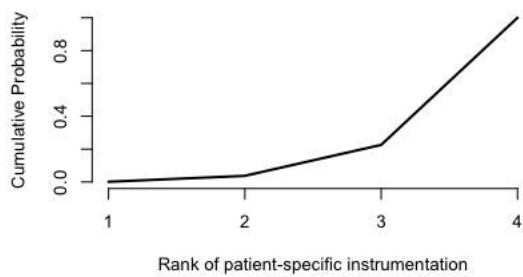
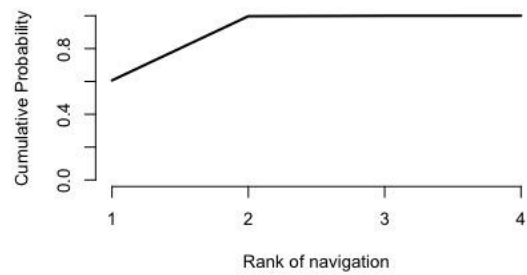
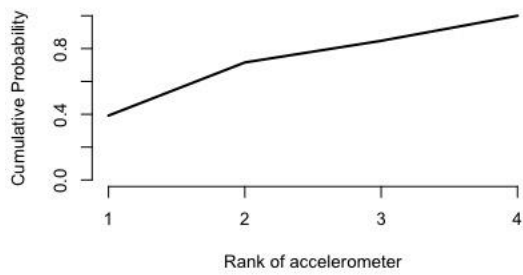
The Surface Under the Cumulative Ranking Curve (SUCRA)

These images show surface under the cumulative ranking curve SUCRA of each cutting guide for each secondary radiological outcomes. The x axis represents the ranking. The y axis represents the probability of rank: from 0% to 100%. The higher the SUCRA value, and the closer to 100%, the higher the likelihood that a therapy is in the top rank or one of the top ranks; the closer to 0 the SUCRA value, the more likely that a therapy is in the bottom rank, or one of the bottom ranks; PSI = patient-specific instrumentation.

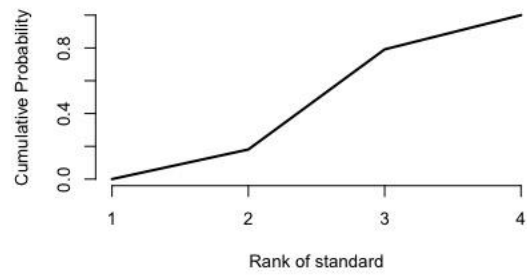
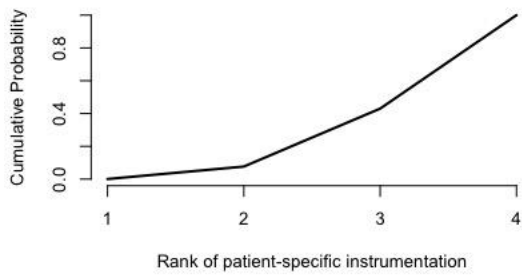
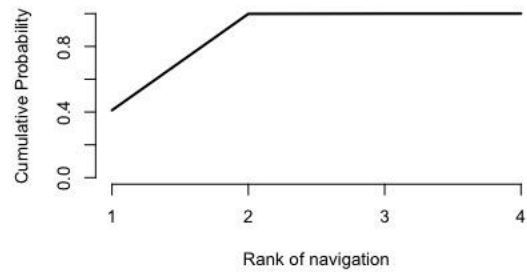
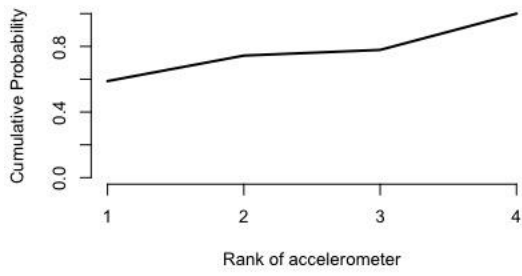
Frontal femoral misalignment



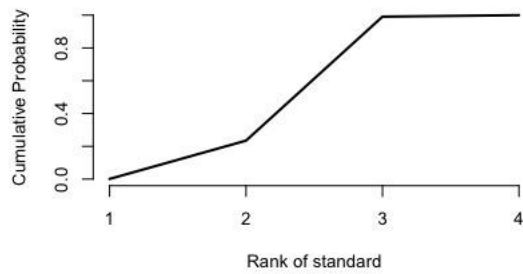
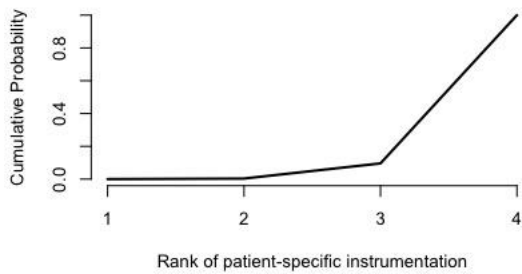
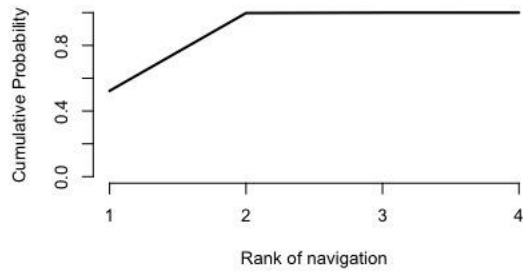
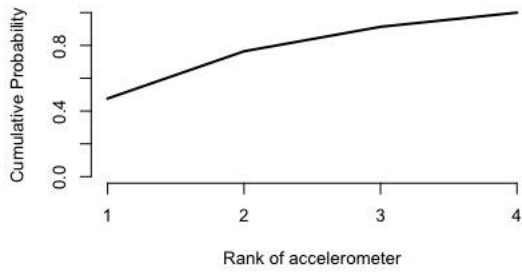
Frontal tibial misalignment



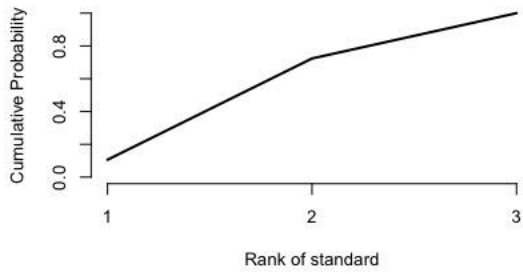
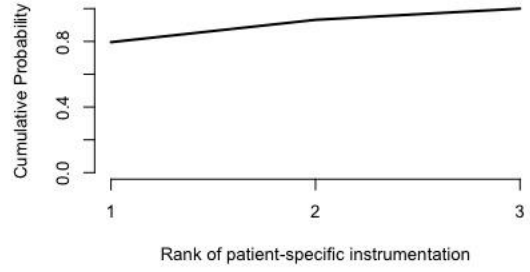
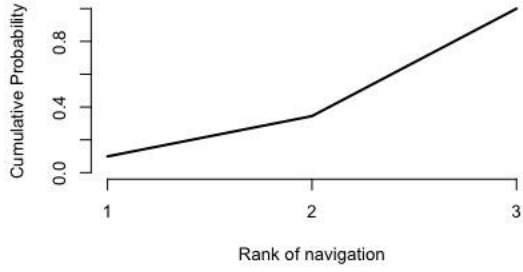
Sagittal femoral misalignment



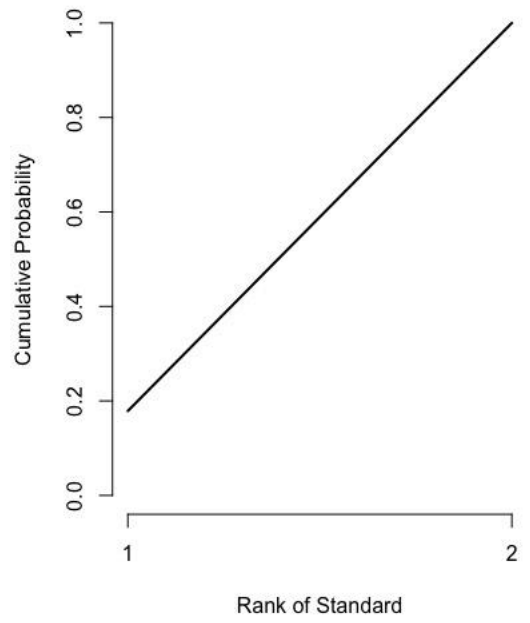
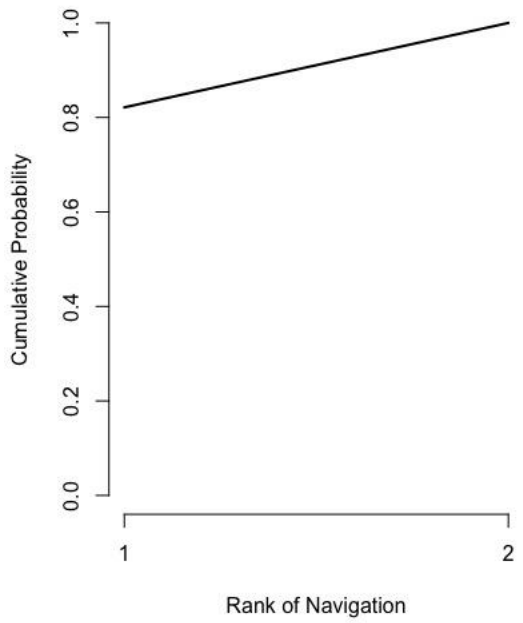
Sagittal tibial misalignment



Femoral rotation misalignment



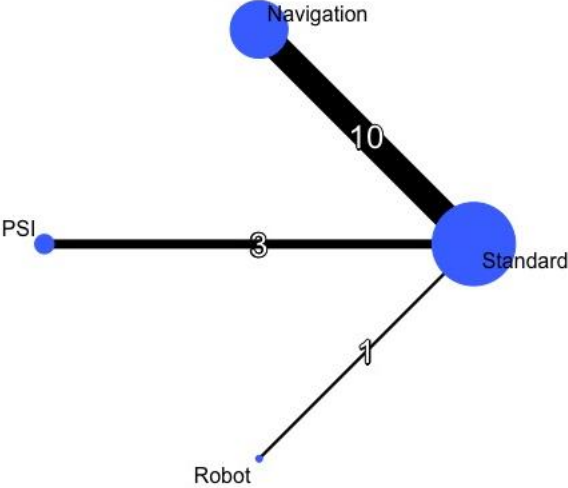
Tibial rotation misalignment



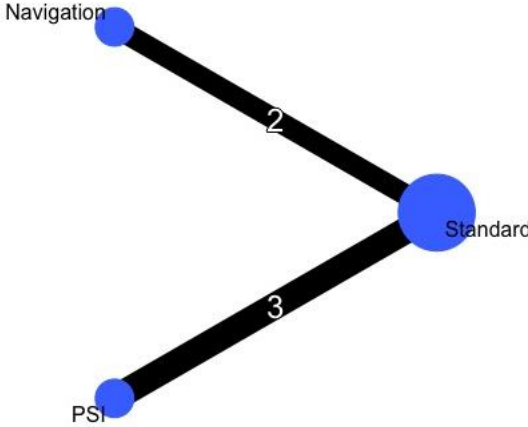
Appendix 8. Secondary outcomes: Clinicals and procedure relative outcome

Network diagrams

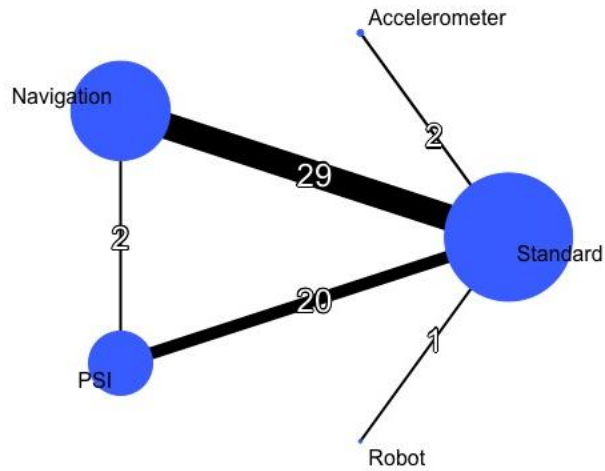
KSS within 1 and 2 years post-surgery network



WOMAC within 1 and 2 years post-surgery network

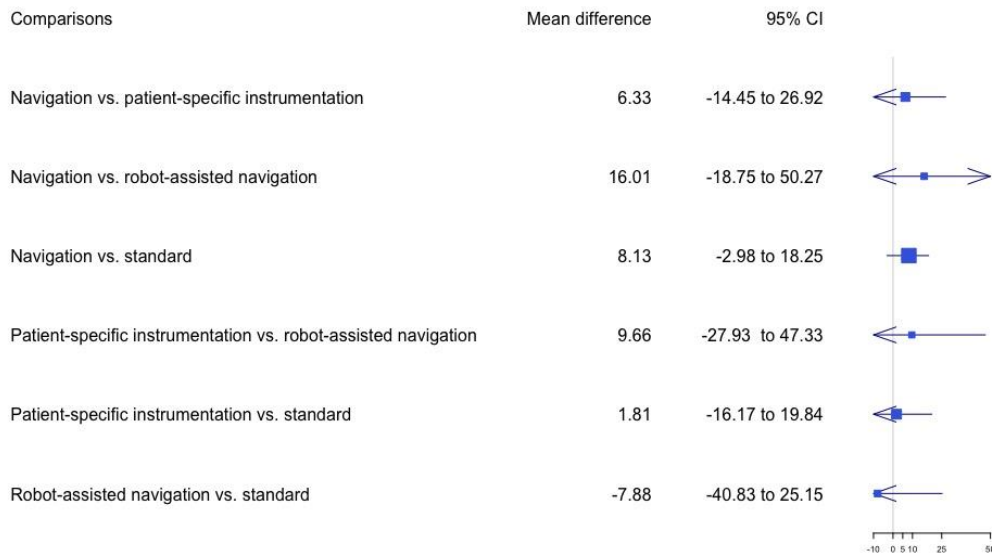


Surgical time network

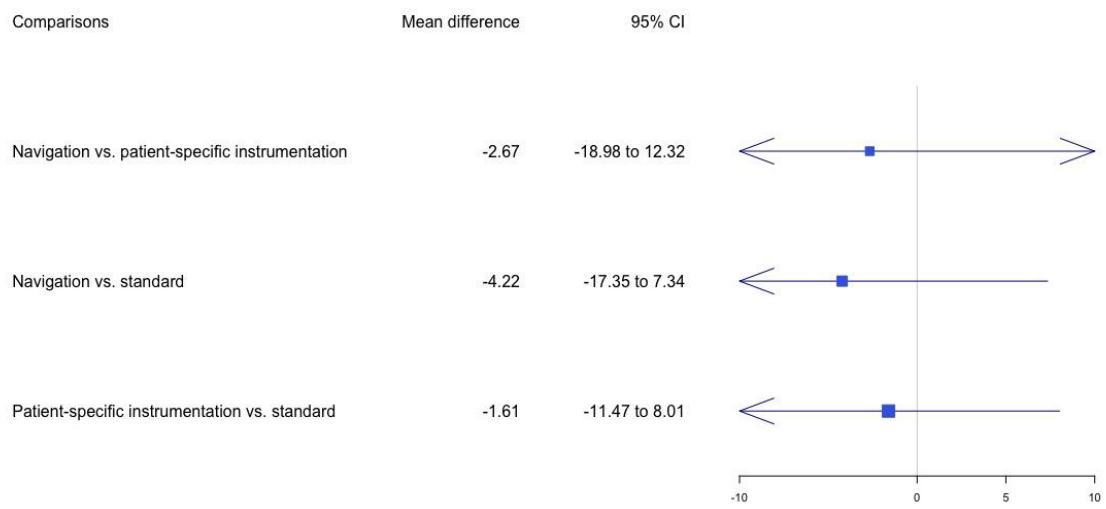


Relative Effects

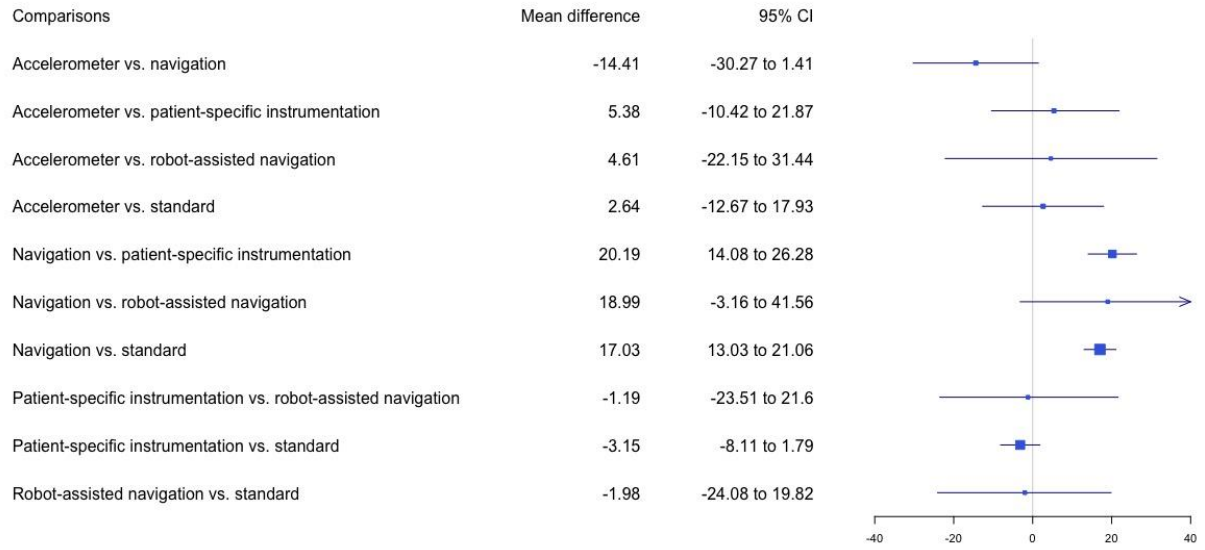
Forest plots in the network meta-analysis of the mean difference of KSS within 1 and 2 years postoperatively



Forest plots in the network meta-analysis of the mean difference of WOMAC within 1 and 2 years postoperatively



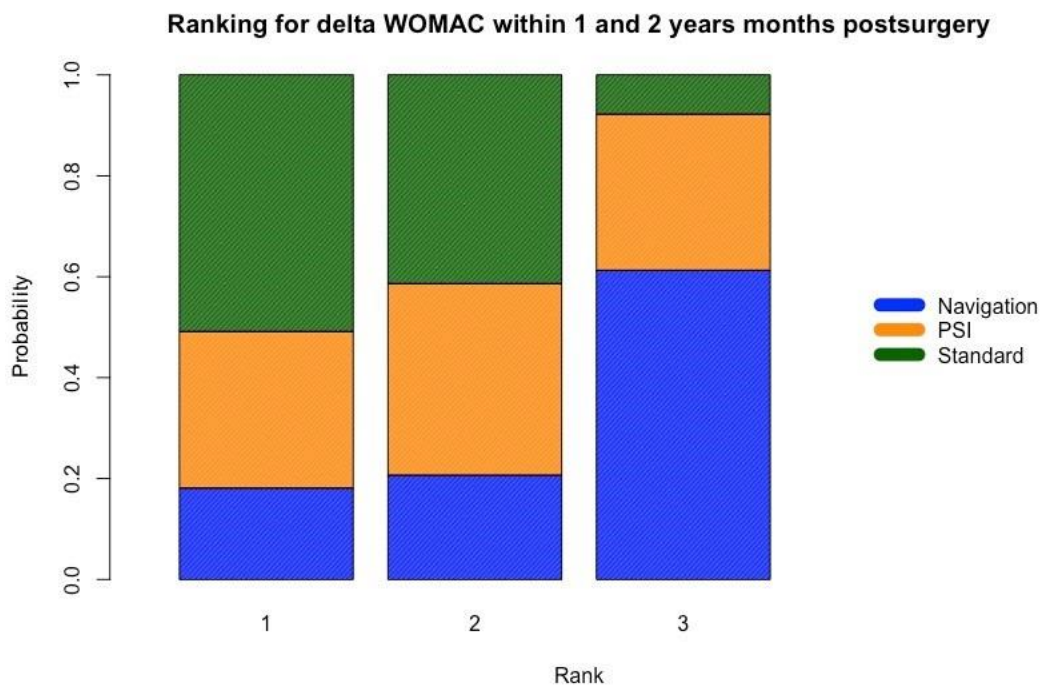
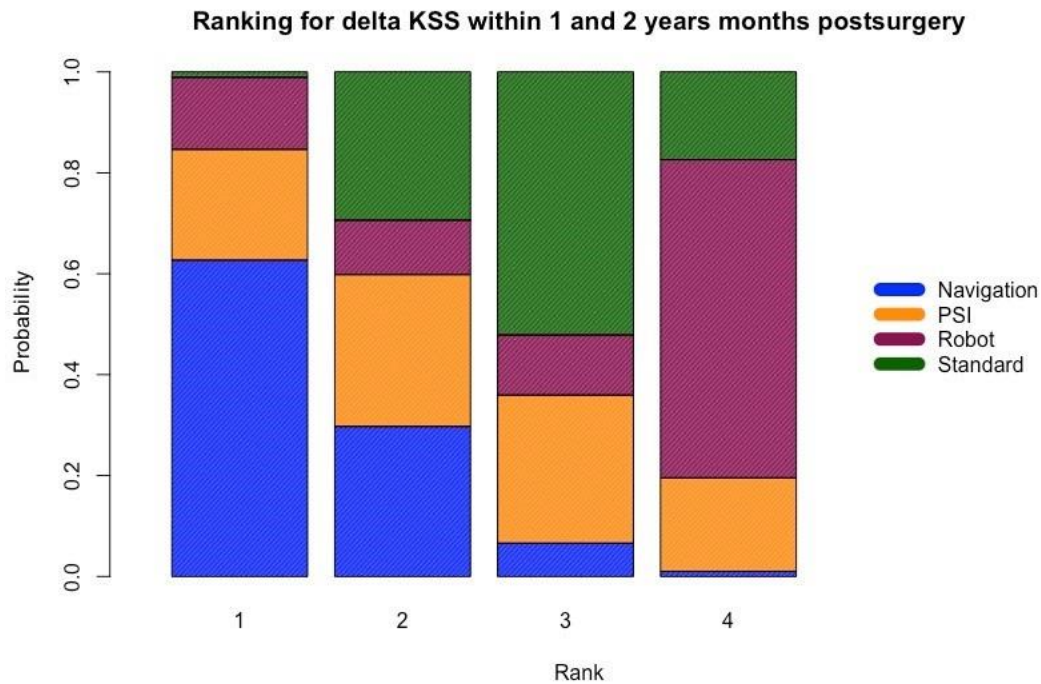
Forest plots in the network meta-analysis of the mean difference of surgical time

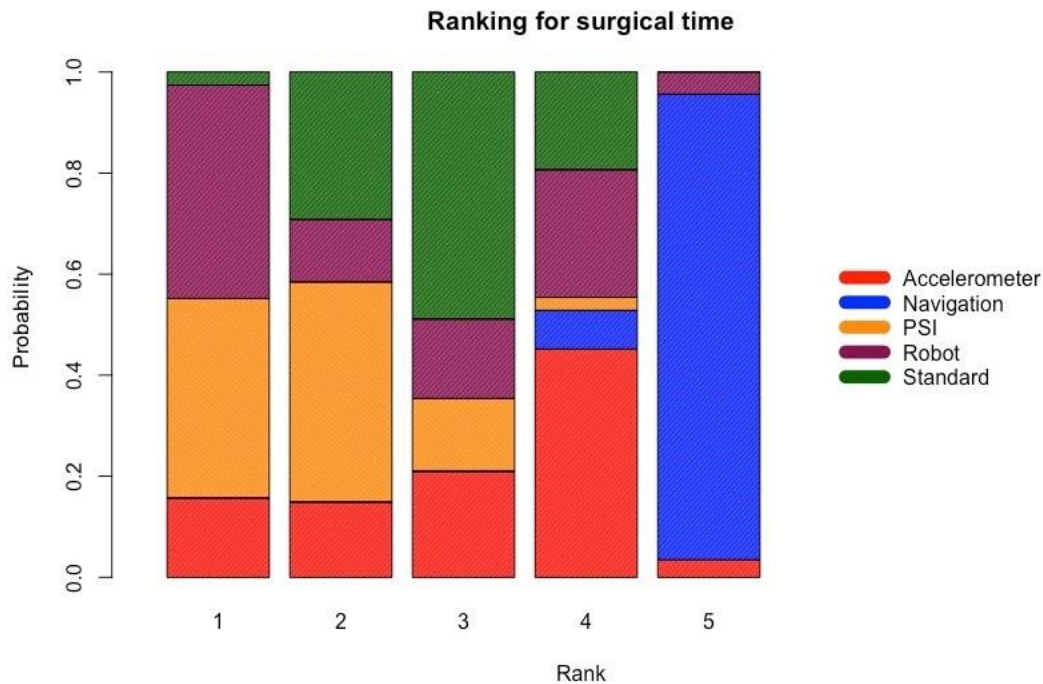


Ranking

Mean rank

These images show rank probabilities of secondary clinical outcomes for each cutting guide. The x axis represents the ranking. The y axis represents the probability of rank: 0% to 100%. Each color represents one cutting guide: red for accelerometer-based navigation, blue for navigation, yellow for patient-specific instrumentation, purple for robot navigation, and green for the standard cutting guide; PSI = patient-specific instrumentation.

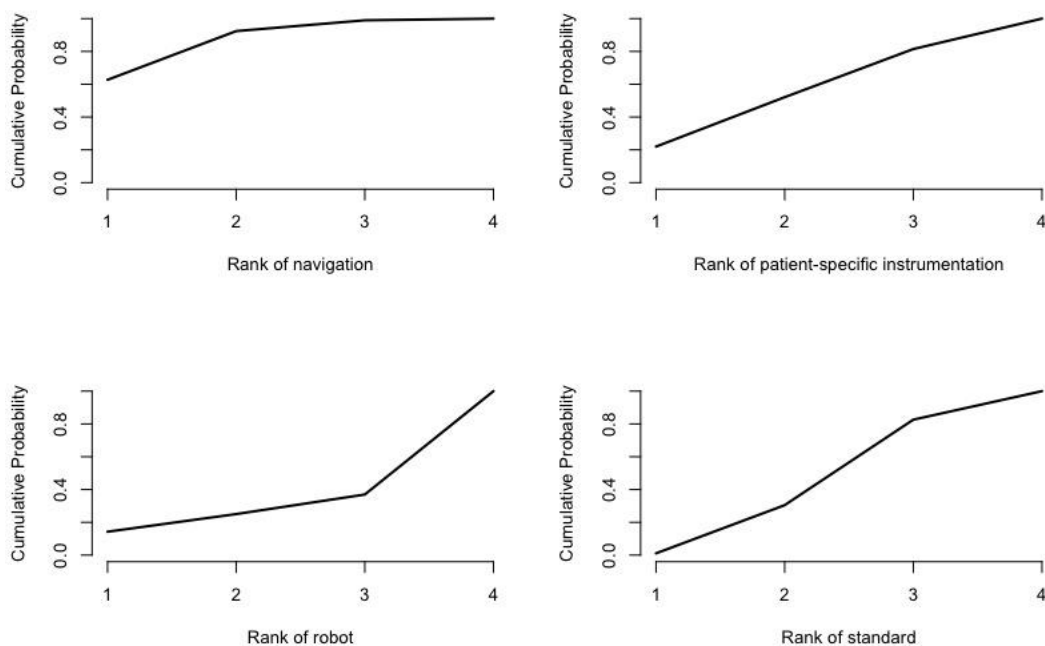




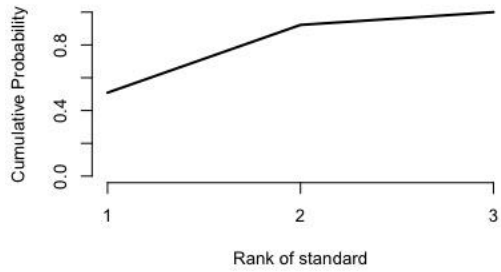
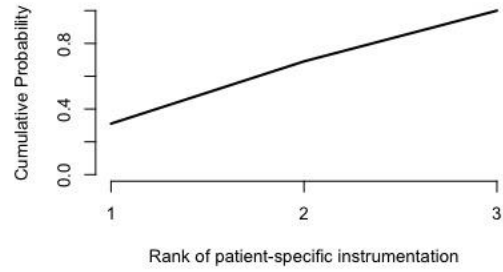
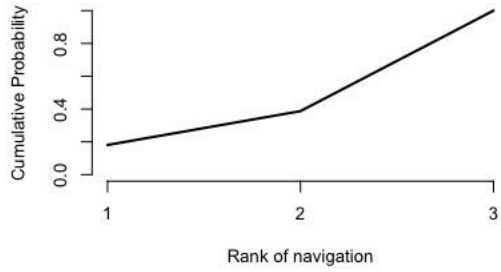
The Surface Under the Cumulative Ranking Curve (SUCRA)

These images show surface under the cumulative ranking curve SUCRA of each cutting guide for each secondary clinical outcomes. The x axis represents the ranking. The y axis represents the probability of rank: from 0% to 100%. The higher the SUCRA value, and the closer to 100%, the higher the likelihood that a therapy is in the top rank or one of the top ranks; the closer to 0 the SUCRA value, the more likely that a therapy is in the bottom rank, or one of the bottom ranks; PSI = patient-specific instrumentation.

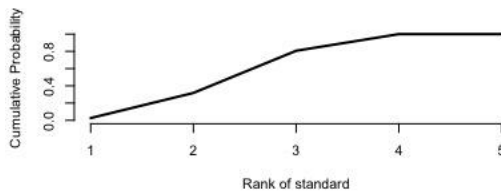
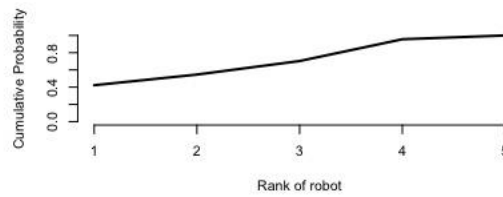
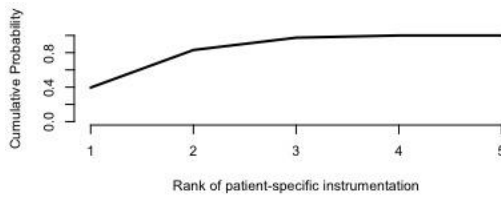
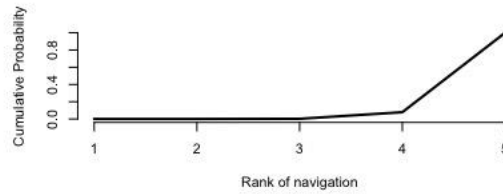
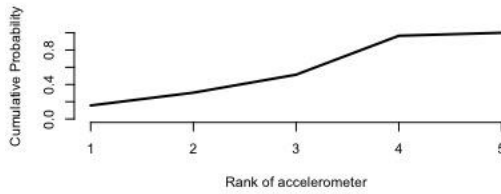
Δ KSS within 1 and 2 years postoperatively



Δ WOMAC within 1 and 2 years



Surgical time



Appendix 9. Complications

Types of complications	PSI	Standard	Navigation	Robot	Accelerometer
Total	9% (34 of 393)	4.7% (125 of 2659)	4.6% (101 of 2187)	16.1% (5 of 31)	2.1% (3 of 146)
Infections	2.5% (10 of 393)	0.8% (21 of 2527)	0.6% (12 of 2048)	6.5% (2 of 31)	0.7% (1 of 146)
Others	1.5% (6 of 393)	0.7% (18 of 2527)	0.3% (6 of 2048)	3.2% (1 of 31)	0% (0 of 146)
Mechanical	0.5% (2 of 393)	0.6% (14 of 2527)	2.3% (47 of 2048)	0% (0 of 31)	0% (0 of 146)
Skin	0% (0 of 393)	0.2% (5 of 2527)	0.3% (6 of 2048)	0% (0 of 31)	0.7% (1 of 146)
Stiffness	2.5% (10 of 393)	1.2% (30 of 2527)	0.4% (8 of 2048)	0% (0 of 31)	0% (0 of 146)
Vascular	1.3% (5 of 393)	0.9% (24 of 2527)	0.6% (12 of 2048)	6.5% (2 of 31)	0% (0 of 146)
Fractures	0% (0 of 393)	0.1% (3 of 2527)	0.2% (4 of 2048)	0% (0 of 31)	0.7% (1 of 146)

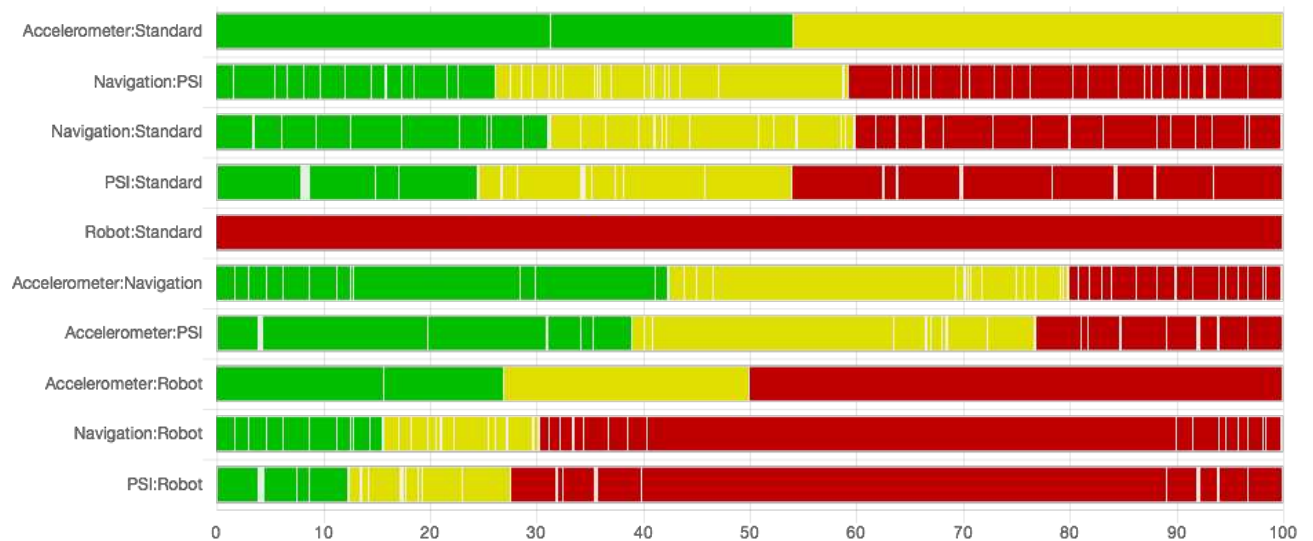
PSI = patient-specific instrumentation.

Appendix 10. Grade evaluation

The quality of evidence of the network meta-analysis has been assessed using the GRADE framework [1]. The online tool CiNeMA [2] was used to evaluate study limitations. It combines judgments about direct comparison with their contribution in the network meta-analysis results. Below, our GRADE evaluation for HKA outliers $\pm 3^\circ$, KSS at 6 months postoperative and WOMAC at 6 months postoperative are presented; PSI = patient-specific instrumentation.

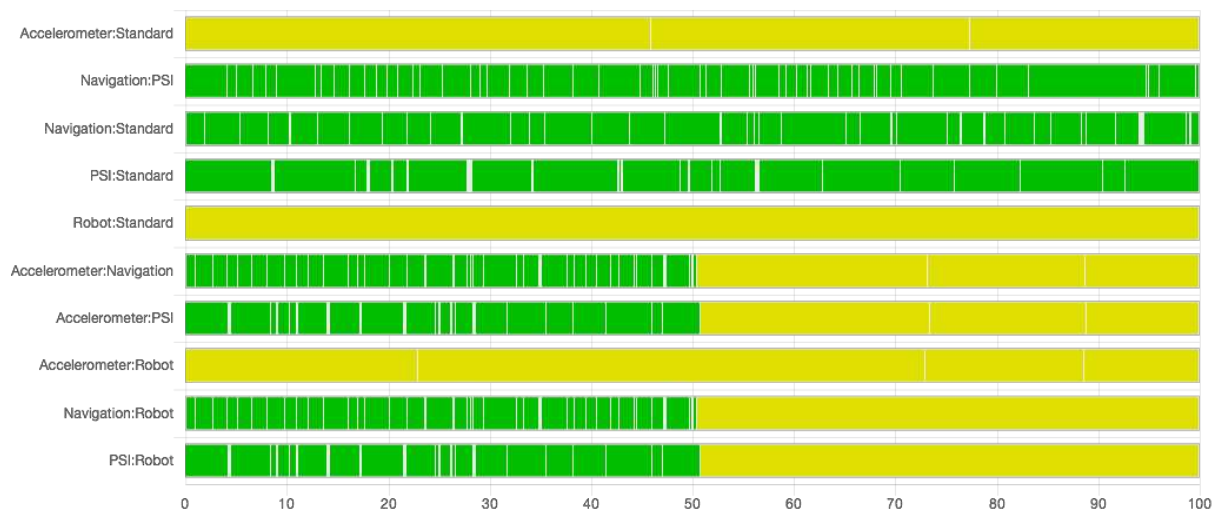
Grade evaluation for HKA outliers $\pm 3^\circ$

Study limitations



The bar-graph shows for the network estimate of every pairwise comparison how much information comes from low (green), moderate (yellow) or high (red) risk of bias trials.

Indirectness

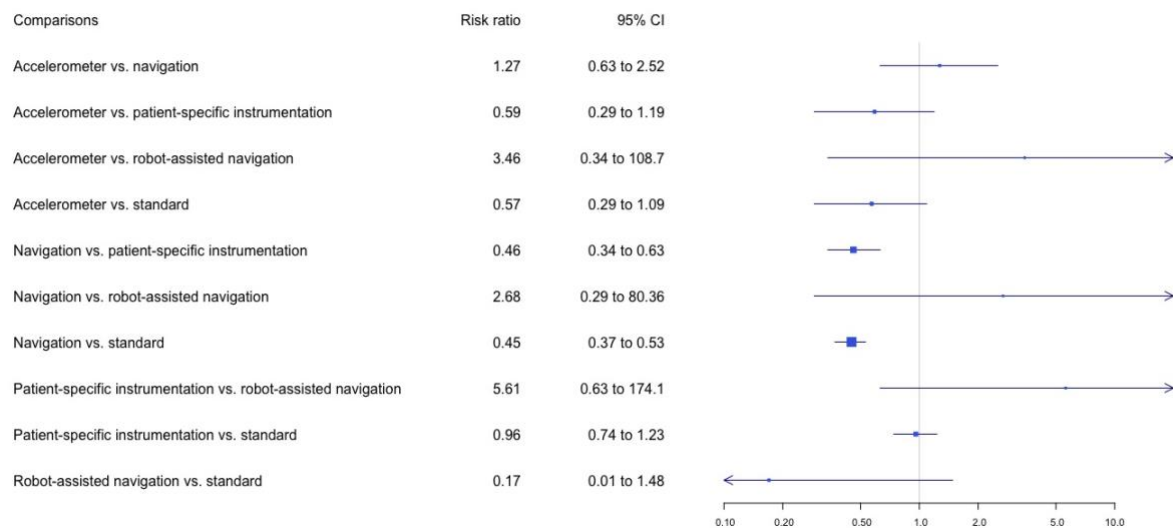


1. Salanti G, Giovane CD, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. *PLoS One*. 2014;9:e99682.

2. Confidence in Network Meta-Analysis (software). Institute of Social and Preventive Medicine, University of Bern. Available at: cinema.ispm.unibe.ch. Accessed June 7, 2018.

Imprecision

We had some concerns for accelerometer/standard because the credible intervals were large, and we had major concerns for robot/standard because there was only one study and the credible intervals was very large.



Heterogeneity/inconsistency

We had some concerns for heterogeneity of navigation/standard comparison ($I^2 = 51\%$). We did not find inconsistency for this network.

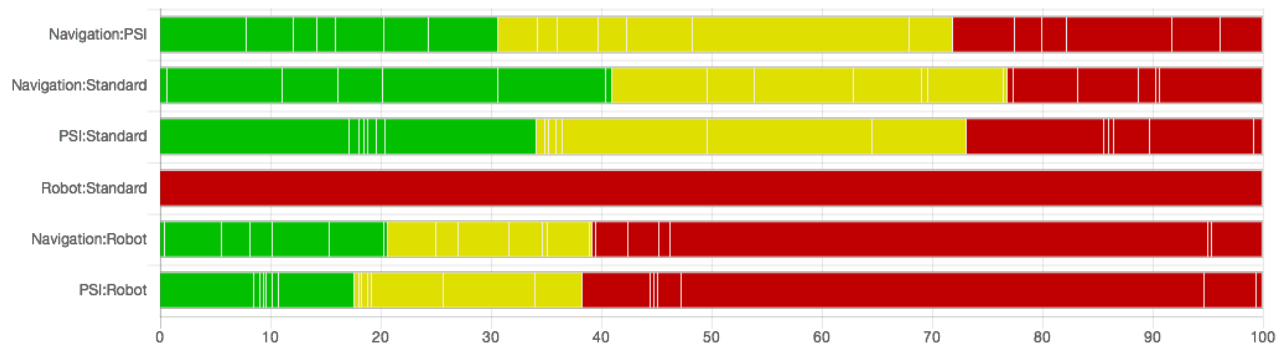
Publication Bias

We searched in ClinicalTrials.gov and in The World Health Organization International Clinical Trials Registry Platform (ICTRP) to locate potential unpublished (or gray) literature. So, we do not have serious concerns about the presence of publications bias.

Final Assessments

Grade evaluation for KSS within first 6 months postoperative

Study limitations



The bar-graph shows for the network estimate of every pairwise comparison how much information comes from low (green), moderate (yellow) or high (red) risk of bias trials; PSI = patient-specific instrumentation.

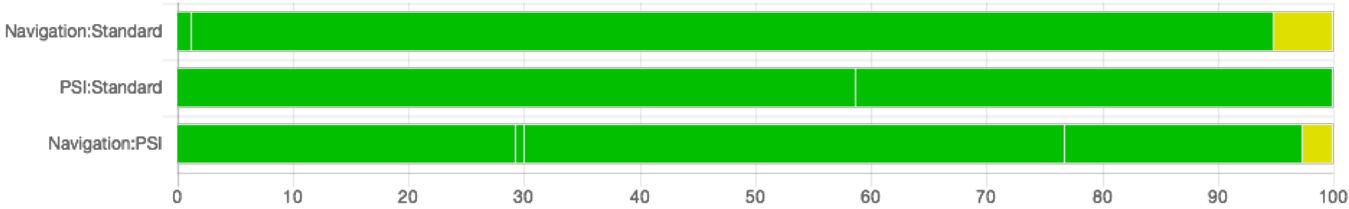
Indirectness

Comparison	Number of studies	Study limitations	Imprecision	Heterogeneity	Incoherence	Indirectness	Publication bias	Confidence rating
Accelerometer: standard	3	No concerns	Some concerns	No concerns	No concerns	Major concerns	Undetected	Very low
Navigation: PSI	1	Some concerns	No concerns	No concerns	No concerns	No concerns	Undetected	Moderate
Navigation: standard	38	Some concerns	No concerns	Some concerns	No concerns	No concerns	Undetected	Moderate
PSI: standard	20	Some concerns	No concerns	No concerns	No concerns	No concerns	Undetected	Moderate
Robot: standard	1	Major concerns	Major concerns	No concerns	No concerns	Major concerns	Undetected	Very low

Navigati on: PSI	2	Some concern s	Some concerns	No concerns	No concerns	No concerns	Undetect ed	Moderat e
Navigati on: standard	13	Some concern s	Some concerns	Major concerns	No concerns	No concerns	Undetect ed	Very Low
PSI: standard	7	Some concern s	Some concerns	Some concerns	No concerns	No concerns	Undetect ed	Low
Robot: standard	1	Major concern s	Major concerns	No concerns	No concerns	Some concerns	Undetect ed	Very low

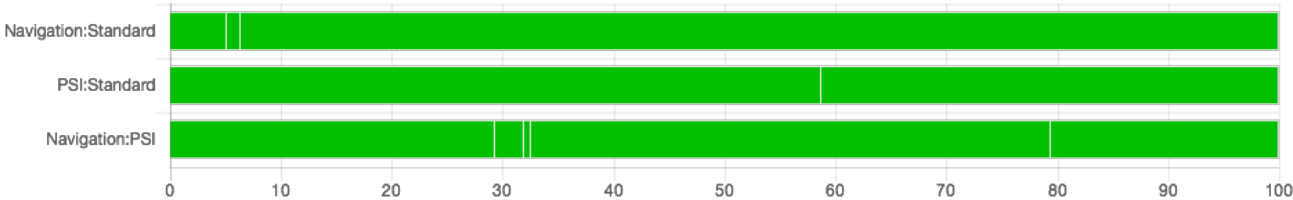
Grade evaluation for WOMAC at the first 6 months post-surgery

Study limitations



The bar-graph shows for the network estimate of every pairwise comparison how much information comes from low (green), moderate (yellow) or high (red) risk of bias trials. We had no concerns for all comparisons because most studies of each comparisons were in the no concerns category; PSI = patient-specific instrumentation.

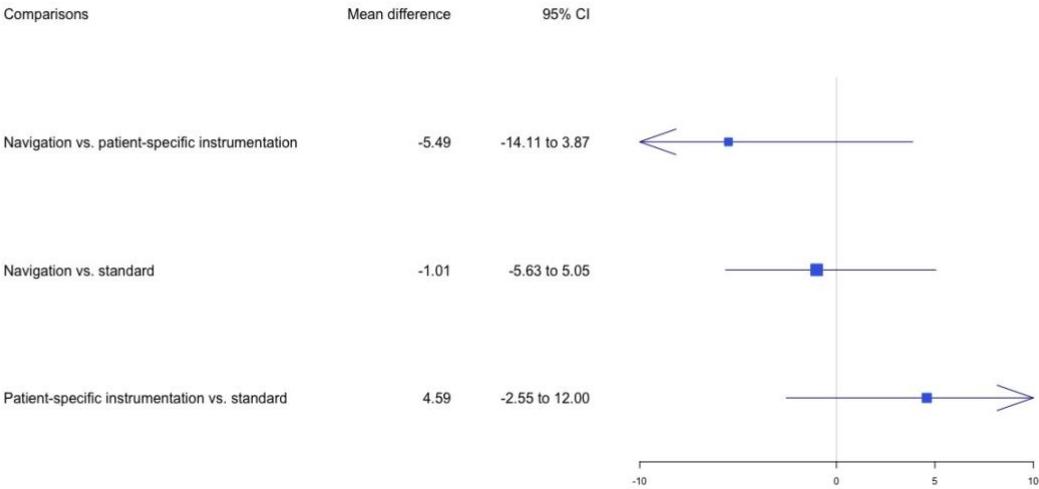
Indirectness



We had no concerns because the characteristics of the population of each comparison were overall similar.

Imprecision

We had some concerns about all comparisons because the credible intervals were large.



Heterogeneity/inconsistency

We did not find heterogeneity and inconsistency for this network.

Publication bias

We did not find small study effects, and we searched in ClinicalTrials.gov and in The World Health Organization International Clinical Trials Registry Platform (ICTRP) to locate potential unpublished (or gray) literature. So, we do not have serious concerns about the presence of publications bias.

Final assessments

Comparison	Number of studies	Study limitations	Imprecision	Heterogeneity	Incoherence	Indirectness	Publication bias	Confidence rating
Navigational: standard	3	No concerns	Some concerns	No concerns	No concerns	No concerns	Undetected	Moderate
PSI: standard	2	No concerns	Some concerns	No concerns	No concerns	No concerns	Undetected	Moderate

Study	Author	Date of publication	Country	Sub-speciality
Minimally invasive lumbar decompression-the surgical learning curve.	Ahn	2016	USA	Spine
Learning Curve for Open Reduction and Internal Fixation of Displaced Intra-articular Calcaneal Fracture by Extensile Lateral Approach Using the Cumulative Summation Control Chart.	Ahn	2019	Korea	Trauma
The learning curve of proximal femoral nailing.	Altintas	2014	Germany	Trauma
Validating a global rating scale to monitor individual resident learning curves during arthroscopic knee meniscal repair.	Alvand	2013	UK	Arthroscopy
Percutaneous Endoscopic Lumbar Discectomy Assisted by O-Arm-Based Navigation Improves the Learning Curve.	Ao	2019	China	Spine
Learning curve for the two-incision total hip replacement.	Archibeck	2004	USA	prosthesis
Learning curve analysis of 3D-fluoroscopy image-guided pedicle screw insertions in lumbar single-level fusion procedures.	Balling	2018	Germany	Spine
Comparative learning curves for early complications in anatomical and reverse shoulder arthroplasty.	Beazley	2018	UK	prosthesis
Hueter anterior approach for hip resurfacing: assessment of the learning curve.	Benoit	2009	Canada	prosthesis
Total hip arthroplasty with accolade/trident through the direct minimally invasive anterior approach without traction table: Learning curve and results after a minimum of 5 years.	Berndt	2019	Switzerland	prosthesis
Learning curves in hip fracture surgery.	Bjorgul	2011	USA	Trauma
Results of the learning curve for interventional hip arthroscopy: a prospective study.	Boden	2014	UK	Arthroscopy
102 lumbar pedicle subtraction osteotomies: one surgeon's learning curve.	Bourghli	2018	France	Spine
The direct anterior minimal invasive approach in total hip replacement: a prospective departmental study on the learning curve	Brun	2018	Norway	prosthesis
Evaluation of the learning curve associated with uncemented primary porous-coated anatomic total hip arthroplasty.	Callaghan	1992	USA	prosthesis
Arthroscopic Latarjet procedure: analysis of the learning curve.	Castricini	2013	Italy	Arthroscopy
The effect of learning curve on the results of percutaneous transforaminal endoscopic lumbar discectomy.	Chaichankul	2012	Thailand	Spine
A learning curve of total knee arthroplasty (TKA) based on surgical volume analysis.	Cheng	2011	Taiwan	prosthesis
Analysis of learning curve of minimally invasive total knee arthroplasty: A single surgeon's experience with 4017 cases over a 9-year period.	Cheng	2019	Taiwan	prosthesis
Total knee arthroplasty using patient-specific guides: Is there a learning curve?	Chinnappa	2015	Australia	prosthesis
Comparison of arthroscopic and open Latarjet with a learning curve analysis.	Cunningham	2016	Switzerland	Arthroscopy
Complications of limb lengthening. A learning curve.	Dahl	1994	USA	Pediatric
Progress in Knee Arthroscopy Skills of Residents and Medical Students: A Prospective Assessment of Simulator Exercises and Analysis of Learning Curves.	Dammerer	2018	Austria	Arthroscopy
Is the Latarjet procedure risky? Analysis of complications and learning curve.	Dauzere	2016	France	Arthroscopy

Value of the cumulative sum test for the assessment of a learning curve: Application to the introduction of patient-specific instrumentation for total knee arthroplasty in an academic department	DeGori	2017	France	prosthesis
What is the learning curve for the anterior approach for total hip arthroplasty?	deSteiger	2015	Australia	prosthesis
Does a Learning Curve Exist for the Surgical Treatment of Proximal Humerus Fractures?	Egol	2018	USA	Trauma
Navigation improves the learning curve of transforaminal percutaneous endoscopic lumbar discectomy.	Fan	2017	China	Spine
What Is the Expected Learning Curve in Computer-assisted Navigation for Bone Tumor Resection?	Farfalli	2017	Argentina	Tumor
The learning curve with a new cephalomedullary femoral nail.	Fine	2017	UK	Trauma
Learning curve of thoracic pedicle screw placement using the free-hand technique in scoliosis: how many screws needed for an apprentice?	Gang	2012	China	Spine
Perioperative aspects of endoscopic anterior scoliosis surgery: the learning curve for a consecutive series of 100 patients.	Gatehouse	2007	Australia	Spine
Soft tissue balancing in total knee arthroplasty using sensor-guided assessment: is there a learning curve?	Gharaibeh	2018	Australia	prosthesis
The learning curve of pedicle screw placement: how many screws are enough?	Gonzalvo	2009	Australia	Spine
Hip-fracture osteosynthesis training: exploring learning curves and setting proficiency standards.	Gustafsson	2019	Denmark	Trauma
Arthroscopic rotator cuff repair: the learning curve.	Guttmann	2005	USA	Arthroscopy
Learning curve with minimally invasive unicompartmental knee arthroplasty.	Hamilton	2010	USA	prosthesis
The learning curve for the direct anterior approach for total hip arthroplasty: a single surgeon's first 500 cases.	Hartford	2017	USA	prosthesis
The knee arthroscopy learning curve: quantitative assessment of surgical skills.	Hodgins	2014	Canada	Arthroscopy
Tunnel positioning in anterior cruciate ligament reconstruction: how long is the learning curve?	Hohmann	2010	Australia	Arthroscopy
Retention of arthroscopic shoulder skills learned with use of a simulator. Demonstration of a learning curve and loss of performance level after a time delay.	Howells	2009	UK	Arthroscopy
What is the learning curve for robotic-assisted pedicle screw placement in spine surgery?	Hu	2014	USA	Spine
Learning curves in orthopaedic surgery: a case for super-specialisation?	Jain	2007	UK	Pediatric
Learning curve in navigated total knee replacement. A multi-centre study comparing experienced and beginner centres.	Jenny	2008	Europe	prosthesis
Is there a new learning curve with transition to a new resurfacing system?	Johnson	2011	UK	prosthesis
Preliminary Results and Learning Curve of the Minimally Invasive Chevron Akin Operation for Hallux Valgus.	JowettCRJ	2017	Australia	Foot
Adoption of Robotic vs Fluoroscopic Guidance in Total Hip Arthroplasty: Is Acetabular Positioning Improved in the Learning Curve?	Kamara	2017	USA	prosthesis
Accuracy of cup orientation and learning curve of the accelerometer-based portable navigation system for total hip arthroplasty in the supine position.	Kamenaga	2019	Japan	prosthesis
Minimally invasive surgical technique in total knee arthroplasty: a learning curve.	Kashyap	2009	UK	prosthesis
Hip arthroscopy learning curve: a prospective single-surgeon study.	Kautzner	2018	Czech Republic	Arthroscopy

Robotic-arm assisted total knee arthroplasty has a learning curve of seven cases for integration into the surgical workflow but no learning curve effect for accuracy of implant positioning.	Kayani	2019	UK	prosthesis
The learning curve associated with robotic-arm assisted unicompartmental knee arthroplasty: a prospective cohort study.	Kayani	2018	UK	prosthesis
A complication-based learning curve from 200 reverse shoulder arthroplasties.	Kempton	2011	USA	prosthesis
Arthroscopic treatment for limitation of motion of the elbow: the learning curve.	Kim	2011	Korea	Arthroscopy
Minimally invasive total knee arthroplasty compared with traditional total knee arthroplasty. Assessment of the learning curve and the postoperative recuperative period.	King	2007	USA	prosthesis
Hip arthroscopy: analysis of a single surgeon's learning experience.	Konan	2011	UK	Arthroscopy
Adopting the direct anterior approach: experience and learning curve in a Chinese patient population.	Kong	2019	China	prosthesis
Arthroscopic Latarjet Stabilization: Analysis of the Learning Curve in the First 90 Primary Cases: Early Clinical Results and Computed Tomography Evaluation.	Kordasiewicz	2019	Poland	Arthroscopy
Learning curve for a modified Watson-Jones minimally invasive approach in primary total hip replacement: analysis of complications and early results versus the standard-incision posterior approach.	Laffosse	2006	France	prosthesis
The Learning Curve by Operative Time for Soft Tissue Balancing in Total Knee Arthroplasty Using Electronic Sensor Technology.	Lakra	2019	USA	prosthesis
Learning curve and clinical outcomes of minimally invasive transforaminal lumbar interbody fusion: our experience in 86 consecutive cases.	Lee	2012	Korea	Spine
Learning curve of a complex surgical technique: minimally invasive transforaminal lumbar interbody fusion (MIS TLIF).	Lee	2014	Singapore	Spine
Perioperative complications and learning curve of the Mobility Total Ankle System	Lee	2013	Korea	prosthesis
Cervical pedicle screw placement using the "key slot technique": the feasibility and learning curve.	Lee	2012	Korea	Spine
Learning curve of acetabular cup positioning in total hip arthroplasty using a cumulative summation test for learning curve (LC-CUSUM).	Lee	2014	Korea	prosthesis
Learning curve of basic hip arthroscopy technique: CUSUM analysis.	Lee	2013	Korea	Arthroscopy
Learning Curve of Internal Fixation for Nondisplaced Femoral Neck Fractures: A Cumulative Sum Analysis.	Lee	2018	Korea	Trauma
Learning Curve and Clinical Outcomes of Performing Surgery with the InterTan Intramedullary Nail in Treating Femoral Intertrochanteric Fractures.	Li	2017	China	Trauma
Surgeon learning curve for pediatric supracondylar humerus fractures	Liu	2011	USA	Trauma
Thoracic pedicle screw instrumentation: the learning curve and evolution in technique in the treatment of adolescent idiopathic scoliosis	Lonner	2009	USA	Spine
The learning curve associated with thoracoscopic spinal instrumentation.	Lonner	2005	USA	Spine
Minimally invasive surgery in total knee arthroplasty: the learning curve.	Lubowitz	2007	USA	prosthesis
The learning curve associated with anteromedial portal drilling in ACL reconstruction.	Luthringer	2016	USA	Arthroscopy
Anterior Cervical Discectomy and Fusion: The Surgical Learning Curve.	Mayo	2016	USA	Spine
Defining the Learning Curve for Hip Arthroscopy: A Threshold Analysis of the Volume-Outcomes Relationship.	Mehta	2018	USA	Arthroscopy

First experiences with the direct anterior approach in lateral decubitus position: learning curve and 1 year complication rate.	Melman	2015	Netherlands	prosthesis
Minimally invasive transforaminal lumbar interbody fusion: one surgeon's learning curve.	Nandyala	2014	USA	Spine
Defining the pediatric spinal thoracoscopy learning curve: sixty-five consecutive cases.	Newton	2000	USA	Spine
The learning curve of lateral access lumbar interbody fusion in an Asian population: a prospective study.	Ng	2015	Singapore	Spine
Re-learning curve for conventional total knee arthroplasty following 30 consecutive computer-assisted total knee arthroplasties.	Nicholson	2013	USA	prosthesis
The learning curve for adopting hip resurfacing among hip specialists.	Nunley	2010	USA	prosthesis
The AO distal locking aiming device: an evaluation of efficacy and learning curve.	Pardiwala	2001	India	Trauma
Learning Curve for Lumbar Decompressive Laminectomy in Biportal Endoscopic Spinal Surgery Using the Cumulative Summation Test for Learning Curve.	Park	2019	Korea	Spine
The learning curve for performing three-column osteotomies in adult spinal deformity patients: one surgeon's experience with 197 cases.	Raad	2019	USA	Spine
Performance of medical students on a virtual reality simulator for knee arthroscopy: an analysis of learning curves and predictors of performance.	Rahm	2016	Switzerland	Arthroscopy
The learning curve associated with robotic-assisted total hip arthroplasty.	Redmond	2015	USA	prosthesis
Incidence of Complications During Initial Experience with Revision of the Agility and Agility LP Total Ankle Replacement Systems: A Single Surgeon's Learning Curve Experience.	Roukis	2015	USA	prosthesis
Quantitative analysis of a spinal surgeon's learning curve for scoliosis surgery.	Ryu	2016	Korea	Spine
Learning curve for placement of thoracic pedicle screws in the deformed spine.	Samdani	2010	USA	Spine
Comparison of the short-term results of the first and last 50 Scandinavian total ankle replacements: assessment of the learning curve in a consecutive series.	Schimmel	2014	Netherlands	prosthesis
Anterior-supine minimally invasive total hip arthroplasty: defining the learning curve.	Seng	2009	USA	prosthesis
Restoring Sesamoid Position in Scarf Osteotomy: A Learning Curve.	Seng	2015	Singapore	Foot
Does computer-assisted surgery improve accuracy and decrease the learning curve in hip resurfacing? A radiographic analysis.	Seyler	2008	USA	prosthesis
Computer navigated total knee arthroplasty: the learning curve.	Smith	2010	Scotland	prosthesis
Fluoroscopy Learning Curve in Hip Arthroscopy-A Single Surgeon's Experience.	Smith	2017	USA	Arthroscopy
The Learning Curve Associated with Robotic Total Knee Arthroplasty.	Sodhi	2018	USA	prosthesis
Video-assisted thoracoscopic surgery in idiopathic scoliosis: evaluation of the learning curve.	Son-Hing	2007	USA	Spine
Evaluation of the Learning Curve When Transitioning From Posterolateral to Direct Anterior Hip Arthroplasty: A Consecutive Series of 1000 Cases.	Stone	2018	USA	prosthesis
Learning Curve for Percutaneous Endoscopic Lumbar Discectomy in Bi-needle Technique Using Cumulative Summation Test for Learning Curve.	Sun	2019	China	Spine
Cup positioning in primary total hip arthroplasty using an imageless navigation device: is there a learning curve?	Thorey	2009	Germany	prosthesis
Identifying the learning curve for total ankle replacement using a mobile bearing prosthesis.	Usuelli	2017	Italy	prosthesis

Learning curve of direct anterior total hip arthroplasty : a single surgeon experience.	VanDenEeden	2018	Netherlands	prosthesis
Learning curve analysis of the Collum Femoris Preserving total hip surgical technique.	VanOldenrijk	2013	International	prosthesis
Is there a learning curve in foot and ankle surgery?	Walton	2012	UK	Foot
An evaluation of the learning curve for a complex surgical technique: the full endoscopic interlaminar approach for lumbar disc herniations.	Wang	2011	USA	Spine
Learning curve for percutaneous endoscopic lumbar discectomy depending on the surgeon's training level of minimally invasive spine surgery.	Wang	2013	China	Spine
Learning from the learning curve in total hip resurfacing: a radiographic analysis.	Withes	2009	Netherlands	prosthesis
Learning curves of percutaneous endoscopic lumbar discectomy in transforaminal approach at the L4/5 and L5/S1 levels: a comparative study.	Wu	2016	China	Spine
Quantification of the learning curve for arthroscopic suprascapular nerve decompression: an evaluation of 300 cases.	Yamakado	2015	Japan	Arthroscopy
Quantification of the Learning Curve for Arthroscopic Os Trigonum Excision.	Yamakado	2018	Japan	Arthroscopy
Clinical and Radiographic Outcomes With Assessment of the Learning Curve in Arthroscopically Assisted Latissimus Dorsi Tendon Transfer for Irreparable Posterosuperior Rotator Cuff Tears.	Yamakado	2017	Japan	Arthroscopy
Orthopaedic trauma surgeons and direct anterior total hip arthroplasty: evaluation of learning curve at a level I academic institution.	York	2017	USA	Trauma
Clinical accuracy of cervical pedicle screw insertion using lateral fluoroscopy: a radiographic analysis of the learning curve.	Yoshimoto	2009	Japan	Spine
The learning curve for minimally invasive Oxford phase 3 unicompartmental knee arthroplasty: cumulative summation test for learning curve (LC-CUSUM).	Zhang	2014	China	prosthesis
Learning and teaching abilities of a newly inaugurated operation technique. Analysis of learning curves and transferability exemplified by Kypho-IORT	Bludau	2013	Germany	Spine
Computer-assisted total-knee arthroplasty. Comparison of two successive systems. Learning curve	Bove	2008	France	prosthesis
Characteristics of the learning curve in total hip endoprosthesis exemplified by the BiContact prosthesis	Flamme	2001	Germany	prosthesis
Arthroscopic repair of anterior relapsing shoulder dislocation. Is there a learning curve?	Garcia lamas	2014	Spain	Arthroscopy
Learning curve of arthroscopic hip surgery	Vilchez	2010	Spain	Arthroscopy