Predictors of clinical outcome in total hip and knee replacement : a methodological appraisal of implants and patient factors
Keurentjes, J.C.

Citation
Keurentjes, J. C. (2014, September 30). *Predictors of clinical outcome in total hip and knee replacement : a methodological appraisal of implants and patient factors*. Retrieved from https://hdl.handle.net/1887/28958

Cover Page

Universiteit Leiden

Author: Keurentjes, Johan Christiaan
Title: Predictors of clinical outcome in total hip and knee replacement : a methodological appraisal of implants and patient factors
Issue Date: 2014-09-30

6

Minimal Clinically Important Differences in Health-Related Quality of Life after Total Hip or Knee Replacement

JC Keurentjes[1], FR Van Tol[1], M Fiocco[2], JW Schoones[3], RG Nelissen[1]

1 Orthopaedic Surgery, Leiden University Medical Center.

2 Medical Statistics and BioInformatics, Leiden University Medical Center.

3 Walaeus Library, Leiden University Medical Center.

Abstract

Objectives We aimed first to summarise minimal clinically important differences (MCIDs) after total hip (THR) or knee replacement (TKR) in health-related quality of life (HRQoL), measured using the Short-Form 36 (SF-36). Secondly, we aimed to improve the precision of MCID estimates by means of meta-analysis.

Methods We conducted a systematic review of English and non-English articles using MEDLINE, the Cochrane Controlled Trials Register (1960 – 2011), EMBASE (1991 – 2011), Web of Science, Academic Search Premier and Science Direct. Bibliographies of included studies were searched in order to find additional studies. Search terms included MCID or minimal clinically important change, THR or TKR and Short-Form 36. We included longitudinal studies that estimated MCID of SF-36 after THR or TKR.

Results Three studies met our inclusion criteria, describing a distinct study population: primary THR, primary TKR and revision THR. No synthesis of study results can be given.

Conclusions Although we found MCIDs in HRQoL after THR or TKR have limited precision and are not validated using external criteria, these are still the best known estimates of MCIDs in HRQoL after THR and TKR to date. We therefore advise these MCIDs to be used as absolute thresholds, but with caution.

Introduction

Total hip (THR) and knee replacement (TKR) are effective surgical interventions, which alleviate pain and improve function and health-related quality of life (HRQoL) in patients with end-stage degeneration of the hip or knee joint, respectively.[190] Typically, studies report the mean improvement in HRQoL at the population level, which provides information for the average patient in a population. However, this information may not be meaningful for individual patients encountered in clinical practice, who will be concerned with the likelihood that they will experience a meaningful improvement in return for the risk undertaken when undergoing an intervention.[191] More relevant for the individual patient therefore is the minimal clinically important difference (MCID), defined as the

minimal difference in scores of an outcome measure that is perceived by patients as beneficial or harmful.[167, 168] The MCID enables patients to be classified as either a responder or a non-responder to a particular therapy, based on their own assessment of their pre- and post-operative HRQoL. Additionally, the MCID allows an estimation of the probability of a relevant improvement in HRQoL of a particular therapy.

Expected benefits of treatment must be weighed against its adverse effects, inconvenience and costs.[192] Therefore, there is not necessarily a single MCID value for any one outcome measure of HRQoL, which can be used for all applications and patient samples.[193] For instance, the benefits of treatment in patients suffering from end-stage osteoarthritis are considerably larger for THR and TKR compared with rehabilitational interventions. On the other hand, the risk of adverse effects is also considerably higher. These differences complicate the direct use of MCIDs in HRQoL as established for rehabilitational interventions,[194] in THR or TKR patients. The use of specific MCIDs in HRQoL after THR or TKR should be encouraged.

MCIDs can be established using two different methods. Anchor-based approaches use an external indicator to assign patients into several groups reflecting different amounts of change in health status.[193] The within-person global change rating is often used as an anchor, which is measured using Likert scales, ranging from five to 15 options.[193] Positive MCIDs are usually estimated by the mean difference between pre- and post-operative scores of patients, who indicate that their condition is "somewhat better"; negative MCIDs are usually estimated by the mean difference between pre- and post-operative scores of patients who indicate that their condition is "somewhat worse".[192, 195] Distribution-based methods offer another approach in the estimation of MCIDs, which interpret results in terms of the relation between the magnitude of effect and some measure of variability in results.[192] Individual effect size standards are often used to estimate the MCID, which is defined as the difference between a patient's pre- and post-operative HRQoL scores, normed to the standard deviation of the pre-operative scores.[195] Generally accepted individual effect size standards are

equal to the group effect size standards, as defined by Cohen.[196] Therefore, the MCID is calculated by multiplying the standard deviation of patients at baseline by 0.5.

Recently Quintana et al[169] and Escobar et al[170] have estimated MCIDs for the SF-36 after THR and TKR. However, these authors have advised against using the found estimates of MCIDs as absolute thresholds, due to the imprecision of these estimates caused by small sample sizes. The precision of an estimate can be enhanced by pooling results of multiple studies in a meta-analysis. Therefore, the purpose of our study was to enhance the precision of the MCIDs after THR and TKR, by means of a systematic review and meta-analysis.

Methods

This systematic review was performed in November 2011, using the PRISMA-Statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) as a guideline in the development of the study protocol and the report of the current study.[36] The inclusion criteria and methods of analysis were specified in advance and documented in a protocol.

Information sources and search strategy   Longitudinal studies that estimate the MCID in HRQol, measured using the Short-Form 36 (SF-36), after primary or revision THR or TKR, were eligible for inclusion. No language, publication date, or publication status restrictions were imposed. Studies were identified by searching electronic databases. No limits were applied for language and foreign papers were translated. This search strategy was applied to PubMed, MEDLINE, Embase, Web of Science, COCHRANE, ScienceDirect and Academic Search Premier. The search was run on 8 November 2011. The following search terms were used in PubMed, and were adapted for the other databases:

(Mcid[tw] OR cid[tw] OR "Minimal clinically important differences" OR "Minimal clinically important difference" OR "clinically important differences" OR "clinically important difference" OR MCIC[tw] OR "Minimal clinically important changes" OR "Minimal

clinically important change" OR "clinically important changes" OR "clinically important change" OR "Minimal clinical important differences" OR "Minimal clinical important difference" OR "clinical important differences" OR "clinical important difference" OR "clinical important changes" OR "minimal detectable change" OR "minimal detectable changes" OR "minimally detectable change" OR "meaningful changes" OR "meaningful change") AND (tka[tw] OR "knee replacement arthroplasty" OR "knee arthroplasty" OR "knee replacement" OR "knee prosthesis" OR tha[tw] OR "hip replacement arthroplasty" OR "hip arthroplasty" OR "hip replacement" OR "hip prosthesis" OR "knee" OR "knees" OR "hip" OR "hips") AND ("SF36" OR "SF-36" OR "short form 36" OR "shortform 36")

Study selection Two authors (JCK and FRvT) independently screened titles and abstracts of the papers resulting from the database search using predefined eligibility criteria. Papers were considered eligible for inclusion if they met two criteria; they were to concern primary or revision THR or TKR and should include an estimate of a MCID. The full text of all included papers, based on titles and abstracts, were screened using the same inclusion criteria. Disagreements between reviewers were resolved by consensus.

HRQoL measured using SF-36 The SF-36 consists of 36 items, covering eight domains (physical function, role physical, bodily pain, general health, vitality, social function, role emotional, and mental health), for which a transformed score is calculated (100 indicating no symptoms and 0 indicating extreme symptoms).[197]

Data collection process and data items Both authors extracted the data independently, using a predefined data extraction form. Areas of disagreement or uncertainty were resolved by consensus. Estimates of MCIDs were extracted from included studies. For anchor-based estimates of MCIDs, we extracted the number of patients, on which the estimate was based, and the standard deviation. For distribution-based estimates, we extracted the number of patient on which the estimate was based. Additionally, study characteristics, concerning follow-up period, sample size, proportion of patients who underwent joint replacement for osteoarthritis, proportion of males,

mean patient age and proportion lost to follow-up, were collected.

Risk of bias in individual studies   We assessed the risk of bias in the included studies through a modified Newcastle-Ottawa Quality Assessment Scale,[198] which included the following questions: "which approach was used to estimate the MCID?" (anchor-based versus distribution-based); "was any form of additional validation performed?" (yes / no); "was the study population representative of THR or TKR in general?" (truly representative / somewhat representative / selected population / not enough information given); "was the follow-up adequate?" (no loss to follow-up / < 5% lost to follow-up (unlikely to bias results) / > 5% lost to follow-up (results possibly biased)). We chose the cut-off point of 5% lost to follow-up according to Pijls et al,[37] who established this threshold for observational studies in orthopaedic literature, using a Delphi approach to form consensus between a group of experts in the fields of THR, TKR or evidence-based medicine.

Summary measures and planned methods of analysis   The primary outcome measure was the MCID in HRQoL, measured using SF-36, for primary THR, primary TKR, revision THR and revision TKR. Whenever possible, estimates of MCIDs were pooled using inverse variance weighting. 95% confidence intervals (CI) were calculated for all MCID estimates.

Results

Study selection   The search strategy revealed a total of 126 results (fig. 6.1 on the facing page). After removal of duplicate entries, 114 unique papers remained. Screening of titles and abstracts revealed 29 papers eligible for inclusion. Further assessment of eligibility, based on full-text papers, led to the exclusion of 26 papers: two did not address THR or TKR and 24 presented no estimation of an MCID. This left three papers, describing three studies, for further analysis.[169, 170, 199]

Study characteristics   An overview of the study characteristics of all included studies is presented in table 6.1 (p. 88). Quintana et al[169] describe the MCID in SF-36 after
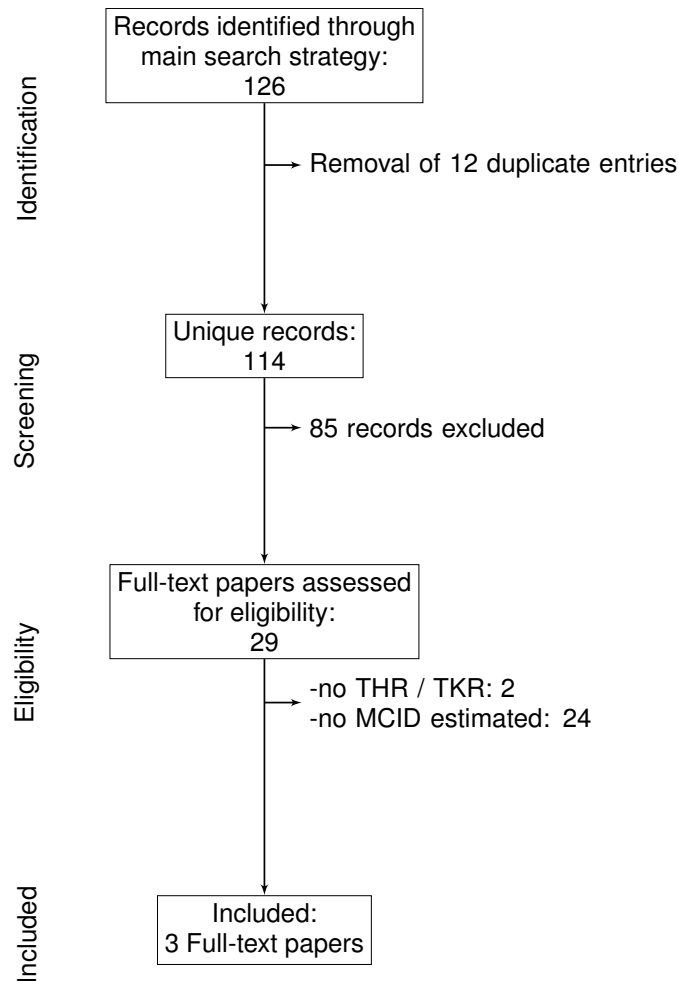
Figure 6.1: **Flow-chart of study inclusion.** THR, total hip replacement; TKR, total knee replacement; MCID, minimal clinically important difference.

primary THR at follow-up periods of six months and two years; Escobar et al[170] describe the MCID in SF-36 after primary TKR at follow-up periods of six months and two years; and Shi et al[199] describe the MCID in SF-36 after revision THR at a follow-up period of six months (table 6.1 on the following page). All included studies were multi-center studies. All studies estimated positive MCIDs (i.e. the minimal difference in scores of the SF-36 that is perceived by patients as beneficial); no study estimated negative MCIDs (the minimal difference in scores of an outcome measure that is perceived by patients as harmful). The sample for the estimation of the MCIDs was 43 patients after six months and 33 after two years for Quintana et al[169]; 76 after six months and 65 after two years for Escobar et al[170]; and 67 after six months for Shi et al[199] (table 6.1 on the next page). The indication for joint replacement was

Ref.	Intervention	Follow-up	Setting	+/-[1]	Study Sample Size	Effective Sample size[2]	OA[3] (%)	Male (%)	Mean age (yrs)	Lost to follow-up (%)
[169]	Primary THR	6 months	Multi-centre	+	485	43	100	49,3	69,4	21,86
	Primary THR	2 years	Multi-centre	+	310	33	100	49,3	69,4	36,08
[170]	Primary TKR	6 months	Multi-centre	+	423	76	100	25	71,6	22,39
	Primary TKR	2 years	Multi-centre	+	364	65	100	25	71,6	33,21
[199]	Revision THR	6 months	Multi-centre	+	67	67	n/a	56,7	70,2	16,25

Table 6.1: **Study characteristics of the three included studies.** 1: Positive or negative MCID estimated; 2: Sample size for MCID estimation; 3: Osteoarthritis.

Authors	Follow-up	MCID methodology	Additional validation	Representativeness of study population	Adequacy of follow-up
Quintana[169]	6 months	Anchor-based	No	Truly representative	>5% lost
	2 years	Anchor-based	No	Truly representative	>5% lost
Escobar[170]	6 months	Anchor-based	No	Truly representative	>5% lost
	2 years	Anchor-based	No	Truly representative	>5% lost
Shi[199]	6 months	Distribution-based	No	Description of cohort incomplete	>5% lost

Table 6.2: **Risk of bias within the three included studies.**

osteoarthritis in all patients of Quintana et al[169] and Escobar et al,[170] while Shi et al[199] offered no statement of the indication for joint replacement (table 6.1). In all studies, some patients were lost to follow-up.

Risk of bias within studies An overview of the risk of bias within studies is presented in table 6.2. Two studies used anchor-based approaches to estimate the MCID,[169, 170] while the other used a distribution-based approach.[199] No study performed any form of additional validation. The study populations of Quintana et al[169] and Escobar et al[170] are truly representative of THR and TKR patients in general, while Shi et al[199] did not provide enough information to assess the representativeness by leaving out the indication of joint replacement. All studies lost >5% of patients to follow-up, rendering a possibility of biased results.

Synthesis of results All studies have described a distinct study population, precluding any meaningful synthesis of study results. An overview of the results of all individual studies is presented in table 6.3 and figure 6.2 and figure 6.3. The MCIDs are presented with 95% confidence intervals for each of the SF-36 domains in primary

| | SF-36 domain | MCID (95% confidence interval) | |
		At six months	At two years
Primary THR[169]	Physical functioning	20.40 (14.4 – 26.4)	8.29 (-1.8 – 18.4)
	Role physical	10.78 (1.5 – 20.0)	11.00 (-1.3 – 23.3)
	Bodily pain	14.67 (6.8 – 22.6)	18.34 (9.1 – 27.6)
	General health	0.40 (-5.2 – 6.0)	-6.37 (-10.9 – -1.9)
	Vitality	10.14 (3.1 – 17.2)	14.51 (6.4 – 22.6)
	Social functioning	8.63 (0.9 – 16.4)	17.97 (7.8 – 28.1)
	Role emotional	-6.45 (-24.5 – 11.6)	20.83 (-0.6 – 42.3)
	Mental health	8.99 (2.3 – 15.7)	16.15 (9.0 – 23.3)
Primary TKR[170]	Physical functioning	11.57 (6.5 – 16.7)	11.07 (5.8 – 16.3)
	Role physical	11.69 (3.8 – 19.6)	13.16 (3.5 – 22.8)
	Bodily pain	16.86 (9.7 – 24.0)	6.69 (-0.4 – 13.8)
	General health	0.85 (-3.2 – 4.9)	-7.30 (-11.3 – -3.3)
	Vitality	3.86 (-1.7 – 9.4)	3.44 (-2.2 – 9.1)
	Social functioning	11.66 (3.7 – 19.6)	6.15 (-1.7 – 14.0)
	Role emotional	7.65 (-4.5 – 19.8)	2.42 (-9.2 – 14.1)
	Mental health	-0.32 (-5.5 – 4.9)	4.02 (-1.7 – 9.7)
Revision THR[199]	Physical functioning	3.25 (2.8 – 3.9)	-
	Role physical	4.78 (4.1 – 5.8)	-
	Bodily pain	14.91 (12.7 – 18.0)	-
	General health	14.12 (12.1 – 17.0)	-
	Vitality	22.81 (19.5 – 27.5)	-
	Social functioning	15.83 (13.5 – 19.1)	-
	Role emotional	19.98 (17.1 – 24.1)	-
	Mental health	12.37 (10.6 – 14.9)	-

Table 6.3: **Minimal clinically important differences (MCIDs) in Short-Form 36 (SF-36) domains after primary and revision total hip replacement (THR) and primary total knee replacement (TKR).**

TKR and primary and revision THR at six months (fig. 6.2 on the following page, table 6.3)[169, 170, 199] and for primary TKR and THR at two years post-operatively (fig. 6.3 (p. 91), table 6.3).[169, 170]

Discussion

We have found one study describing MCIDs in SF-36 after primary THR,[169] one after primary TKR[170] and one after revision THR[199]; we did not find any studies describing MCIDs after revision TKR. As all studies have described a distinct study population, no synthesis of study results can be given. Therefore, we were unable to improve the precision of each MCID estimate.
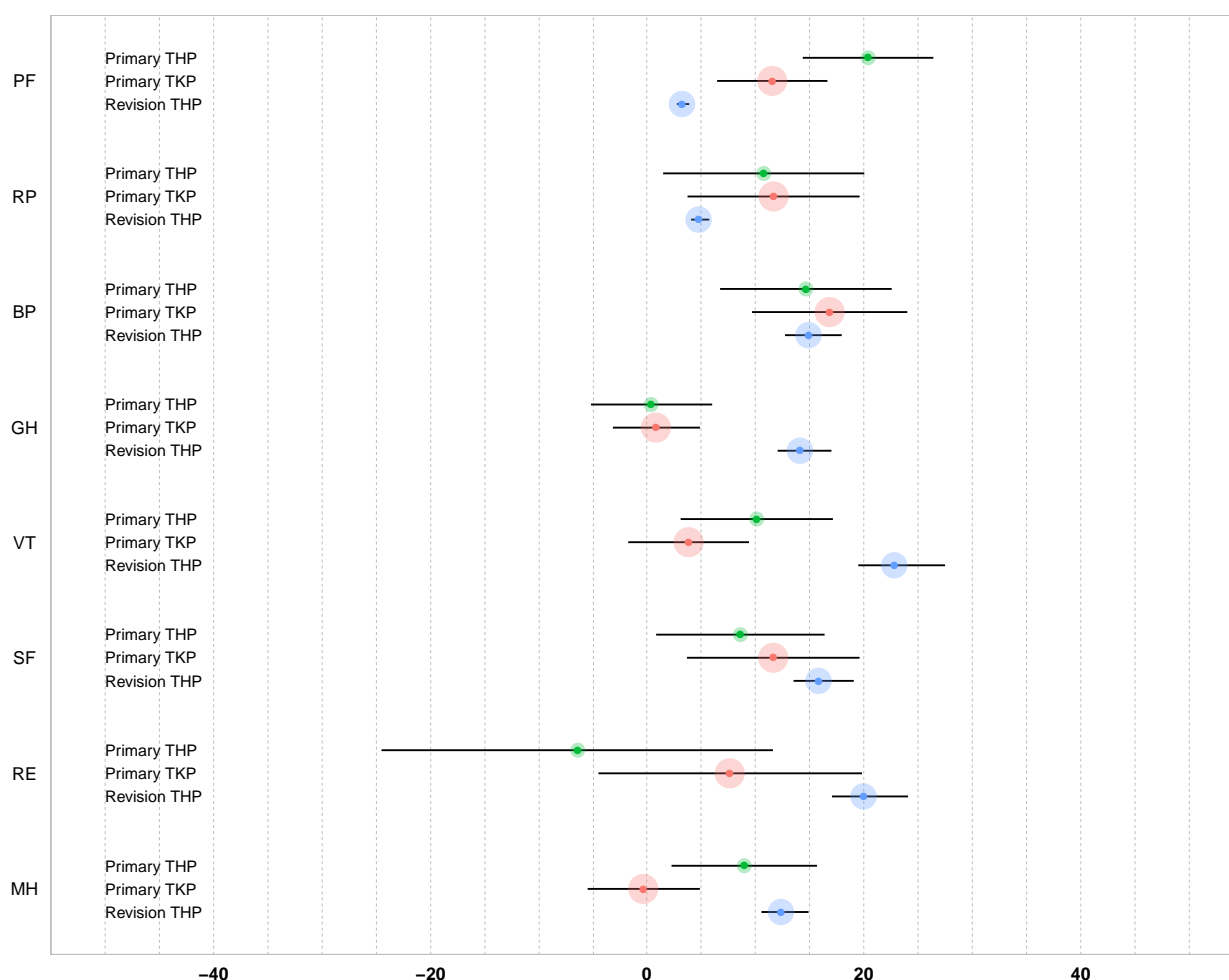
Figure 6.2: **Graph showing the minimal clinically important differences (MCIDs) in the domains of the Short-Form 36 (SF-36) at six months after primary total hip (THR)[169] and total knee replacement (TKR)[170] and revision THR.[199]** The size of the coloured circles represents the sample sizes used to estimate the MCID, and the error bars denote the 95% confidence intervals. PF, physical functioning; RP, role physical; BP, bodily pain; GH, general health; VT, vitality; SF, social functioning; RE, role emotional; MH, mental health.

However, in order to visualise the precision of all MCID estimates, we calculated 95% confidence intervals of all MCID estimates, which were not presented in the original studies. These confidence intervals are presented in figure 6.2 and figure 6.3 on the next page.

The findings of this systematic review underline the need to identify MCIDs for each specific population. As can be seen from figure 6.2 and figure 6.3 on the next page, MCIDs differ both between SF-36 subscales and patient populations. The use of a
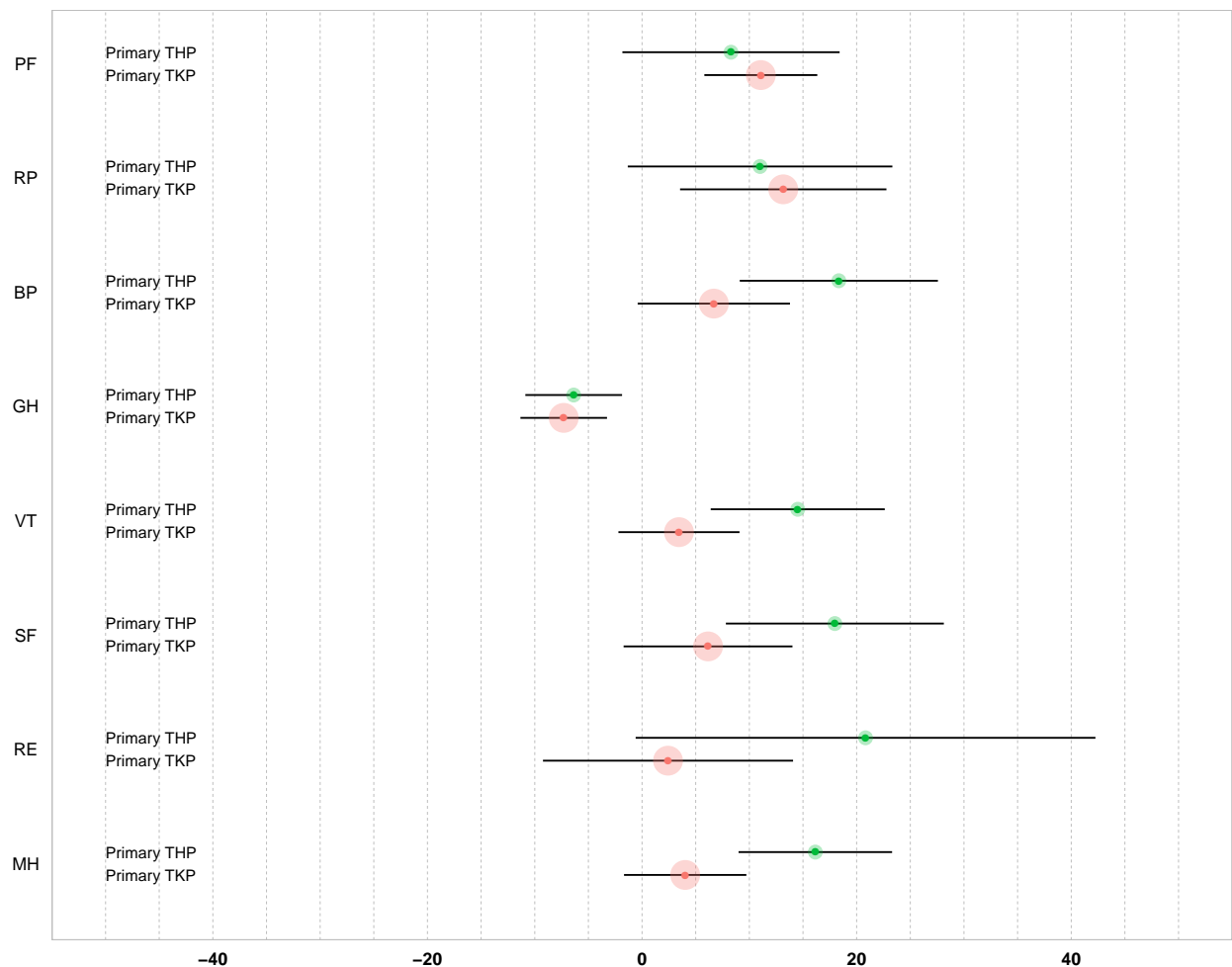
Figure 6.3: **Graph showing the minimal clinically important differences (MCIDs) in the domains of the Short-Form 36 (SF-36) at two years after primary total hip (THR)[169] and total knee replacement (TKR).[170]** The size of the coloured circles represents the sample sizes used to estimate the MCID, and the error bars denote the 95% confidence intervals. PF, physical functioning; RP, role physical; BP, bodily pain; GH, general health; VT, vitality; SF, social functioning; RE, role emotional; MH, mental health.

"one-size-fits-all" MCID does not appear justified, as patients suffering from osteoarthritis of the hip and knee, which are regarded as similar disease entities, have different MCIDs in HRQoL.[200–202] In order to study patient-relevant improvements in HRQoL at the individual level in revision TKR patients, MCIDs need to be established in this particular population as well.

Limitations of the included studies include imprecision as a result of small sample sizes, the lack of validation of the MCID estimates and the rates of loss to follow-up. Anchor-based approaches in particular suffer from imprecision due to small sample

sizes, as this approach uses only a part of all data to estimate the MCID. A precise estimation of the MCID is further hampered by the clinical success of joint replacement: typically, one expects a large effect of THR or TKR.[203] The group sizes of patients who indicate that their condition has "somewhat improved" are therefore expected to be small, which contributes to an imprecise estimation of MCIDs. Unfortunately, there are only two ways to improve the precision of anchor-based MCID estimates: one can either perform larger studies, or pool study results in a meta-analysis. To date, the only studies that have established anchor-based MCIDs in HRQoL after primary THR or TKR were those of Quintana et al[169] and Escobar et al.[170] More research is required to improve the precision of MCIDs in HRQoL. Estimates with higher precision are generated by distribution-based approaches, which use data from the entire population to estimate the MCID. However, these approaches are criticised for the arbitrariness of the individual effect size standards.[192]

A strongly recommended method of determining MCIDs is by triangulation of multiple approaches.[193] None of the included studies has applied any form of additional validation, such as secondary anchor questions; all used a single approach. Besides a limited precision, caused by small group sizes, the accuracy of the MCID estimates might be limited as well due to this lack of additional validation.[204] Therefore, further research is needed to provide external validation of the established MCIDs in HRQoL.

However, until further research is performed, the MCID estimates of these three studies[169, 170, 199] are the best available estimates. Cautious use of these estimates should be encouraged in order to study improvement in HRQoL at the individual level, the most relevant outcome measure for individual patients encountered in clinical practice.