BMJ Open

Predictors of physical functioning after total hip arthroplasty: a systematic review.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010725
Article Type:	Research
Date Submitted by the Author:	30-Nov-2015
Complete List of Authors:	Buirs, Leon; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery van Beers, Loes; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Scholtes, Vanessa; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Pastoors, Tom; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Sprague, Sheila; McMaster University, Department of Clinical Epidemiology and Biostatistics Poolman, Rudolf; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery
Primary Subject Heading:	Surgery
Secondary Subject Heading:	Rheumatology
Keywords:	total hip, arthroplasty, functional outcome, systematic review, predictors

Predictors of physical functioning after total hip arthroplasty: a systematic review

Buirs LD[1], MD; Van Beers LWAH[1], MSc; Scholtes VAB[1], PHD; Pastoors T[1]; Sprague S[2], PHD; Poolman RW[1], MD PHD

[1]Department of Orthopaedic Surgery, Onze Lieve Vrouwe Gasthuis, Amsterdam, the Netherlands.

[2]Division of Orthopedic Surgery, Centre for Evidence-Based Orthopedics, Department of Surgery, McMaster University, Hamilton, Ontario, Canada.

Objective: The objective of this systematic review of the literature was to identify the predictors of functional outcome after total hip arthroplasty.

Method: A systematic literature search in Web of Science, Cinahl, Embase and PubMed was conducted on June 23, 2015. The articles were selected based upon their quality, relevance and measurement of the predictive factor. The level of evidence of all studies was determined using the GRADE rating scheme.

Results: The initial search resulted in 1092 citations. After application of the inclusion and exclusion criteria, 33 articles met our eligibility criteria and were graded. All of the included studies were classified as level of evidence I or II. Of the included studies, 18 evaluated body mass index (BMI), 17 evaluated pre-operative status, 15 evaluated age, 15 evaluated gender, and 13 evaluated co-morbidity. There was strong evidence suggesting an association between BMI, age, comorbidity, pre-operative status, and mental health with functional outcome after THA. There was weak evidence suggesting an association between quadriceps strength and education with functional outcome after THA. The evidence was inconsistent for associations with gender and socio-economic status and functional outcome following THA. We did not find any evidence suggesting that alcohol

consumption, vitamin-D insufficiency and allergies were predictors of functional outcome following THA.

Conclusion: We have identified multiple predictors of functional outcome after THA, which will enable general practitioners and orthopedic surgeons to better predict the improvement in physical functioning for their THA patients. They can use this information to provide patient specific advice regarding the referral for THA and the expected outcomes after THA. Further research with consistent measurement tools, outcomes, and duration of follow-up across studies is needed to confirm the influence of these factors.

Keywords: total hip arthroplasty - functional outcome - systematic review - predictors

Strengths and limitations of this study

- We have carried out a comprehensive and robust systematic review in accordance with the PRISMA guidelines.
- We included all patient related predictors and did not limit ourselves to the most common predictors. This led to a complete overview of all predictors evaluated.
- We screened a large number of literature sources, and all reviewing and data extraction was carried out by one author (LDB) and double checked by a second author (LWAHB).
- Because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up we could not apply a meta-analysis.
- Some predictors as quadriceps strength, education, socioeconomic status and alcohol consumption are reported only a few times and therefore conclusions cannot be reached.

Introduction

Total hip arthroplasty (THA) is a surgical procedure performed to reduce pain and improve function in patients with osteoarthritis (OA) of the hip. According to the Agency for Healthcare Research and Quality, more than 305.000 total hip replacements are performed each year in the United States[1]. Following THA, the majority of patients experience reductions in pain, improvements in function, and better health related quality of life [2].However, not all patients achieve the same level of functional improvement after THA. Specifically, greater than 30 percent of patients undergoing THA report moderate to severe activity limitations two years post THA [3]. It is unclear which factors are causing these limitations in function [4;5].

In the last decade, many studies have been published investigating the predictors of functional outcome after THA. Most studies measured patient related predictors (e.g. gender, body mass index (BMI) and age) and non-patient related predictors (e.g. type of implant, procedure, surgeon). Some authors have considered mental health, as it has a big impact on the performance and pain after surgery [6-8]. The studies evaluating predictors of functional outcome report different predictors, use different follow-up periods, and use a variety of outcome measures. Young et al. published a systematic review on this topic more than 15 years ago[9], which did not include all relevant predictors. In addition, considerable research has been published on predictors of functional outcome has been published in the past 15 years, which justifies a new systematic review on the topic. Therefore, we conducted a systematic review to identify predictors of mid-term and long-term functional outcome after THA.

Methods

Registration

This systematic review is registered at Prospero (http://www.crd.york.ac.uk/PROSPERO/) with registry number CRD42015016929.

Selection criteria

Studies that met the following criteria were included in our review: (1) included patients undergoing a THA; (2) included physical functioning was an outcome measure; and (3) had at least one variable that was considered as a predictor of physical functioning.

Search strategy

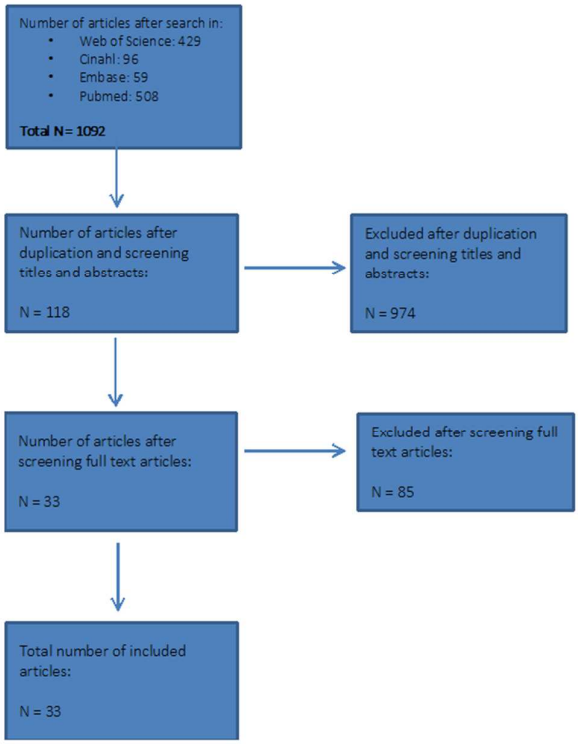
With the guidance of an independent medical librarian we conducted a literature search through four medical databases: Web of Science; Cinahl; Embase ,and PubMed. This literature search was performed on June 23, 2015. In Web of Science we used the following search terms: TOPIC: (total hip arthroplasty) *AND* TOPIC: (predictor*). In Cinahl we searched for: (MM "Arthroplasty, Replacement, Hip") AND predictor*. In Embase we searched for: exp hip arthroplasty/ exp prediction/ or exp predictor variable/ exp prognosis/ or exp functional assessment/ or exp treatment outcome/ or exp daily life activity/. In PubMed we searched for ("Arthroplasty, Replacement, Hip"[Majr] OR "Hip Prosthesis"[Majr]) AND (predictor* OR risk Factor* OR risk assessment OR predictive value of tests OR prognostic factor* OR Prognostic*) AND (hoos OR "hip disability and osteoarthritis outcome score " OR womac OR "Western Ontario and McMaster Universities Arthritis Index" OR "harris hip score" OR HHS OR SF-12 OR short form 12 OR SF 36 OR "short form 36" OR trendelenburg OR TUG OR "timed up and go" OR "oxford hip score" OR "IOWA hip score" OR "functional recovery score" OR FRS OR AFI OR "hospital for special surgery" OR aaos OR "charnley hip score" OR HSS OR LEGS OR "mayo

clinical hip score"). The results of these four different searches were combined in Reference Manager and duplicates were removed.

Study selection

Two of the authors (LWAHB and TP) independently screened the titles and abstracts of all the articles, using the above mentioned selection criteria. Both reviewers screened the full-text articles of the articles found eligible in the first round. A third author (LDB) compared these results and in case of different opinions, consensus was reached. The study selection procedure is schematically presented in Figure 1.

Figure 1: Flowchart of the study selection procedure



Data extraction

One of the authors (LDB) extracted the data, double checked by a second author (LWAHB). From each articles, the following information was extracted: (1) predictor variable; (2) author (3) year of publication; (4) level of evidence; (5) number of patients; (6) measurement tools that are used; (7) follow-up period; (8) significance level; (9) association between predictor variable and outcome

measure; (10) predictor level of measurement (Table 1) . The results were categorized by predictor variable.

Methodological quality assessment

The level of evidence of all studies was determined by one of the authors (LDB) with the GRADE rating scheme (http://www.gradeworkinggroup.org).

Measurement tools

We aimed to include all predictors mentioned in previous studies, and will not limit ourselves to the most common predictors. The measurement tools most used to define the functional outcome are the Harris Hip Score (HHS)[10], Oxford Hip Score (OHS)[11;12], Short Form-36 (SF-36)[13], LEFS (Lower Extremity Functional Scale)[14], Timed Up and Go test (TUG)[15;16] and the Western Ontario and McMaster Universities OA Index (WOMAC)[17]. We will use all mentioned measurement tools as outcome in this study.

Best evidence synthesis

A follow-up period up to 24 months was considered as 'short term' and a follow-up period of more than 24 months was considered as 'long term'. Results were divided into four categories of evidence: Strong evidence: at least 60% of the studies, with a minimum of 3 studies, describe the same significant (p-value <.05) association. Weak evidence: a) only 2 studies describe the same significant association; b)3 studies describe the same association of which 2 are significant and 1 is not significant (p-value >.05). No evidence: a) only 1 study available; b) more studies available of which none found a significant association. Inconsistent evidence: all other scenarios. No conclusions can be drawn in this literature review when no or inconsistent evidence is available.

This systematic review conforms to the PRISMA statement[18].

Results

Selection and methodological quality

The initial search resulted in 1092 citations (Figure 1) and 33 articles met our eligibility criteria. The articles included were designated as level of evidence low(11), moderate(17) or high(5)(Table 1)

TABLE 1 Methodological quality of included studies

Study	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	GRADE
Kessler, 2007	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Villalobos, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Nankaku, 2013	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Slaven, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Moran, 2005	observational study	n.a.	not serious	not serious	not serious	strong association	moderate
Stevens, 2012	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Wang, 2010	observational study	not serious	not serious	not serious	not serious	none	moderate
Dowsey, 2010	observational study	serious	not serious	not serious	not serious	strong association	low
Judge, 2014	observational study	not serious	not serious	not serious	not serious	very strong association	high
Bergschmidt, 2010	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Jones, 2012	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Smith, 2012	observational study	not serious	not serious	serious	not serious	strong association	moderate
Judge, 2013	observational study	not serious	not serious	not serious	not serious	very strong association	high
Bischoff, 2004	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Gandhi, 2010	observational study	serious	not serious	not serious	not serious	none	low
Nilsdotter, 2003	observational study	not serious	serious	not serious	not serious	strong association	low
Davis, 2012	observational study	not serious	not serious	not serious	not serious	very strong association	high
Hamilton, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Quintana, 2009	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Nilsdotter, 2002	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Dowsey, 2014	observational study	not serious	not serious	not serious	not serious	very strong association	high
Lavernia, 2010	observational study	serious	not serious	not serious	not serious	strong association	low
Mahomed, 2002	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Vogl, 2014	observational study	not serious	serious	not serious	not serious	n.a.	low
Clement, 2010	observational study	not serious	not serious	not serious	not serious	very strong association	high
Johansson, 2010	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Fortin, 2002	observational study	not serious	not serious	not serious	serious	strong association	low
Badura-Brzoza, 2009	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Holstege, 2011	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Schafer, 2010	observational study	not serious	not serious	not serious	n.a.	strong association	low
Graves, 2014	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Lavernia, 2012	observational study	not serious	not serious	not serious	n.a.	none	low
Lavernia, 2013	observational study	not serious	not serious	not serious	not serious	strong association	moderate

GRADE Grading recommendations assessment development and evaluation
High: true effect lies close to the estimate of the effect
Moderate: true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low: true effect may be substantially different from the estimate of effect
Very low: true effect is likely to be substantially different from the estimate of effect

Measures of functional outcome

Multiple outcome measures were used across these studies including the HHS, OHS, SF-36 PF, LEFS, TUG and the WOMAC score. The follow-up period ranged from 3 to 72 months with an average of 18 (SD17) months.

Predictive factors of functional outcome

BMI

Eighteen articles evaluated BMI as a potential predictor of functional outcome after THA[19-36](Table 2). A total of 14432 patients were included in all articles concerning the impact of BMI, with a mean follow-up time of 22 months. The applied levels of measurement of BMI were continuous, dichotomous or categorical.

The measurements tools used to determine the functional outcome are the WOMAC score, HHS, OHS, LEFS, SF-12 PF and the ambulatory status. The classification of a high BMI ranged from >28kg/m2 to >35kg/m2.

Of the 18 articles, 13 found a significant association [19-21;24;25;27;29-33;35;36]. Twelve articles evaluated the short-term functional outcome of which eight articles[19;22;24;27;30;32;35;36] found a significant negative association and one article a significant positive association[33]. Of the seven articles evaluating the long-term functional outcome, five articles found a significant negative association [20;21;25;29;31]. All articles were designated as level of evidence low(5), moderate(9) or high(4).

Because more than 60% of the articles report a significant negative association, there is strong evidence of a negative association between BMI and short-term and long-term functional outcome after THA.

TABLE 2: Articles reporting BMI as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	0.49	no	*cont (BMI)*
Villalobos, 2012	low	63	SF-12 PCS WOMAC HHS OHS	ST (3m)	**0.004*** **0.041*** 0.793* 0.428*	pos pos no no	*dich* *(1: BMI >28 2: BMI ≤28)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	0.06	no	*cont (BMI)*
Slaven, 2012	low	40	LEFS	ST (6m)	**n.a.**	neg	*dich* *(1: BMI >34 2: BMI ≤34)*
Moran, 2005	moderate	749	HHS	ST (6m) ST (18m)	**0.02** **0.001**	neg neg	*cont (BMI)*
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.001**	neg	*cont (BMI)*
Wang, 2010	moderate	97	WOMAC	ST (12m)	0.11	no	*cont (BMI)*
Dowsey, 2010	low	471	HHS SF-12 PCS	ST (12m)	**<0.01** **0.05**	neg neg	*cat (3)* *(1: BMI <30 2: BMI 30-39 3: BMI >30*
Dowsey, 2014	high	835	HHS	ST (12m)	**<0.0001**	neg	*cont (BMI)*
Judge, 2014	high	4413	OHS	ST (12m)	**0.003**	neg	*cat (5)* *(1:BMI 18.5-25 2:BMI 25-30 3:BMI 30-35 4:BMI 35-40 5:BMI >40*
Bergschmidt, 2010	moderate	100	HHS	ST (24m)	**0.007**	neg	*cat (3)* *(1: BMI <26 2: BMI 29-29 3: BMI >29*
Jones, 2012	moderate	231	WOMAC	ST (6m) LT (36m)	**0.001** no	neg no	*dich* *(1: BMI >35 2: BMI ≤35)*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.01**	neg	*cont (BMI)*
Judge, 2013	high	1431	OHS	LT (36m)	**<0.001**	neg	*cont (BMI)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	neg	*cont (BMI)*
Gandhi, 2010	low	636	WOMAC	LT (39m)	0.06	no	*cont (BMI)*
Nilsdotter,2003	low	211	WOMAC PF	LT (42m)	**0.03**	neg	*cont (BMI)*
Davis,2011	high	1617	HHS	LT (60m)	**<0.001**	neg	*cont (BMI)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dig= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities OA Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Age

Fifteen articles evaluated age as a possible predictor of functional outcome after THA [19;20;23;25;26;28-32;34;36-39](Table 3). A total of 9234 patients were included in all studies that identified age as a possible predictor, with a mean follow-up time of 19 months. The applied levels of measurement of age were continuous, dichotomous or categorical.

The measurements tools used to determine the functional outcome are the WOMAC score, HHS, OHS, SF-36 PF, SF-12 PF and the ambulatory status. Different classifications of high age were used, ranging from >60 to >75.

Of the 15 articles, 10 found a significant association[23;25;26;28;29;31;32;36;38;39]. Ten articles evaluated the short-term functional outcome of which six articles found a significant negative association [26;28;32;36;38;39]. Of the six articles evaluating the long-term functional outcome, five articles found a significant negative association [23;25;31;38;39]. All articles were designated as level of evidence low(4), moderate(9) or high(2).

Because more than 60% of the articles report a significant negative association, there is strong evidence of a negative association between high age and short-term and long-term functional outcome after THA.

TABLE 3: Articles reporting age as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	**0.03**	neg	*cont (age)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	yes	neg	*dich (1: age >67.5 2: age ≤67.5)*
Slaven, 2012	low	40	LEFS	ST (6m)	no	no	*dich (1: age >68.5 2: age ≤68.5)*
Hamilton, 2012	low	1410	OHS SF-12	ST (6m) ST (12m)	x x	no no	*cont (age)*
Quintana, 2009	moderate	788	WOMAC PF	ST (6m) ST (24m)	0.41 **0.001**	no neg	*dich (1: age >70 2: age ≤70)*
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.01**	neg	*cont (age)*
Wang, 2010	moderate	97	WOMAC	ST (12m)	no	no	*cont (age)*
Dowsey, 2014	high	835	HHS SF-12 PCS	ST (12m)	**<0.0001** **0.003**	neg neg	*cont (age)*
Nilsdotter, 2002	moderate	148	WOMAC PF SF-36	ST (12m)	**0.004** **0.002**	neg neg	*dich (1: age >72 2: age ≤72)*
Bergschmidt, 2010	moderate	100	HHS WOMAC SF-12	ST (12m)	>0.097 >0.097 >0.097	no no no	*cat (3) (1: age <60 2: age 60-69 3: age >69*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	x	no	*dich (1: age >75 2: age ≤75)*
Judge, 2013	high	1431	OHS	LT (36m)	**n.a.**	neg/pos	*cat (3) (1: age <50 2: age 50-60 3: age >60*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	neg	*cont (age)*
Nilsdotter, 2003	low	211	WOMAC PF	LT (43m)	**0.002**	neg	*cont (age)*
Gandhi, 2010	low	636	WOMAC SF-36	LT (39m)	**<0.05** **<0.05**	neg	*cont (age)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dig= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities OA Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Gender

Fifteen articles evaluated gender as a possible predictor of functional outcome after THA [19;20;23;24;26;28-32;34;36;38-40](Table 4). A total of 7156 patients were included in all articles that evaluated gender as a possible predictor, with a mean follow-up time of 23.3 months. The applied level of measurement of gender was dichotomous.

The measurements tools used to determine the functional outcome included the WOMAC score HHS, LEFS, SF-36 and the ambulatory status.

Of the 15 articles, seven found a statistically significant association between preoperative status and functional outcome [23;30-32;34;39;40]. Nine articles evaluated the short-term functional outcome of which four articles found a significant association [30;32;34;39]. Six articles evaluated the long-term functional outcome of which three found a significant association [23;31;40]. All articles were designated as level of evidence low(5), moderate(9) or high(1).

In four of the seven articles with a significant outcome, being male predicted a better outcome [31;32;34;39] whereas three articles reported being female as a predictor of better functional outcome [23;30;40]. This demonstrates inconsistent evidence for an association between gender and functional outcome after THA.

TABLE 4: Articles reporting gender as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	n.a.	no	*dich (1: men 2: woman)*
Nilsdotter, 2002	moderate	148	WOMAC SF-36	ST (3m) ST (12m)	0.7	no	*dich (1: men 2: woman)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	0.10	no	*dich (1: men 2: woman)*
Slaven, 2012	low	40	LEFS	ST (6m)	**0.039**	pos, woman	*dich (1: men 2: woman)*
Quintana, 2009	moderate	788	SF-36 PF	ST (6m) ST (24m)	**n.a.** n.a.	pos, men no	*dich (1: men 2: woman)*
Bergschmidt, 2010	moderate	100	HHS	ST (12m)	n.a.	no	*dich (1: men 2: woman)*
Stevens, 2012	low	653	WOMAC	ST (12m)	**0.002**	pos, men	*dich (1: men 2: woman)*
Dowsey, 2014	high	835	HHS	ST (12m)	0.06	no	*dich (1: men 2: woman)*
Wang, 2010	moderate	97	WOMAC	ST (16.8m)	**0.0001**	pos, men	*dich (1: men 2: woman)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	no	no	*dich (1: men 2: woman)*
Jones, 2012	moderate	231	WOMAC	LT (36m)	0.118	no	*dich (1: men 2: woman)*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	pos, men	*dich (1: men 2: woman)*
Gandhi, 2010	low	636	WOMAC SF-36 PF	LT (39m)	no **<0.05**	no pos, woman	*dich (1: men 2: woman)*
Lavernia, 2010	low	532	WOMAC PF	LT (42m)	**<0.001***	pos, woman	*dich (1: men 2: woman)*
Nilsdotter,2003	low	211	WOMAC PF	LT (66m)	0.37	no	*dich (1: men 2: woman)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; dig= dichotomous;**
SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities OA Index;
LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term;
LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable
pos= positive; neg= negative

Pre-operative status

Seventeen articles evaluated pre-operative status as a possible predictor of functional outcome after THA [19;25;27-31;34;36-39;41-45](Table 5). A total of 9689 patients were included in all articles that evaluated pre-operative status, with a mean follow-up time of 16 months. The applied levels of measurement of preoperative status were continuous, dichotomous or categorical.

The WOMAC score [17] is the measurement tool most used to determine the preoperative status [19;29;34;38;39;41-43;45]. Other measurement tools used are the HHS, TUG, OHS, SF-36, SF-12 and the ambulatory status.

Of the 17 articles, 16 articles found a statistically significant correlation between pre-operative status and functional outcome. Fourteen articles evaluated the short-term outcome of which 13 reported a significant association. Three articles evaluated the long-term outcome; all three found a significant association. The only study that did not report a significant association, is a study with a small patient group that used the TUG to determine the preoperative status[30]. All articles were designated as level of evidence low(5), moderate(9) or high(3).

As more than 60% of the articles report a significant negative association, there is strong evidence of a short-term and long-term association between the preoperative status and the functional outcome after THA.

TABLE 5: Articles reporting preoperative status as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Quintana, 2009	moderate	788	WOMAC PF SF-36 PF	ST (6m)	**<0.001**	yes	*cont (WOMAC + SF-36*
Slaven, 2012	low	40	TUG	ST (6m)	n.a.	no	*dich (successful/unsuccessful)*
Mahomed, 2002	moderate	103	WOMAC PF+P SF36 PF	ST (6m)	**<0.05** **<0.05**	yes	*cont (WOMAC + SF-36)*
Hamilton, 2012	low	1410	OHS SF-12	ST (6m) ST (12m)	yes	yes	*cont (OHS)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	**n.a.**	yes	*dich (TUG score 10)*
Vogl, 2014	low	281	WOMAC	ST (6m)	**n.a.**	yes	*cont (WOMAC)*
Bergschmidt, 2010	moderate	100	WOMAC SF-36	ST (12m)	<0.022 0.003	yes	*cat (3) 1: HHS ,<48 2: HHS 48-59 3: HHS >59*
Clement, 2010	high	1312	OHS SF-12	ST (12m)	**0.001***	yes	*cont (OHS)*
Johansson, 2010	moderate	75	HHS WOMAC SF-36	ST (12m)	**≤0.006** **<0.001** **≤0.005**	yes yes yes	*cat (3) 1: HHS ,<45 2: HHS 45-55 3: HHS >55*
Nilsdotter, 2002	moderate	148	WOMAC SF-36	ST (12m)	**<0.0001**	yes	*dich low quartile vs high quartile WOMAC*
Dowsey, 2014	high	835	HHS	ST (12m)	**<0.0001**	yes	*cont (HHS)*
Wang, 2010	moderate	97	WOMAC	ST (16.8m)	**0.0001**	yes	*cont (WOMAC PF)*
Moran, 2005	moderate	749	HHS	ST (18m)	**n.a.**	yes	*cont*
Fortin, 2002	low	222	WOMAC SF-36	ST (24m)	**n.a.** **n.a.**	yes yes	*dig (1: high WOMAC 2. low WOMAC*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	yes	*cont (HHS)*
Nilsdotter,2003	low	211	WOMAC PF	LT (42m)	**0.007**	yes	*dich low quartile vs high quartile SF-36 PF*
Judge, 2013	high	1431	OHS	LT (60m)	**<0.001**	yes	*cont (OHS)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dig= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities OA Index; LoE: level of evidence; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Comorbidity

Thirteen articles evaluated comorbidity as a possible predictor of functional outcome after THA (Table 6). A total of 9363 patients were included in all studies that evaluated comorbidity as a possible predictor, with a mean follow-up time of 23.3 months. The applied levels of measurement of preoperative status were continuous, dichotomous or categorical.

The measurements tools used to determine the functional outcome are the WOMAC score, HHS, LEFS, SF-36 and the ambulatory status. Most articles used the number of comorbidities as predictor of functional outcome. Other articles used the presence of a specific comorbidity as a predictor.

Of the 13 articles, 11 found a significant negative association [20;23;24;27;29;31;32;34-36;39;41;44]. Seven articles evaluated the short-term outcome of which six reported a significant negative association[24;24;25;27;32;34;36;41;44]. Six articles evaluated the long-term outcome, of which five found a significant negative association[20;23-25;31]. All articles were designated as level of evidence low(2), moderate(8) or high(3).

Because more than 60% of the articles report a significant negative association, there is strong evidence of a negative association between comorbidities and short-term and long-term functional outcome after THA.

TABLE 6: Articles reporting comorbidity status as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Quintana, 2009	moderate	788	WOMAC PF SF-36 PF	ST (6m)	n.a. n.a.	no	*cat(3)* *1: 0 comorb 2: 1-2 comorb 3: >2 comorb*
Mahomed, 2002	moderate	103	WOMAC PF+P	ST (6m)	**<0.05**	neg	*cont* *(number of comorbidities)*
Moran, 2005	moderate	749	HHS	ST (6m) ST (18m)	**n.a.**	neg	*dich* *(presence of coronary heart disease and previous thromo-embolism)*
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.01**	neg	*cat(3)* *1: 0 comorb 2: 1-2 comorb 3: >2 comorb*
Clement, 2010	high	1312	OHS SF-12	ST (12m)	**0.01**	neg	*cont* *(number of comorbidities)*
Dowsey, 2014	high	835	HHS	ST (12m)	**0.0001**	neg	*cont* *(age adjusted CCI)*
Wang, 2010	moderate	97	WOMAC	ST (16.8m)	**0.0246**	neg	*dich* *(1: >0 comorbidities 2: 0 comorbidities)*
Jones, 2012	moderate	231	WOMAC	LT (36m)	**0.012**	neg	*dig* *(1; 0 cardiac diseases* *2: >0 cardiac diseases)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	neg	*dich* *(1; >2 chron diseases* *2. 0-1 chronic diseases)*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	neg	*cont* *(asa grade)*
Gandhi, 2010	low	636	WOMAC SF-36 PF	LT (39m)	**<0.05**	neg	*cont* *(number of comorbidities)*
Nilsdotter,2003	low	211	WOMAC PF	LT (42m)	0.08	no	*dich* *(1: >1 comorbidities 2: 0-1 comorbidities)*
Judge, 2013	high	1431	OHS	LT (60m)	**0.001**	neg	*cont* *(number of comorbidities)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dig= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities OA Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Other predictors

The predictors that are evaluated in five articles or less will be described below and will be displayed all together in Table 7.

Five articles evaluated *mental health* as a possible predictor of functional outcome after THA, with a total of 3563 patients [20;25;36;39;46]. All four articles evaluating the short-term functional outcome

found a significant negative association [25;36;39;46]. Both articles that evaluated the long-term outcome found a significant negative association. Because more than 60% of the articles report a significant positive association, there is strong evidence of a positive association between mental health and short-term functional outcome after THA. Because only two studies evaluated the long-term outcome, this evidence is weak.

Two studies evaluated *alcohol consumption* as a predictor of functional outcome [20;47]. None of them found a significant result and therefore none show evidence of an association. The two articles evaluating quadriceps strength as a possible predictor [28;48] looked at the short-term functional outcome and both found a significant association. Therefore the evidence for an association is weak.

All three articles that evaluated education as a possible predictor, found a significant association [20;41;49]. Two articles evaluated the short-term outcome and both found a significant association [41;49]. Bischoff et al evaluated the long-term effect and found a significant association [20]. All three studies used the WOMAC score to measure the functional outcome. These results show weak evidence for a short-term association, and incomplete evidence for a long-term association.

Dowsey et al. reported *socio-economic status* as a predictor, using the socio-economic status score (SES) as measurement tool[36]. They did not find a significant result and therefore show no evidence of an association.

The influence of having more than 3 *allergies* on the short-term functional outcome was reported by Graves et al[50]. Those patients had diminished improvements on SF-36 PCS and WOMAC scores, 6,5 months after THA. This result shows no evidence of an association between having more than 3 allergies and functional outcome.

Lavernia et al evaluated vitamin-D insufficiency as a predictor of functional outcome after THA[51]. A preoperative 25-hydroxyvitamin-D3 plasma level of under 30 ng/ml, predicted a worse HHS

11 months postoperative. Because no other studies evaluated vitamin-D insufficiency as a possible predictor, this result shows no evidence of an association.

TABLE 7: All predictors that are evaluated in five articles or less

predictor	Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Mental health	Badura-Brzoza, 2009	moderate	102	SF-36 PCS	ST (6m)	**0.005**	neg	*cont (anxiety as a trait)*
	Quintana, 2009	moderate	788	SF-36 PF WOMAC P	ST (6m) ST (24m)	**<0.001** **0.002**	yes	*cont (SF-36 MH score)*
	Dowsey, 2014	high	835	HSS	ST (12m)	**<0.0001**	yes	*cont (SF-12 MH score)*
	Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	yes	*dich (1: >60 pts on the SF-36 MH score 2: ≤60pts on SF-36 MH score)*
	Judge, 2013	high	916	OHS	ST (12m) LT (60m)	**0.045**	yes	*cont (SF-36 MH score)*
Alcohol consumption	Bischoff, 2004	moderate	914	WOMAC PF	LT (36m)	**n.a.**	no	*dich (1: >1 alcoholic drinks per day 2: 0-1 alcoholic drinks per day)*
	Lavernia, 2012	low	191	WOMAC	LT (72m)	**n.a.**	no	*cat (3) (1: non drinkers 2: occasional drinkers 3: moderate drinkers)*
Quadriceps strength	Holstege, 2011	moderate	55	WOMAC PF	ST (3m)	**0.004**	pos	*cont (knee extensor strength)*
	Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	**n.a.**	pos	*dich (1: > 1.25 N m/kg 2: ≤1.25 m/kg knee extensor strength)*
Education	Schafer, 2010	low	1007	WOMAC	ST (6m)	**n.a.**	pos	*dich (1; >12 years school 2: <9 years school)*
	Mahomed, 2002	moderate	103	WOMAC PF+P	ST (6m)	**0.007**	pos	*cont (level of education)*
	Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	pos	*dich (1:college education 2:less than college education)*
Socio economic status	Dowsey, 2014	high	835	HHS	LT (12m)	0.63	no	*cont (SES score)*
Allergies	Graves, 2014	moderate	459	WOMAC PF SF-36 PCS	ST (6.5m)	**0.04** **0.0002**	neg	*dich (>3 allergies)*
Vitamin-D insufficiency	Lavernia, 2013	moderate	60	HHS WOMAC	ST (11m)	**0.002** 0.478	neg	*dich (25-hydroxyvitamin-D3) (1; >30 ng/ml 2: <30 ng/ml)*

*** All significant results are bold; cont= continuous; dig= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities OA Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Discussion

In this systematic literature review we sought to provide a clear overview of all patient related predictors of functional outcome after THA.

Key findings

Our review found strong evidence of an association between BMI, age, comorbidity, preoperative score, and mental health with functional outcome after THA. Weak evidence was found for the predictors quadriceps strength and education. Inconsistent evidence was found for the predictors gender and socio-economic status. No evidence was found for the predictors alcohol consumption, vitamin-D insufficiency, and allergies.

In our review thirteen studies found a significant negative association between BMI and functional outcome after THA.. A prior review of Young et al [9] found the same significant negative association. Although the review of Young et al and our current review come to the same conclusion, the clinical impact of this outcome is still questionable. A large study by Judge et al. showed a small significant correlation between a high BMI and a worse functional outcome, but conclude that the total improvement in function outweighs the small lack of improvement caused by a high BMI[35].

Although our review shows strong evidence of an association between BMI and functional outcome, different classifications of high BMI were used. Because of these different classifications, it is difficult to define a specific BMI that predicts who will do well after THA. A meta-analysis is needed to specify which BMI will have a high risk of worse outcome. We could not conduct a meta-analysis since different classifications of BMI are used and there was heterogeneity in outcome instruments. Therefore future research on the impact of BMI should use clearly defined outcomes that are consistent across studies.

In our review eight of the 14 studies found an association between higher age and poorer functional outcome, therefore age is an important factor predicting functional outcome. When determining a specific age limit, it is not clear how to apply this result clinically. Some articles used a linear regression analysis for age. When looking at age, it is not only interesting to see the effect of high age, but also of low age. Therefore linear regression analysis might not be the best statistical analysis with variables as age or BMI. For future research on the impact of age on functional outcome after THA, more consistent outcomes must be used across studies. There is no consensus among studies about what specific age limit is recommended for THA. This current review shows inconclusive evidence of an association between gender and functional outcome because six out of 14 articles found a statistically significant result.

Three studies reported being female led to a better functional outcome [23;30;40]. The other four significant articles found the opposite result where being male has a positive association with functional outcome after THA [31;32;34;39]. The results are very contradictory and the differences may be partially attributable to confounding factors.

The pre-operative status was found to be a conclusive predictor. Only one study -- Slaven et al. -- did not find a statistically significant correlation[30]. This might be due to the fact that they used the TUG score as measurement tool[30]. The WOMAC score was the measurement tool most used to define the pre-operative status (9 times)[19;29;34;38;39;41-43;45]. Other pre-operative measurement tools that have proven to be good predictors of functional outcome are the HHS score, OHS score, SF-12 PF score, SF-36 PF score, and ambulatory status.

Of the 13 articles that evaluated comorbidity as a possible predictor of functional outcome, 11 found a significant negative association [20;23-25;27;31;32;34;36;39;41;44]. Comorbidity can be measured in several ways, for example: the number of comorbidities, the presence of a specific comorbidity, the Charlson index [52] and the Elixhauser comorbidity measure [53]. Comorbidities can affect the true functional outcome after THA but can also affect the score on the measurement tool. For example: if a patient

is unable to walk to the grocery store after a THA due to a lung disease, his functional outcome score will be lower despite a possible good functioning total hip. Except for one article, all articles found a significant negative effect. And therefore having comorbidities can be seen as a predictor of negative functional outcome.

All five studies that evaluated mental health as a predictor of functional outcome found a statistically significant positive association. Four of these studies used SF-36 MH [54] as measurement tool to define mental health[20;25;36;39]. These results show strong evidence of a positive association between mental health and short-term functional outcome after THA. The two studies reporting quadriceps strength as a predictor had both small sample sizes which can affect the external validity of the studies[28;48]. Therefore this evidence is weak and more research must be done on the effect of quadriceps strength.

Three articles evaluated education as predictor of functional outcome. Mahomed et al[41] and Bischoff et al[20] used the level of school education as a predictor, and Schafer et al[49] used years of education as a predictor. Because education is in part a surrogate of socioeconomic status, this might also indicate that low socioeconomic status is a factor associated with poor functional outcome. Dowsey et al however did not find a correlation between socioeconomic status and functional outcome[36]. Future research is needed on various components of socioeconomic status to specify the impact on functional outcome. Because only one study evaluated allergies[50] and vitamin-D insufficiency[51] as possible predictors of functional outcome, no conclusions can be drawn.

Previous systematic reviews

The previous systematic review of Young et al. concluded that important research remains to be done to examine the magnitude and interaction of patient factors on the outcome of THA[9]. The review of Young et al. used only one database (MEDLINE) and is more than 15 years old. In our systematic review we used multiple databases (Web of Science; Cinahl; Embase and PubMed) and reported all predictors evaluated in literature.

Strengths and Limitations

We included all patient related predictors and did not limit ourselves to the most common predictors. This led to a complete overview of all predictors evaluated. The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up. Not all studies used in this review adjusted their outcomes for potential confounders. Therefore some outcomes may be due to confounding factors. A limitation of our review is that we looked at functional outcome without including pain. Some patients will not see an improvement in their function after THA, but will lose the hip related pain. For this reason especially people with a high BMI and older age can benefit from THA, without improving the function of the hip. Some predictors as quadriceps strength, education, socioeconomic status and alcohol consumption are reported only a few times and therefore conclusions cannot be reached. More research in large datasets is needed to draw definitive conclusions on these predictors.

Implications for practice

Our review provides a clear overview of the current literature on the predictors for physical functioning after THA. Orthopedic surgeons and general practitioners can use this information to predict the improvement in physical functioning for their patients and it enables them to provide patient specific advice on THA surgery.

Implications for future research

In the future, we suggest studies that evaluate possible predictors of functional outcome after THA to use equal measurement tools, outcomes and durations of follow-up. In that way a meta-analysis can be applied and the influence of these factors can be specified.

Conclusion

This review shows that several patient related characteristics can predict the functional outcome after THA. It shows strong evidence of an association between BMI, age, comorbidity, preoperative score and mental health with functional outcome after THA. Weak evidence suggested that quadriceps strength and education were predictive of functional outcomes after THA.. Inconsistent evidence was found for the predictors gender and socio-economic status. Alcohol consumption, vitamin-D insufficiency and allergies were not found to be predictive out outcomes after THA. Understanding predictors will help orthopedic surgeons and general practitioners predict the outcomes in physical functioning after THA and it will enable them provide patient specific advice regarding the decision to undergo THA. Large clinical trials are necessary to confirm these findings.

Acknowledgement

We would like to thank the medical librarian Bert Berenschot at Onze Lieve Vrouwe Gasthuis for his help with the literature search.

Contributors: LDB, LWAHB, VABS, TP, SS, RWP, made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data. LDB, LWAHB, VABS, SS and RWP have been involved in drafting the manuscript or revising it critically for important intellectual content. All authors read and approved the final manuscript.

Funding: The study was not funded and non of the contributors recieved funding for this study.

Competing interests:"All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; L.W.A.H. Beers reports grants from Link / Lima, grants from Stryker, outside the submitted work. Dr. Sprague reports employment/salary from McMaster University, employment/salary from Global Research Solutions, outside the submitted work. Dr. Scholtes reports grants from Link / Lima, grants from Stryker, grants from NuVasive, grants from Zonmw, grants from Achmea and grants from Tornier, outside the submitted work. Dr. Poolman reports grants from Link / Lima, grants from Stryker, grants from NuVasive, grants from Zonmw, grants from Achmea and grants from Tornier, outside the submitted work; no other relationships or activities that could appear to have influenced the submitted work."

Ethical approval: Not required.

Transparency: The lead author (LDB) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned and registered have been explained.

Data sharing: No additional data are available, though details on statistical analysis are available from the corresponding author (LDB) on request.

based in whole or part on the on the Contribution, iv) to exploit all subsidiary rights to exploit all subsidiary rights that currently exist or as may exist in the future in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above. All research articles will be made available on an Open Access basis (with authors being asked to pay an open access fee—see http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse). The terms of such Open Access shall be governed by a Creative Commons licence—details as to which Creative Commons licence will apply to the research article are set out in our worldwide licence referred to above.

Reference List

(1) *Bernstein J, Derman P. Dramatic increase in total knee replacement utilization rates cannot be fully explained by a disproportionate increase among younger patients. Orthopedics 2014; 37(7):e656-e659.*

(2) *Judge A, Cooper C, Williams S, Dreinhoefer K, Dieppe P. Patient-reported outcomes one year after primary hip replacement in a European Collaborative Cohort. Arthritis Care Res (Hoboken) 2010; 62(4):480-488.*

(3) *Singh JA, Lewallen DG. Predictors of Activity Limitation and Dependence on Walking Aids After Primary Total Hip Arthroplasty. Journal of the American Geriatrics Society 2010; 58(12):2387-2393.*

(4) *Williams O, Fitzpatrick R, Hajat S, Reeves BC, Stimpson A, Morris RW et al. Mortality, morbidity, and 1-year outcomes of primary elective total hip arthroplasty. J Arthroplasty 2002; 17(2):165-171.*

(5) *MacWilliam CH, Yood MU, Verner JJ, McCarthy BD, Ward RE. Patient-related risk factors that predict poor outcome after total hip replacement. Health Serv Res 1996; 31(5):623-638.*

(6) *Caumo W, Schmidt AP, Schneider CN, Bergmann J, Iwamoto CW, Bandeira D et al. Risk factors for preoperative anxiety in adults. Acta Anaesthesiol Scand 2001; 45(3):298-307.*

(7) *Feeney SL. The relationship between pain and negative affect in older adults: anxiety as a predictor of pain. J Anxiety Disord 2004; 18(6):733-744.*

(8) Giraudet-Le Quintrec J-S, Coste J, Vastel L, Pacault V, Jeanne L, Lamas JP et al. Positive effect of patient education for hip surgery: a randomized trial. Clin Orthop Relat Res 2003;(414):112-120.

(9) Young NL, Cheah D, Waddell JP, Wright JG. Patient characteristics that affect the outcome of total hip arthroplasty: A review. Canadian Journal of Surgery 1998; 41(3):188-195.

(10) Nilsdotter A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. Arthritis Care Res (Hoboken) 2011; 63 Suppl 11:S200-S207.

(11) Gosens T, Hoefnagels NHM, de Vet RCW, Dhert WJA, van Langelaan EJ, Bulstra SK et al. The "Oxford Heup Score": the translation and validation of a questionnaire into Dutch to evaluate the results of total hip arthroplasty. Acta Orthop 2005; 76(2):204-211.

(12) Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996; 78(2):185-190.

(13) Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992; 305(6846):160-164.

(14) Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. Phys Ther 1999; 79(4):371-383.

(15) Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39(2):142-148.

(16) Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the "get-up and go" test. Arch Phys Med Rehabil 1986; 67(6):387-389.

(17) Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988; 15(12):1833-1840.

(18) Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009; 151(4):264-9, W64.

(19) Bergschmidt P, Bader R, Finze S, Tokar I, Kundt G, Mittelmeier W. Impact of preoperative variables on the functional and radiological outcome of an uncemented femoral stem: a prospective two-year follow-up. Hip Int 2010; 20(2):187-197.

(20) Bischoff-Ferrari HA, Lingard EA, Losina E, Baron JA, Roos EM, Phillips CB et al. Psychosocial and geriatric correlates of functional status after total hip replacement. Arthritis & Rheumatism-Arthritis Care & Research 2004; 51(5):829-835.

(21) Davis AM, Wood AM, Keenan AC, Brenkel IJ, Ballantyne JA. Does body mass index affect clinical outcome post-operatively and at five years after primary unilateral total hip

replacement performed for osteoarthritis? A multivariate analysis of prospective data. J Bone Joint Surg Br 2011; 93(9):1178-1182.

(22) *Dowsey MM, Liew D, Stoney JD, Choong PFM. The impact of obesity on weight change and outcomes at 12 months in patients undergoing total hip arthroplasty. Medical Journal of Australia 2010; 193(1):17-21.*

(23) *Gandhi R, Dhotar H, Davey JR, Mahomed NN. Predicting the Longer-term Outcomes of Total Hip Replacement. J RHEUMATOL 2010; 37(12):2573-2577.*

(24) *Jones CA, Cox V, Jhangri GS, Suarez-Almazor ME. Delineating the impact of obesity and its relationship on recovery after total joint arthroplasties. Osteoarthritis Cartilage 2012; 20(6):511-518.*

(25) *Judge A, Arden NK, Batra RN, Thomas G, Beard D, Javaid MK et al. The association of patient characteristics and surgical variables on symptoms of pain and function over 5 years following primary hip-replacement surgery: a prospective cohort study. Bmj Open 2013; 3(3).*

(26) *Kessler S, Kafer W. Overweight and obesity: two predictors for worse early outcome in total hip replacement? Obesity (Silver Spring) 2007; 15(11):2840-2845.*

(27) *Moran M, Walmsley P, Gray A, Brenkel IJ. Does body mass index affect the early outcome of primary total hip arthroplasty? J Arthroplasty 2005; 20(7):866-869.*

(28) *Nankaku M, Tsuboyama T, Akiyama H, Kakinoki R, Fujita Y, Nishimura J et al. Preoperative prediction of ambulatory status at 6 months after total hip arthroplasty. Phys Ther 2013; 93(1):88-93.*

(29) *Nilsdotter AK, Petersson IF, Roos EM, Lohmander LS. Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study. Annals of the Rheumatic Diseases 2003; 62(10):923-930.*

(30) *Slaven EJ. Prediction of Functional Outcome at Six Months Following Total Hip Arthroplasty. PHYS THER 2012; 92(11):1386-1394.*

(31) *Smith GH, Johnson S, Ballantyne JA, Dunstan E, Brenkel IJ. Predictors of excellent early outcome after total hip arthroplasty. J Orthop Surg Res 2012; 7:13.*

(32) *Stevens M, Paans N, Wagenmakers R, van Beveren J, van Raay JJAM, van der Meer K et al. The Influence of Overweight/Obesity on Patient-Perceived Physical Functioning and Health-Related Quality of Life After Primary Total Hip Arthroplasty. Obesity Surgery 2012; 22(4):523-529.*

(33) *Villalobos PA, Navarro-Espigares JL, Hernandez-Torres E, Martinez-Montes JL, Villalobos M, Arroyo-Morales M. Body Mass Index as Predictor of Health-Related Quality-of-Life Changes After Total Hip Arthroplasty: A Cross-Over Study. Journal of Arthroplasty 2013; 28(4):666-670.*

(34) *Wang W, Morrison TA, Geller JA, Yoon RS, Macaulay W. Predicting short-term outcome of primary total hip arthroplasty:a prospective multivariate regression analysis of 12 independent factors. J Arthroplasty 2010; 25(6):858-864.*

(35) Judge A, Batra RN, Thomas GE, Beard D, Javaid MK, Murray DW et al. Body mass index is not a clinically meaningful predictor of patient reported outcomes of primary hip replacement surgery: prospective cohort study. Osteoarthritis Cartilage 2014; 22(3):431-439.

(36) Dowsey MM, Nikpour M, Choong PFM. Outcomes following large joint arthroplasty: does socio-economic status matter? BMC Musculoskelet Disord 2014; 15:148.

(37) Hamilton D, Henderson GR, Gaston P, MacDonald D, Howie C, Simpson AH. Comparative outcomes of total hip and knee arthroplasty: a prospective cohort study. Postgrad Med J 2012; 88(1045):627-631.

(38) Nilsdotter AK, Lohmander LS. Age and waiting time as predictors of outcome after total hip replacement for osteoarthritis. Rheumatology (Oxford) 2002; 41(11):1261-1267.

(39) Quintana JM, Escobar A, Aguirre U, Lafuente I, Arenaza JC. Predictors of Health-related Quality-of-life Change after Total Hip Arthroplasty. Clinical Orthopaedics and Related Research 2009; 467(11):2886-2894.

(40) Lavernia CJ, Alcerro JC, Contreras JS, Rossi MD. Patient perceived outcomes after primary hip arthroplasty: does gender matter? Clin Orthop Relat Res 2011; 469(2):348-354.

(41) Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. J RHEUMATOL 2002; 29(6):1273-1279.

(42) Johansson HR, Bergschmidt P, Skripitz R, Finze S, Bader R, Mittelmeier W. Impact of preoperative function on early postoperative outcome after total hip arthroplasty. J Orthop Surg (Hong Kong) 2010; 18(1):6-10.

(43) Fortin PR, Penrod JR, Clarke AE, St Pierre Y, Joseph L, Belisle P et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002; 46(12):3327-3330.

(44) Clement ND, Muzammil A, MacDonald D, Howie CR, Biant LC. Socioeconomic status affects the early outcome of total hip replacement. J Bone Joint Surg Br 2011; 93(4):464-469.

(45) Vogl M, Wilkesmann R, Lausmann C, Hunger M, Plotz W. The impact of preoperative patient characteristics on health states after total hip replacement and related satisfaction thresholds: a cohort study. Health Qual Life Outcomes 2014; 12:108.

(46) Badura-Brzoza K, Zajac P, Kasperska-Zajac A, Brzoza Z, Matysiakiewicz J, Piegza M et al. Anxiety and depression and their influence on the quality of life after total hip replacement: preliminary report. International Journal of Psychiatry in Clinical Practice 2008; 12(4):280-284.

(47) Lavernia CJ, Villa JM, Contreras JS. Alcohol use in elective total hip arthroplasty: risk or benefit? Clin Orthop Relat Res 2013; 471(2):504-509.

(48) Holstege MS, Lindeboom R, Lucas C. Preoperative quadriceps strength as a predictor for short-term functional outcome after total hip replacement. Arch Phys Med Rehabil 2011; 92(2):236-241.

(49) Schafer T, Krummenauer F, Mettelsiefen J, Kirschner S, Gunther KP. Social, educational, and occupational predictors of total hip replacement outcome. Osteoarthritis Cartilage 2010; 18(8):1036-1042.

(50) Graves CM, Otero JE, Gao Y, Goetz DD, Willenborg MD, Callaghan JJ. Patient reported allergies are a risk factor for poor outcomes in total hip and knee arthroplasty. J Arthroplasty 2014; 29(9 Suppl):147-149.

(51) Lavernia CJ, Villa JM, Iacobelli DA, Rossi MD. Vitamin D insufficiency in patients with THA: prevalence and effects on outcome. Clin Orthop Relat Res 2014; 472(2):681-686.

(52) Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40(5):373-383.

(53) van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care 2009; 47(6):626-633.

(54) McHorney CA, Ware JEJ, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31(3):247-263.

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both. *Comment:* *Tittle*: *Predictors of physical functioning after total hip arthroplasty: a systematic review.*	Page 1.
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. *Comment:* *We performed a structured summary including all mentioned aspects.*	Page 2.
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. *Comment:* *Rationale is described in detail in the Introduction.*	Page 3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). *Comment:* *This systematic review aims to evaluate the predictors of functional outcome after total hip arthroplasty.* *P = subjects with osteoarthritis* *I = total hip arhtroplasty* *C = -* *O = functional outcome after THA* (HHS, OHS, SF-36, LEFS, TUG, WOMAC) *S= systematic review*	
METHODS			

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. *Comment:* *The protocol can be retrieved electronically through the first author.*	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. *Comment:* *All follow-up lengths and languages were included. The study characteristics can be found at the selection criteria.*	Page 4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. *Comment:* *With the help of an independent medical librarian we conducted a literature search through four different databases: Web of Scienel; Cinahl; Embase and PubMed. This literature search was performed on June 23, 2015. The exact search strategy can be found in the methods chapter.*	Page 4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. *Comment:* *Details of the flowchart and entire search strategy are described in Figure 1*	Figure 1.
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). *Comment:* *Two of the authors (LvB and TP) first independently screened the titles and abstracts of all the articles, using the above mentioned selection criteria. Both reviewers screened the full-text articles of the articles found eligible in the first round. A third author (LDB) compared these results and in case of different opinions, consensus was reached. The study selection procedure is schematically presented in figure 1.*	Page 5
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. *Comment:* *One of the authors (LDB) extracted the data, double checked by a second author. The extracted data of all included studies are summarized in table 1. These results are categorized by predictor variable. From all the articles, the following information was extracted: (1) predictor variable; (2) author (3) year of publication; (4) level of evidence; (5)*	Page 5

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

		number of patients; (6) measurement tools that are used; (7) follow-up period; (8) significance level; (9) association between predictor variable and outcome measure; (10) predictor level of measurement.	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. *Comment:* *See methods for the complete search strategy.*	Page 4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. *Comment:* *Risk of bias and our attempt to reduce the risk of bias in the individual studies was described in both the Subjects and Methods section and in the Results.*	Page 4,5,7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). *Comment:* *Best-evidence synthesis were described in the Subjects and Methods section. As this systematic review was a qualitative synthesis of the available evidence. In view of the heterogeneity of the target population, the variability of study objectives and differences in methodological quality, a meta-analysis could not be performed.* *In the Results we described in detail our findings with regard to the predictors of functional outcome after THA. All predictors evaluated in literature are reported.*	Page 6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis. *Comment:* *The performed best-evidence syntheses were described in the Methods section.* *In the results and Table 2-7 we described in detail our findings of the predictors of functional outcome after THA. This systematic review was a qualitative synthesis of the available evidence, a meta-analysis could not be performed. The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up..*	

Section/topic	#	Checklist item	Reported on page #

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). *Comment:* *Risk of bias in the individual studies was determined by the GRADE approach and displayed in Table 8.*	Table 1
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. *Comment:* *Not applicable.*	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. *Comment:* *Figure 1 shows the flow of information through the different phases of the systematic review.*	Figure 1.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. *Comments:* *See Table 2-7 for all extracted data.*	Table 2-7
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). *Comment:* *See Table 1 for the GRADE rating scheme.*	Table 1
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. *Comment:* *Details about the individual studies are described in the Results and in detail in Table 2-7*	Table 2-7
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency. *Comment:* *The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up.*	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15).	Table 1

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

		Comment: *See Table 1 for the GRADE rating scheme.*	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). *Comment:* *Not applicable.*	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). *Comment:* *The main findings and their implications are described in the Discussion section..*	Page 20
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). *Comment:* *Limitations of the review were described in detail in the discussion section.*	Page 20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research. *Comment:* *We described that methodological well-conducted, randomized, controlled trials in larger groups of subjects with more equal distribution and extensive measurements methods are necessary to investigate the pain sensitivity and pain perception in obese subjects vs non-obese subjects.* *In addition we advised to study the unknown variables of influence to pain sensitivity and pain perception in obese subjects.*	Page 23
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. *Comment:* *This systematic review was performed without any funding and the authors have no disclosure of conflicts of interest.* *The authors have no disclosure of conflicts of interest regarding the systematic review.*	24

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

BMJ Open

Predictors of physical functioning after total hip arthroplasty: a systematic review.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010725.R1
Article Type:	Research
Date Submitted by the Author:	22-Feb-2016
Complete List of Authors:	Buirs, Leon; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery van Beers, Loes; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Scholtes, Vanessa; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Pastoors, Tom; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Sprague, Sheila; McMaster University, Department of Clinical Epidemiology and Biostatistics Poolman, Rudolf; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery
Primary Subject Heading:	Surgery
Secondary Subject Heading:	Rheumatology
Keywords:	total hip, arthroplasty, functional outcome, systematic review, predictors

Predictors of physical functioning after total hip arthroplasty: a systematic review

Buirs LD[1], MD; Van Beers LWAH[1], MSc; Scholtes VAB[1], PHD; Pastoors T[1]; Sprague S[2], PHD; Poolman RW[1], MD PHD

[1]Department of Orthopaedic Surgery, Onze Lieve Vrouwe Gasthuis, Amsterdam, the Netherlands.

[2]Division of Orthopedic Surgery, Centre for Evidence-Based Orthopedics, Department of Surgery, McMaster University, Hamilton, Ontario, Canada.

Objective: The objective of this systematic review of the literature was to identify the predictors of functional outcome after total hip arthroplasty.

Method: A systematic literature search in Web of Science, Cinahl, Embase and PubMed was conducted on June 23, 2015. The articles were selected based upon their quality, relevance and measurement of the predictive factor. The level of evidence of all studies was determined using the GRADE rating scheme.

Results: The initial search resulted in 1092 citations. After application of the inclusion and exclusion criteria, 33 articles met our eligibility criteria and were graded. All of the included studies were classified as level of evidence low(11), moderate(17) or high(5). Of the included studies, 18 evaluated body mass index (BMI), 17 evaluated pre-operative physical function, 15 evaluated age, 15 evaluated gender, and 13 evaluated co-morbidity. There was strong evidence suggesting an association between BMI, age, comorbidity, pre-operative physical function, and mental health with functional outcome after THA. There was weak evidence suggesting an association between quadriceps strength and education with functional outcome after THA. The evidence was inconsistent for associations with gender and socio-economic status and functional outcome following THA. We found limited evidence suggesting that alcohol consumption, vitamin-D insufficiency and allergies were predictors of functional outcome following THA.

Conclusion: We have identified multiple predictors of functional outcome after THA, which will enable general practitioners and orthopedic surgeons to better predict the improvement in physical

functioning for their THA patients. They can use this information to provide patient specific advice regarding the referral for THA and the expected outcomes after THA. Further research with consistent measurement tools, outcomes, and duration of follow-up across studies is needed to confirm the influence of these factors.

Keywords: total hip arthroplasty - functional outcome - systematic review - predictors

Strengths and limitations of this study

- We have carried out a comprehensive and robust systematic review in accordance with the PRISMA guidelines.
- We included a range of patient related predictors and did not limit ourselves to the most common predictors. This led to a broad overview of predictors evaluated.
- We screened a large number of literature sources, and all reviewing and data extraction was carried out by one author (LDB) and double checked by a second author (LWAHB).
- Because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up we could not apply a meta-analysis.
- Some predictors as quadriceps strength, education, socioeconomic status and alcohol consumption are reported only a few times and therefore conclusions cannot be reached.

Introduction

Total hip arthroplasty (THA) is a surgical procedure performed to reduce pain and improve function in patients with osteoarthritis (OA) of the hip. According to the Agency for Healthcare Research and Quality, more than 305.000 total hip replacements are performed each year in the United States[1].

Following THA, the majority of patients experience reductions in pain, improvements in function, and better health related quality of life [2].However, not all patients achieve the same level of functional improvement after THA. Specifically, greater than 30 percent of patients undergoing THA report moderate to severe activity limitations two years post THA [3]. It is unclear which factors are associated with these limitations in function [4;5].

In the last decade, many studies have been published investigating the predictors of functional outcome after THA. "Young et al. published a systematic review on this topic in 1998. Since then considerable research has been published on predictors of functional outcome which justifies a new systematic review." [6]. Therefore, we conducted a systematic review to review predictors of mid-term and long-term functional outcome after THA.

Methods

Registration

This systematic review is registered at Prospero (http://www.crd.york.ac.uk/PROSPERO/) with registry number CRD42015016929.

Selection criteria

Studies that met the following criteria were included in our review: (1) included patients undergoing a THA; (2) included physical functioning was an outcome measure; (3) had at least one variable that was considered as a predictor of physical functioning and (4) was written In English. We did not select a time period.

Search strategy

With the guidance of an independent medical librarian we conducted a literature search through four medical databases: Web of Science; Cinahl; Embase ,and PubMed. This literature search was performed on June 23, 2015. In Web of Science we used the following search terms: TOPIC: (total hip arthroplasty) *AND* TOPIC: (predictor*). In Cinahl we searched for: (MM "Arthroplasty, Replacement, Hip") AND predictor*. In Embase we searched for: exp hip arthroplasty/ exp prediction/ or exp predictor variable/ exp prognosis/ or exp functional assessment/ or exp treatment outcome/ or exp daily life activity/. In PubMed we searched for ("Arthroplasty, Replacement, Hip"[Majr] OR "Hip Prosthesis"[Majr]) AND (predictor* OR risk Factor* OR risk assessment OR predictive value of tests OR prognostic factor* OR Prognostic*) AND (hoos OR "hip disability and osteoarthritis outcome score " OR womac OR "Western Ontario and McMaster Universities Arthritis Index" OR "harris hip score" OR HHS OR SF-12 OR short form 12 OR SF 36 OR "short form 36" OR trendelenburg OR TUG OR "timed up and go" OR "oxford hip score" OR "IOWA hip score" OR "functional recovery score" OR FRS OR AFI OR "hospital for special surgery" OR aaos OR "charnley hip score" OR HSS OR LEGS OR "mayo clinical hip score"). The results of these four different searches were combined in Reference Manager and duplicates were removed.

Study selection

Two of the authors (LWAHB and TP) independently screened the titles and abstracts of all the articles, using the above mentioned selection criteria. Both reviewers screened the full-text articles of the articles found eligible in the first round. A third author (LDB) compared these results and in

case of different opinions, consensus was reached. The study selection procedure is schematically presented in Figure 1.

Data extraction

One of the authors (LDB) extracted the data, double checked by a second author (LWAHB). From each articles, the following information was extracted: (1) predictor variable; (2) author (3) year of publication; (4) level of evidence; (5) number of patients; (6) measurement tools that are used; (7) follow-up period; (8) significance level; (9) association between predictor variable and outcome measure; (10) predictor level of measurement (Table 1) . The results were categorized by predictor variable.

Methodological quality assessment

The level of evidence of all studies was determined by one of the authors (LDB) with the GRADE rating scheme (http://www.gradeworkinggroup.org).

Measurement tools

We aimed to include all predictors mentioned in previous studies, and will not limit ourselves to the most common predictors. The measurement tools most used to define the functional outcome are the Harris Hip Score (HHS)[7], Oxford Hip Score (OHS)[8;9], Short Form-36 (SF-36)[10], LEFS (Lower Extremity Functional Scale)[11], Timed Up and Go test (TUG)[12;13] and the Western Ontario and McMaster Universities OA Index (WOMAC)[14]. We will use all mentioned measurement tools as outcome in this study.

Best evidence synthesis

A follow-up period up to 24 months was considered as 'short term' and a follow-up period of more than 24 months was considered as 'long term'. Results were divided into four categories of evidence: Strong evidence: at least 60% of the studies, with a minimum of 3 studies, describe the same

significant (p-value <.05) association. Weak evidence: a) only 2 studies describe the same significant association; b)3 studies describe the same association of which 2 are significant and 1 is not significant (p-value >.05). Limited evidence: a) only 1 study available; b) more studies available of which none found a significant association. Inconsistent evidence: all other scenarios. No conclusions can be drawn in this literature review when no or inconsistent evidence is available.

This systematic review conforms to the PRISMA statement[15].

Results

Selection and methodological quality

The initial search resulted in 1092 citations (Figure 1) and 33 articles met our eligibility criteria. The articles included were designated as level of evidence low(11), moderate(17) or high(5)(Table 1)

Table 1 Methodological quality of included studies

Study	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	GRADE
Kessler, 2007	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Villalobos, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Nankaku, 2013	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Slaven, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Moran, 2005	observational study	n.a.	not serious	not serious	not serious	strong association	moderate
Stevens, 2012	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Wang, 2010	observational study	not serious	not serious	not serious	not serious	none	moderate
Dowsey, 2010	observational study	serious	not serious	not serious	not serious	strong association	low
Judge, 2014	observational study	not serious	not serious	not serious	not serious	very strong association	high
Bergschmidt, 2010	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Jones, 2012	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Smith, 2012	observational study	not serious	not serious	serious	not serious	strong association	moderate
Judge, 2013	observational study	not serious	not serious	not serious	not serious	very strong association	high
Bischoff, 2004	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Gandhi, 2010	observational study	serious	not serious	not serious	not serious	none	low
Nilsdotter, 2003	observational study	not serious	serious	not serious	not serious	strong association	low
Davis, 2012	observational study	not serious	not serious	not serious	not serious	very strong association	high
Hamilton, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Quintana, 2009	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Nilsdotter, 2002	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Dowsey, 2014	observational study	not serious	not serious	not serious	not serious	very strong association	high
Lavernia, 2010	observational study	serious	not serious	not serious	not serious	strong association	low

Mahomed, 2002	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Vogl, 2014	observational study	not serious	serious	not serious	not serious	n.a.	low
Clement, 2011	observational study	not serious	not serious	not serious	not serious	very strong association	high
Johansson, 2010	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Fortin, 2002	observational study	not serious	not serious	not serious	serious	strong association	low
Badura-Brzoza, 2009	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Holstege, 2011	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Schafer, 2010	observational study	not serious	not serious	not serious	n.a.	strong association	low
Graves, 2014	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Lavernia, 2012	observational study	not serious	not serious	not serious	n.a.	none	low
Lavernia, 2013	observational study	not serious	not serious	not serious	not serious	strong association	moderate

GRADE: Grading recommendations assessment development and evaluation

High: true effect lies close to the estimate of the effect

Moderate: true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low: true effect may be substantially different from the estimate of effect

Very low: true effect is likely to be substantially different from the estimate of effect

Measures of functional outcome

Multiple outcome measures were used across these studies including the HHS, OHS, SF-36 PF, LEFS, TUG and the WOMAC score. The follow-up period ranged from 3 to 72 months with an average of 18 (SD17) months.

Predictive factors of functional outcome

BMI

Eighteen studies evaluated BMI as a potential predictor of functional outcome after THA[16-33](Table 2).

TABLE 2: Articles reporting BMI as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	0.49	no	*cont (BMI)*
Villalobos, 2012	low	63	SF-12 PCS	ST (3m)	**0.004***	pos	*dich (1: BMI >28 2: BMI ≤28)*
			WOMAC		**0.041***	pos	
			HHS		0.793*	no	
			OHS		0.428*	no	
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	0.06	no	*cont (BMI)*
Slaven, 2012	low	40	LEFS	ST (6m)	**n.a.**	neg	*dich (1: BMI >34 2: BMI ≤34)*
Moran, 2005	moderate	749	HHS	ST (6m)	**0.02**	neg	*cont (BMI)*
				ST (18m)	**0.001**	neg	
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.001**	neg	*cont (BMI)*
Wang, 2010	moderate	97	WOMAC	ST (12m)	0.11	no	*cont (BMI)*
Dowsey, 2010	low	471	HHS	ST (12m)	**<0.01**	neg	*cat (3) (1: BMI <30 2: BMI 30-39 3: BMI >30*
			SF-12 PCS		**0.05**	neg	
Dowsey, 2014	high	835	HHS	ST (12m)	**<0.0001**	neg	*cont (BMI)*
Judge, 2014	high	4413	OHS	ST (12m)	**0.003**	neg	*cat (5) (1: BMI 18.5-25 2: BMI 25-30 3: BMI 30-*

							35 4: BMI 35-40 5: BMI >40
Bergschmidt, 2010	moderate	100	HHS	ST (24m)	**0.007**	neg	*cat (3) (1: BMI <26 2: BMI 29-29 3: BMI >29*
Jones, 2012	moderate	231	WOMAC	ST (6m)	**0.001**	neg	*dich (1: BMI >35 2: BMI ≤35)*
				LT (36m)	no	no	
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.01**	neg	*cont (BMI)*
Judge, 2013	high	1431	OHS	LT (36m)	**<0.001**	neg	*cont (BMI)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	neg	*cont (BMI)*
Gandhi, 2010	low	636	WOMAC	LT (39m)	0.06	no	*cont (BMI)*
Nilsdotter, 2003	low	211	WOMAC PF	LT (42m)	**0.03**	neg	*cont (BMI)*
Davis,2011	high	1617	HHS	LT (60m)	**<0.001**	neg	*cont (BMI)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dig= dichotomous;**
cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index;
LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term;
LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative

A total of 14432 patients were included in all articles concerning the impact of BMI, with a mean follow-up time of 22 months. The applied levels of measurement of BMI were continuous, dichotomous or categorical.

The measurements tools used to determine the functional outcome are the WOMAC score, HHS, OHS, LEFS, SF-12 PF and the ambulatory status. The classification of a high BMI ranged from >28kg/m2 to >35kg/m2.

Of the 18 studies, 13 found a significant association [16-18;21;22;24;26-30;32;33]. Twelve studies evaluated the short-term functional outcome of which eight studies[16;19;21;24;27;29;32;33] found a significant negative association and one article a significant positive association[30]. Of the seven studies evaluating the long-term functional outcome, five articles found a significant negative association [17;18;22;26;28]. All studies were designated as level of evidence low(5), moderate(9) or high(4).

Because more than 60% of the studies report a significant negative association, there is strong evidence of a negative association between BMI and short-term and long-term functional outcome after THA. Note that these results were consistent, if not better, when we only considered the studies with high or moderate level of evidence according to GRADE.

Age

Fifteen studies evaluated age as a possible predictor of functional outcome after THA [16;17;20;22;23;25-29;31;33-36](Table 3).

TABLE 3: Articles reporting age as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	**0.03**	neg	*cont (age)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	yes	neg	*dich (1: age >67.5 2: age ≤67.5)*
Slaven, 2012	low	40	LEFS	ST (6m)	no	no	*dich (1: age >68.5 2: age ≤68.5)*
Hamilton, 2012	low	1410	OHS	ST (6m)	x	no	*cont (age)*
			SF-12	ST (12m)	x	no	
Quintana, 2009	moderate	788	WOMAC PF	ST (6m)	0.41	no	*dich*
				ST (24m)	**0.001**	neg	*(1: age >70 2: age ≤70)*
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.01**	neg	*cont (age)*
Wang, 2010	moderate	97	WOMAC	ST (12m)	no	no	*cont (age)*
Dowsey, 2014	high	835	HHS	ST (12m)	**<0.0001**	neg	*cont (age)*
			SF-12 PCS		**0.003**	neg	
Nilsdotter, 2002	moderate	148	WOMAC PF	ST (12m)	**0.004**	neg	*dich*
			SF-36		**0.002**	neg	*(1: age >72 2: age ≤72)*

Bergschmidt, 2010	moderate	100	HHS	ST (12m)	>0.097	no	*cat (3) (1: age <60 2: age 60-69 3: age >69*
			WOMAC		>0.097	no	
			SF-12		>0.097	no	
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	x	no	*dich (1: age >75 2: age ≤75)*
Judge, 2013	high	1431	OHS	LT (36m)	**n.a.**	neg	*cat (3) (1: age <50 2: age 50-60 3: age >60*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	neg	*cont (age)*
Nilsdotter, 2003	low	211	WOMAC PF	LT (43m)	**0.002**	neg	*cont (age)*
Gandhi, 2010	low	636	WOMAC	LT (39m)	**<0.05**	neg	*cont (age)*
			SF-36		**<0.05**		

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dig= dichotomous;**
cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index;
LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term;
LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative

A total of 9234 patients were included in all studies that identified age as a possible predictor, with a mean follow-up time of 19 months. The applied levels of measurement of age were continuous, dichotomous or categorical.

The measurements tools used to determine the functional outcome are the WOMAC score, HHS, OHS, SF-36 PF, SF-12 PF and the ambulatory status. Different classifications of high age were used, ranging from >60 to >75.

Of the 15 studies, 10 found a significant association[20;22;23;25;26;28;29;33;35;36]. Ten studies evaluated the short-term functional outcome of which six studies found a significant negative association [23;25;29;33;35;36]. The other four studies did not found a significant association. Of the six studies evaluating the long-term functional outcome, five studies found a significant negative association [20;22;28;35;36]. All studies were designated as level of evidence low(4), moderate(9) or high(2).

Because more than 60% of the studies report a significant negative association, there is strong evidence of a negative association between high age and short-term and long-term functional outcome after THA. Note that these results were consistent, if not better, when we only considered the studies with high or moderate level of evidence according to GRADE.

Gender

Fifteen studies evaluated gender as a possible predictor of functional outcome after THA [16;17;20;21;23;25-29;31;33;35-37](Table 4).

TABLE 4: Articles reporting gender as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	n.a.	no	*dich (1: men 2: woman)*
Nilsdotter, 2002	moderate	148	WOMAC SF-36	ST (3m) ST (12m)	0.7	no	*dich (1: men 2: woman)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	0.10	no	*dich (1: men 2: woman)*
Slaven, 2012	low	40	LEFS	ST (6m)	**0.039**	pos, woman	*dich (1: men 2: woman)*
Quintana, 2009	moderate	788	SF-36 PF	ST (6m) ST (24m)	**n.a.** n.a.	pos, men no	*dich (1: men 2: woman)*
Bergschmidt, 2010	moderate	100	HHS	ST (12m)	n.a.	no	*dich (1: men 2: woman)*
Stevens, 2012	low	653	WOMAC	ST (12m)	**0.002**	pos, men	*dich (1: men 2: woman)*
Dowsey, 2014	high	835	HHS	ST (12m)	0.06	no	*dich (1: men 2: woman)*
Wang, 2010	moderate	97	WOMAC	ST (16.8m)	**0.0001**	pos, men	*dich (1: men 2: woman)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	no	no	*dich (1: men 2: woman)*
Jones, 2012	moderate	231	WOMAC	LT (36m)	0.118	no	*dich (1: men 2: woman)*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	pos, men	*dich (1: men 2: woman)*
Gandhi, 2010	low	636	WOMAC	LT (39m)	no	no	*dich*

			SF-36 PF		**<0.05**	pos, woman	*(1: men 2: woman)*
Lavernia, 2010	low	532	WOMAC PF	LT (42m)	**<0.001***	pos, woman	*dich (1: men 2: woman)*
Nilsdotter,2003	low	211	WOMAC PF	LT (66m)	0.37	no	*dich (1: men 2: woman)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; dig= dichotomous;**
SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index;
LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term;
LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative

A total of 7156 patients were included in all studies that evaluated gender as a possible predictor, with a mean follow-up time of 23.3 months. The measurements tools used to determine the functional outcome included the WOMAC score HHS, LEFS, SF-36 and the ambulatory status.

Of the 15 studies, seven found a statistically significant association between preoperative physical function and functional outcome [20;27-29;31;36;37]. Nine studies evaluated the short-term functional outcome of which four studies found a significant association [27;29;31;36]. Six studies evaluated the long-term functional outcome of which three found a significant association [20;28;37]. All studies were designated as level of evidence low(5), moderate(9) or high(1).

In four of the seven studies with a significant outcome, being male predicted a better outcome [28;29;31;36] whereas three studies reported being female as a predictor of better functional outcome [20;27;37]. This demonstrates inconsistent evidence for an association between gender and functional outcome after THA.

Pre-operative physical function

Seventeen studies evaluated pre-operative physical function as a possible predictor of functional outcome after THA [16;22;24-28;31;33-36;38-42](Table 5).

TABLE 5: Articles reporting pre-operative physical function as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Quintana, 2009	moderate	788	WOMAC PF SF-36 PF	ST (6m)	**<0.001**	yes	*cont (WOMAC + SF-36*
Slaven, 2012	low	40	TUG	ST (6m)	n.a.	no	*dich (successful/unsuccessful)*
Mahomed, 2002	moderate	103	WOMAC PF+P SF36 PF	ST (6m)	**<0.05** **<0.05**	yes	*cont (WOMAC + SF-36)*
Hamilton, 2012	low	1410	OHS SF-12	ST (6m) ST (12m)	yes	yes	*cont (OHS)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	**n.a.**	yes	*dich (TUG score 10)*
Vogl, 2014	low	281	WOMAC	ST (6m)	**n.a.**	yes	*cont (WOMAC)*
Bergschmidt, 2010	moderate	100	WOMAC SF-36	ST (12m)	<0.022 0.003	yes	*cat (3)* *1: HHS ,<48 2: HHS 48-59* *3: HHS >59*
Clement, 2010	high	1312	OHS	ST (12m)	**0.001***	yes	*cont (OHS)*

			SF-12				
Johansson, 2010	moderate	75	HHS WOMAC SF-36	ST (12m)	**≤0.006** **<0.001** **≤0.005**	yes yes yes	*cat (3)* *1: HHS ,<45 2: HHS 45-55 3: HHS >55*
Nilsdotter, 2002	moderate	148	WOMAC SF-36	ST (12m)	**<0.0001**	yes	*dich* *low quartile vs high quartile WOMAC*
Dowsey, 2014	high	835	HHS	ST (12m)	**<0.0001**	yes	*cont (HHS)*
Wang, 2010	moderate	97	WOMAC	ST (16.8m)	**0.0001**	yes	*cont (WOMAC PF)*
Moran, 2005	moderate	749	HHS	ST (18m)	**n.a.**	yes	*cont*
Fortin, 2002	low	222	WOMAC SF-36	ST (24m)	**n.a.** **n.a.**	yes yes	*dig (1: high WOMAC 2. low WOMAC*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	yes	*cont (HHS)*
Nilsdotter,2003	low	211	WOMAC PF	LT (42m)	**0.007**	yes	*dich* *low quartile vs high quartile SF-36 PF*
Judge, 2013	high	1431	OHS	LT (60m)	**<0.001**	yes	*cont (OHS)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dig= dichotomous;**
cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index;
LoE: level of evidence; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index;
HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative

A total of 9689 patients were included in all studies that evaluated pre-operative physical function, with a mean follow-up time of 16 months. The applied levels of measurement of preoperative physical function were continuous, dichotomous or categorical.

The WOMAC score [14] is the measurement tool most used to determine the preoperative physical function [16;26;31;35;36;38-40;42]. Other measurement tools used are the HHS, TUG, OHS, SF-36, SF-12 and the ambulatory status.

Of the 17 studies, 16 found a statistically significant correlation between pre-operative physical function and functional outcome. Fourteen studies evaluated the short-term outcome of which 13 reported a significant association. Three studies evaluated the long-term outcome; all three found a significant association. The only study that did not report a significant association, is a study with a small patient group that used the TUG to determine the preoperative physical function[27]. All studies were designated as level of evidence low(5), moderate(9) or high(3).

As more than 60% of the studies report a significant negative association, there is strong evidence of a short-term and long-term association between the preoperative physical function and the functional outcome after THA.

Comorbidity

Thirteen studies evaluated comorbidity as a possible predictor of functional outcome after THA (Table 6).

TABLE 6: Articles reporting comorbidity status as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Quintana, 2009	moderate	788	WOMAC PF	ST (6m)	n.a.	no	*cat(3) 1: 0 comorb 2: 1-2 comorb 3: >2 comorb*
			SF-36 PF		n.a.		
Mahomed, 2002	moderate	103	WOMAC PF+P	ST (6m)	**<0.05**	neg	*cont (number of comorbidities)*
Moran, 2005	moderate	749	HHS	ST (6m)	**n.a.**	neg	*dich (presence of coronary heart disease and previous thromo-embolism)*
				ST (18m)			
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.01**	neg	*cat(3) 1: 0 comorb 2: 1-2 comorb 3: >2 comorb*
Clement, 2010	high	1312	OHS	ST (12m)	**0.01**	neg	*cont (number of comorbidities)*
			SF-12				
Dowsey, 2014	high	835	HHS	ST (12m)	**0.0001**	neg	*cont (age adjusted CCI)*
Wang,	moderate	97	WOMAC	ST (16.8m)	**0.0246**	neg	*dich*

2010							*(1: >0 comorbidities 2: 0 comorbidities)*
Jones, 2012	moderate	231	WOMAC	LT (36m)	**0.012**	neg	*dig (1; 0 cardiac diseases 2: >0 cardiac diseases)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	neg	*dich (1; >2 chron diseases 2. 0-1 chronic diseases)*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	neg	*cont (asa grade)*
Gandhi, 2010	low	636	WOMAC SF-36 PF	LT (39m)	**<0.05**	neg	*cont (number of comorbidities)*
Nilsdotter, 2003	low	211	WOMAC PF	LT (42m)	0.08	no	*dich (1: >1 comorbidities 2: 0-1 comorbidities)*
Judge, 2013	high	1431	OHS	LT (60m)	**0.001**	neg	*cont (number of comorbidities)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dig= dichotomous;**
cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index;
LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term;
LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative

A total of 9363 patients were included in all studies that evaluated comorbidity as a possible predictor, with a mean follow-up time of 23.3 months. The applied levels of measurement of preoperative status were continuous, dichotomous or categorical.

The measurements tools used to determine the functional outcome are the WOMAC score, HHS, LEFS, SF-36 and the ambulatory status. Most studies used the number of comorbidities or asa grade as predictor of functional outcome. Other studies used the presence of a specific comorbidity as a predictor like cardiac diseases, coronary heart diseases and thrombo-embolisms.

Of the 13 studies, 11 found a significant negative association [17;20;21;24;26;28;29;31-33;36;38;41]. Seven studies evaluated the short-term outcome of which six reported a significant negative association[21;21;22;24;29;31;33;38;41]. Six studies evaluated the long-term outcome, of which five found a significant negative association[17;20-22;28]. All articles were designated as level of evidence low(2), moderate(8) or high(3).

Because more than 60% of the studies report a significant negative association, there is strong evidence of a negative association between comorbidities and short-term and long-term functional outcome after THA.

Other predictors

The predictors that are evaluated in five studies or less will be described below and will be displayed all together in Table 7.

TABLE 7: All predictors that are evaluated in five articles or less

predictor	Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Mental health	Badura-Brzoza, 2009	moderate	102	SF-36 PCS	ST (6m)	**0.005**	neg	*cont (anxiety as a trait)*
	Quintana, 2009	moderate	788	SF-36 PF WOMAC P	ST (6m) ST (24m)	**<0.001** **0.002**	yes	*cont (SF-36 MH score)*
	Dowsey, 2014	high	835	HSS	ST (12m)	**<0.0001**	yes	*cont (SF-12 MH score)*
	Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	yes	*dich (1: >60 pts on the SF-36 MH score 2: ≤60pts on SF-36 MH score)*
	Judge, 2013	high	916	OHS	ST (12m) LT (60m)	**0.045**	yes	*cont (SF-36 MH score)*
Alcohol consumption	Bischoff, 2004	moderate	914	WOMAC PF	LT (36m)	**n.a.**	no	*dich (1: >1 alcoholic drinks per day 2: 0-1 alcoholic drinks per day)*
	Lavernia, 2012	low	191	WOMAC	LT (72m)	**n.a.**	no	*cat (3) (1: non drinkers 2: occasional drinkers 3: moderate drinkers)*
Quadriceps strength	Holstege, 2011	moderate	55	WOMAC PF	ST (3m)	**0.004**	pos	*cont (knee extensor strength)*
	Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	**n.a.**	pos	*dich (1: > 1.25 N m/kg 2: ≤1.25 m/kg*

								knee extensor strength)
Education	Schafer, 2010	low	1007	WOMAC	ST (6m)	**n.a.**	pos	*dich (1; >12 years school 2: <9 years school)*
	Mahomed, 2002	moderate	103	WOMAC PF+P	ST (6m)	**0.007**	pos	*cont (level of education)*
	Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	pos	*dich (1:college education 2:less than college education)*
Socio eco-nomic status	Dowsey, 2014	high	835	HHS	LT (12m)	0.63	no	*cont (SES score)*
Allergies	Graves, 2014	moderate	459	WOMAC PF SF-36 PCS	ST (6.5m)	**0.04** **0.0002**	neg	*dich (>3 allergies)*
Vitamin-D insufficiency	Lavernia, 2013	moderate	60	HHS WOMAC	ST (11m)	**0.002** 0.478	neg	*dich (25-hydroxyvitamin-D3) (1; >30 ng/ml 2: <30 ng/ml)*

*** All significant results are bold; cont= continuous; dig= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function;**
WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index; LoE: level of evidence; LEFS=lower extremity functional scale;
OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score;
N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative

Five studies evaluated *mental health* as a possible predictor of functional outcome after THA, with a total of 3563 patients [17;22;33;36;43]. All four studies evaluating the short-term functional outcome found a significant positive association [22;33;36;43]. Both studies that evaluated the long-term outcome found a significant positive association. Because more than 60% of the studies report a significant positive association, there is strong evidence of an association between good mental health and better short-term physical function outcome after THA. Because only two studies evaluated the long-term outcome, this evidence is weak.

Two studies evaluated *alcohol consumption* as a predictor of functional outcome [17;44]. None of them found a significant result and therefore none show evidence of an association. The two studies evaluating quadriceps strength as a possible predictor [25;45] looked at the short-term functional outcome and both found a significant association. Therefore the evidence for an association is weak.

All three studies that evaluated education as a possible predictor, found a significant association [17;38;46]. Two studies evaluated the short-term outcome and both found a significant association [38;46]. Bischoff et al evaluated the long-term effect and found a significant association [17]. All three studies used the WOMAC score to measure the functional outcome. These results show weak evidence for a short-term association, and incomplete evidence for a long-term association.

Dowsey et al. reported *socio-economic status* as a predictor, using the socio-economic status score (SES) as measurement tool[33]. They did not find a significant result and therefore show limited evidence of an association.

The influence of having more than 3 *allergies* on the short-term functional outcome was reported by Graves et al[47]. Those patients had diminished improvements on SF-36 PCS and WOMAC scores, 6,5 months after THA. This result shows limited evidence of an association between having more than 3 allergies and functional outcome.

Lavernia et al evaluated vitamin-D insufficiency as a predictor of functional outcome after THA[48]. A preoperative 25-hydroxyvitamin-D3 plasma level of under 30 ng/ml, predicted a worse HHS 11 months postoperative. Because no other studies evaluated vitamin-D insufficiency as a possible predictor, this result shows limited evidence of an association.

Discussion

In this systematic literature review we sought to provide a clear overview of a range of patient related predictors of functional outcome after THA.

Key findings

Our review found strong evidence of an association between BMI, age, comorbidity, preoperative physical function, and mental health with functional outcome after THA. Weak evidence was found for the predictors quadriceps strength and education. Inconsistent evidence was found for the predictors gender and socio-economic status. Limited evidence was found for the predictors alcohol consumption, vitamin-D insufficiency, and allergies.

In our review thirteen studies found a significant negative association between BMI and functional outcome after THA.. A prior review of Young et al [6] found the same significant negative association. Although the review of Young et al and our current review come to the same conclusion, the clinical impact of this outcome is still questionable. A large study by Judge et al. showed a small significant correlation between a high BMI and a worse functional outcome, but conclude that the total improvement in function outweighs the small lack of improvement caused by a high BMI[32].

Although our review shows strong evidence of an association between BMI and functional outcome, different classifications of high BMI were used. Because of these different classifications, it is difficult to define a specific BMI that predicts who will do well after THA. We could not conduct a meta-analysis since different classifications of BMI are used and there was heterogeneity in outcome instruments. Therefore future research on the impact of BMI should use clearly defined outcomes that are consistent across studies.

In our review eight of the 14 studies found an association between higher age and poorer functional outcome, therefore age is an important factor predicting functional outcome. Some articles used a linear regression analysis for age. When looking at age, it is not only interesting to see the effect of high age, but also of low age. Therefore linear regression analysis might not be the best statistical analysis with variables as age or BMI. For future research on the impact of age on functional outcome

after THA, more consistent outcomes must be used across studies. There is no consensus among studies about what specific age limit is recommended for THA. This current review shows inconclusive evidence of an association between gender and functional outcome because six out of 14 studies found a statistically significant result.

Three studies reported being female led to a better functional outcome [20;27;37]. The other four significant articles found the opposite result where being male has a positive association with functional outcome after THA [28;29;31;36]. The results are very contradictory and the differences may be partially attributable to confounding factors.

The pre-operative physical function was found to be a conclusive predictor. With the exception of one study reporting the timed up and go test as an outcome, better pre-operative physical function was consistently associated with better long-term physical function [27]. This might be due to the fact that they used the TUG score as measurement tool[27]. The WOMAC score was the measurement tool most used to define the pre-operative status (9 times)[16;26;31;35;36;38-40;42]. Other pre-operative measurement tools that were good predictors of functional outcome are the HHS score, OHS score, SF-12 PF score, SF-36 PF score, and ambulatory status.

Of the 13 studies that evaluated comorbidity as a possible predictor of functional outcome, 11 found a significant negative association [17;20-22;24;28;29;31;33;36;38;41]. Comorbidity can be measured in several ways, for example: the number of comorbidities, the presence of a specific comorbidity, the Charlson index [49] and the Elixhauser comorbidity measure [50]. Comorbidities can affect the true functional outcome after THA but can also affect the score on the measurement tool. For example: if a patient is unable to walk to the grocery store after a THA due to a lung disease, his functional outcome score will be lower despite a possible good functioning total hip. Except for one article, all studies found a significant negative effect. And therefore having comorbidities can be seen as a predictor of negative functional outcome.

All five studies that evaluated mental health as a predictor of functional outcome found a statistically significant positive association. Four of these studies used SF-36 MH [51] as measurement tool to measure mental health[17;22;33;36]. These results show strong evidence of a positive association between mental health and short-term functional outcome after THA. The two studies reporting quadriceps strength as a predictor had both small sample sizes which can affect the external validity of the studies[25;45]. Therefore this evidence is weak and more research must be done on the effect of quadriceps strength.

Three studies evaluated education as predictor of functional outcome. Mahomed et al[38] and Bischoff et al[17] used the level of school education as a predictor, and Schafer et al[46] used years of education as a predictor. Because education is in part a surrogate of socioeconomic status, this might also indicate that low socioeconomic status is a factor associated with poor functional outcome. Dowsey et al however did not find a correlation between socioeconomic status and functional outcome[33]. Future research is needed on various components of socioeconomic status to specify the impact on functional outcome. Because only one study evaluated allergies[47] and vitamin-D insufficiency[48] as possible predictors of functional outcome, no conclusions can be drawn.

Previous systematic reviews

The previous systematic review of Young et al. concluded that important research remains to be done to examine the magnitude and interaction of patient factors on the outcome of THA[6]. The review of Young et al. used only one database (MEDLINE) and is more than 15 years old. Young et al. also looked at implant survivorship. In our systematic review we used multiple databases (Web of Science; Cinahl; Embase and PubMed) and reported only patient related predictors evaluated in literature.

Strengths and Limitations

We included a range of patient related predictors and did not limit ourselves to the most common predictors. This led to a broad overview of predictors evaluated. The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up. Not all studies used in this review adjusted their outcomes for potential confounders. Therefore some outcomes may be due to confounding factors. A limitation of our review is that we looked at functional outcome without including pain. Some patients will not see an improvement in their function after THA, but will lose the hip related pain. For this reason especially people with a high BMI and older age can benefit from THA, without improving the function of the hip. Some predictors such as quadriceps strength, education, socioeconomic status and alcohol consumption are reported only a few times and therefore conclusions cannot be reached. More research in large datasets is needed to draw definitive conclusions on these predictors.

Implications for practice

Our review provides a clear overview of the current literature on the predictors for physical functioning after THA. Orthopedic surgeons and general practitioners can use this information to predict the improvement in physical functioning for their patients and it enables them to provide patient specific advice on THA surgery.

Implications for future research

In the future, we suggest studies that evaluate possible predictors of functional outcome after THA to use equal measurement tools, outcomes and durations of follow-up. In that way a meta-analysis can be applied and the influence of these factors can be specified.

Conclusion

This review shows that several patient related characteristics can predict the functional outcome after THA. It shows strong evidence of an association between BMI, age, comorbidity, preoperative physical function and mental health with functional outcome after THA. Weak evidence suggested that quadriceps strength and education were predictive of functional outcomes after THA. Inconsistent evidence was found for the predictors gender and socio-economic status. Alcohol consumption, vitamin-D insufficiency and allergies showed limited evidence predicting functional outcome after THA. Understanding predictors will help orthopedic surgeons and general practitioners predict the outcomes in physical functioning after THA they can use this information to provide patient specific advice and target care for patients with THA. Large clinical trials are necessary to confirm these findings.

Acknowledgement

We would like to thank the medical librarian Bert Berenschot at Onze Lieve Vrouwe Gasthuis for his help with the literature search.

Contributors: LDB, LWAHB, VABS, TP, SS, RWP, made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data. LDB, LWAHB, VABS, SS and RWP have been involved in drafting the manuscript or revising it critically for important intellectual content. All authors read and approved the final manuscript.

Funding: The study was not funded and none of the contributors recieved funding for this study.

Competing interests:"All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; L.W.A.H. Beers reports grants from Link / Lima, grants from Stryker, outside the submitted work. Dr. Sprague reports employment/salary from McMaster

University, employment/salary from Global Research Solutions, outside the submitted work. Dr. Scholtes reports grants from Link / Lima, grants from Stryker, grants from NuVasive, grants from Zonmw, grants from Achmea and grants from Tornier, outside the submitted work. Dr. Poolman reports grants from Link / Lima, grants from Stryker, grants from NuVasive, grants from Zonmw, grants from Achmea and grants from Tornier, outside the submitted work; no other relationships or activities that could appear to have influenced the submitted work**."**

Ethical approval: Not required.

Transparency: The lead author (LDB) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned and registered have been explained.

Data sharing: No additional data are available.

Reference List

(1) *Bernstein J, Derman P. Dramatic increase in total knee replacement utilization rates cannot be fully explained by a disproportionate increase among younger patients. Orthopedics 2014; 37(7):e656-e659.*

(2) *Judge A, Cooper C, Williams S, Dreinhoefer K, Dieppe P. Patient-reported outcomes one year after primary hip replacement in a European Collaborative Cohort. Arthritis Care Res (Hoboken) 2010; 62(4):480-488.*

(3) *Singh JA, Lewallen DG. Predictors of Activity Limitation and Dependence on Walking Aids After Primary Total Hip Arthroplasty. Journal of the American Geriatrics Society 2010; 58(12):2387-2393.*

(4) *Williams O, Fitzpatrick R, Hajat S, Reeves BC, Stimpson A, Morris RW et al. Mortality, morbidity, and 1-year outcomes of primary elective total hip arthroplasty. J Arthroplasty 2002; 17(2):165-171.*

(5) *MacWilliam CH, Yood MU, Verner JJ, McCarthy BD, Ward RE. Patient-related risk factors that predict poor outcome after total hip replacement. Health Serv Res 1996; 31(5):623-638.*

(6) *Young NL, Cheah D, Waddell JP, Wright JG. Patient characteristics that affect the outcome of total hip arthroplasty: A review. Canadian Journal of Surgery 1998; 41(3):188-195.*

(7) *Nilsdotter A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. Arthritis Care Res (Hoboken) 2011; 63 Suppl 11:S200-S207.*

(8) *Gosens T, Hoefnagels NHM, de Vet RCW, Dhert WJA, van Langelaan EJ, Bulstra SK et al. The "Oxford Heup Score": the translation and validation of a questionnaire into Dutch to evaluate the results of total hip arthroplasty. Acta Orthop 2005; 76(2):204-211.*

(9) *Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996; 78(2):185-190.*

(10) Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992; 305(6846):160-164.

(11) Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. Phys Ther 1999; 79(4):371-383.

(12) Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39(2):142-148.

(13) Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the "get-up and go" test. Arch Phys Med Rehabil 1986; 67(6):387-389.

(14) Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988; 15(12):1833-1840.

(15) Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009; 151(4):264-9, W64.

(16) Bergschmidt P, Bader R, Finze S, Tokar I, Kundt G, Mittelmeier W. Impact of preoperative variables on the functional and radiological outcome of an uncemented femoral stem: a prospective two-year follow-up. Hip Int 2010; 20(2):187-197.

(17) Bischoff-Ferrari HA, Lingard EA, Losina E, Baron JA, Roos EM, Phillips CB et al. Psychosocial and geriatric correlates of functional status after total hip replacement. Arthritis & Rheumatism-Arthritis Care & Research 2004; 51(5):829-835.

(18) Davis AM, Wood AM, Keenan AC, Brenkel IJ, Ballantyne JA. Does body mass index affect clinical outcome post-operatively and at five years after primary unilateral total hip replacement performed for osteoarthritis? A multivariate analysis of prospective data. J Bone Joint Surg Br 2011; 93(9):1178-1182.

(19) Dowsey MM, Liew D, Stoney JD, Choong PFM. The impact of obesity on weight change and outcomes at 12 months in patients undergoing total hip arthroplasty. Medical Journal of Australia 2010; 193(1):17-21.

(20) Gandhi R, Dhotar H, Davey JR, Mahomed NN. Predicting the Longer-term Outcomes of Total Hip Replacement. J RHEUMATOL 2010; 37(12):2573-2577.

(21) Jones CA, Cox V, Jhangri GS, Suarez-Almazor ME. Delineating the impact of obesity and its relationship on recovery after total joint arthroplasties. Osteoarthritis Cartilage 2012; 20(6):511-518.

(22) Judge A, Arden NK, Batra RN, Thomas G, Beard D, Javaid MK et al. The association of patient characteristics and surgical variables on symptoms of pain and function over 5 years following primary hip-replacement surgery: a prospective cohort study. Bmj Open 2013; 3(3).

(23) Kessler S, Kafer W. Overweight and obesity: two predictors for worse early outcome in total hip replacement? Obesity (Silver Spring) 2007; 15(11):2840-2845.

(24) *Moran M, Walmsley P, Gray A, Brenkel IJ. Does body mass index affect the early outcome of primary total hip arthroplasty? J Arthroplasty 2005; 20(7):866-869.*

(25) *Nankaku M, Tsuboyama T, Akiyama H, Kakinoki R, Fujita Y, Nishimura J et al. Preoperative prediction of ambulatory status at 6 months after total hip arthroplasty. Phys Ther 2013; 93(1):88-93.*

(26) *Nilsdotter AK, Petersson IF, Roos EM, Lohmander LS. Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study. Annals of the Rheumatic Diseases 2003; 62(10):923-930.*

(27) *Slaven EJ. Prediction of Functional Outcome at Six Months Following Total Hip Arthroplasty. PHYS THER 2012; 92(11):1386-1394.*

(28) *Smith GH, Johnson S, Ballantyne JA, Dunstan E, Brenkel IJ. Predictors of excellent early outcome after total hip arthroplasty. J Orthop Surg Res 2012; 7:13.*

(29) *Stevens M, Paans N, Wagenmakers R, van Beveren J, van Raay JJAM, van der Meer K et al. The Influence of Overweight/Obesity on Patient-Perceived Physical Functioning and Health-Related Quality of Life After Primary Total Hip Arthroplasty. Obesity Surgery 2012; 22(4):523-529.*

(30) *Villalobos PA, Navarro-Espigares JL, Hernandez-Torres E, Martinez-Montes JL, Villalobos M, Arroyo-Morales M. Body Mass Index as Predictor of Health-Related Quality-of-Life Changes After Total Hip Arthroplasty: A Cross-Over Study. Journal of Arthroplasty 2013; 28(4):666-670.*

(31) *Wang W, Morrison TA, Geller JA, Yoon RS, Macaulay W. Predicting short-term outcome of primary total hip arthroplasty:a prospective multivariate regression analysis of 12 independent factors. J Arthroplasty 2010; 25(6):858-864.*

(32) *Judge A, Batra RN, Thomas GE, Beard D, Javaid MK, Murray DW et al. Body mass index is not a clinically meaningful predictor of patient reported outcomes of primary hip replacement surgery: prospective cohort study. Osteoarthritis Cartilage 2014; 22(3):431-439.*

(33) *Dowsey MM, Nikpour M, Choong PFM. Outcomes following large joint arthroplasty: does socio-economic status matter? BMC Musculoskelet Disord 2014; 15:148.*

(34) *Hamilton D, Henderson GR, Gaston P, MacDonald D, Howie C, Simpson AH. Comparative outcomes of total hip and knee arthroplasty: a prospective cohort study. Postgrad Med J 2012; 88(1045):627-631.*

(35) *Nilsdotter AK, Lohmander LS. Age and waiting time as predictors of outcome after total hip replacement for osteoarthritis. Rheumatology (Oxford) 2002; 41(11):1261-1267.*

(36) *Quintana JM, Escobar A, Aguirre U, Lafuente I, Arenaza JC. Predictors of Health-related Quality-of-life Change after Total Hip Arthroplasty. Clinical Orthopaedics and Related Research 2009; 467(11):2886-2894.*

(37) *Lavernia CJ, Alcerro JC, Contreras JS, Rossi MD. Patient perceived outcomes after primary hip arthroplasty: does gender matter? Clin Orthop Relat Res 2011; 469(2):348-354.*

(38) *Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. J RHEUMATOL 2002; 29(6):1273-1279.*

(39) *Johansson HR, Bergschmidt P, Skripitz R, Finze S, Bader R, Mittelmeier W. Impact of preoperative function on early postoperative outcome after total hip arthroplasty. J Orthop Surg (Hong Kong) 2010; 18(1):6-10.*

(40) *Fortin PR, Penrod JR, Clarke AE, St Pierre Y, Joseph L, Belisle P et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002; 46(12):3327-3330.*

(41) *Clement ND, Muzammil A, MacDonald D, Howie CR, Biant LC. Socioeconomic status affects the early outcome of total hip replacement. J Bone Joint Surg Br 2011; 93(4):464-469.*

(42) *Vogl M, Wilkesmann R, Lausmann C, Hunger M, Plotz W. The impact of preoperative patient characteristics on health states after total hip replacement and related satisfaction thresholds: a cohort study. Health Qual Life Outcomes 2014; 12:108.*

(43) *Badura-Brzoza K, Zajac P, Kasperska-Zajac A, Brzoza Z, Matysiakiewicz J, Piegza M et al. Anxiety and depression and their influence on the quality of life after total hip replacement: preliminary report. International Journal of Psychiatry in Clinical Practice 2008; 12(4):280-284.*

(44) *Lavernia CJ, Villa JM, Contreras JS. Alcohol use in elective total hip arthroplasty: risk or benefit? Clin Orthop Relat Res 2013; 471(2):504-509.*

(45) *Holstege MS, Lindeboom R, Lucas C. Preoperative quadriceps strength as a predictor for short-term functional outcome after total hip replacement. Arch Phys Med Rehabil 2011; 92(2):236-241.*

(46) *Schafer T, Krummenauer F, Mettelsiefen J, Kirschner S, Gunther KP. Social, educational, and occupational predictors of total hip replacement outcome. Osteoarthritis Cartilage 2010; 18(8):1036-1042.*

(47) *Graves CM, Otero JE, Gao Y, Goetz DD, Willenborg MD, Callaghan JJ. Patient reported allergies are a risk factor for poor outcomes in total hip and knee arthroplasty. J Arthroplasty 2014; 29(9 Suppl):147-149.*

(48) *Lavernia CJ, Villa JM, Iacobelli DA, Rossi MD. Vitamin D insufficiency in patients with THA: prevalence and effects on outcome. Clin Orthop Relat Res 2014; 472(2):681-686.*

(49) *Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40(5):373-383.*

(50) *van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care 2009; 47(6):626-633.*

(51) *McHorney CA, Ware JEJ, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31(3):247-263.*

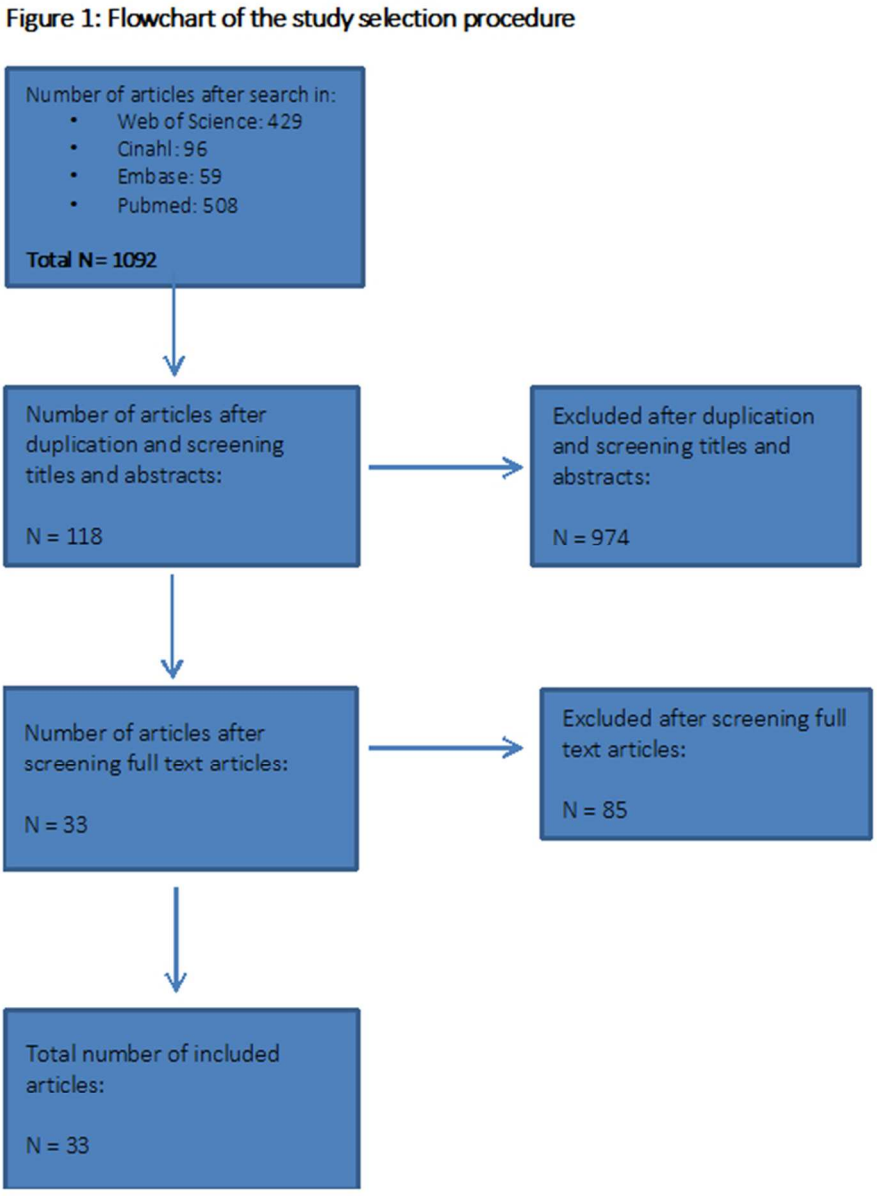


Figure 1: The study selection procedure.

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both. *Comment:* *Tittle*: *Predictors of physical functioning after total hip arthroplasty: a systematic review.*	Page 1.
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. *Comment:* *We performed a structured summary including all mentioned aspects.*	Page 2.
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. *Comment:* *Rationale is described in detail in the Introduction.*	Page 3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). *Comment:* *This systematic review aims to evaluate the predictors of functional outcome after total hip arthroplasty.* *P = subjects with osteoarthritis* *I = total hip arhtroplasty* *C = -* *O = functional outcome after THA* (HHS, OHS, SF-36, LEFS, TUG, WOMAC) *S= systematic review*	
METHODS			

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. *Comment:* *The protocol can be retrieved electronically through the first author.*	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. *Comment:* *All follow-up lengths and languages were included. The study characteristics can be found at the selection criteria.*	Page 4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. *Comment:* *With the help of an independent medical librarian we conducted a literature search through four different databases: Web of Scienel; Cinahl; Embase and PubMed. This literature search was performed on June 23, 2015. The exact search strategy can be found in the methods chapter.*	Page 4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. *Comment:* *Details of the flowchart and entire search strategy are described in Figure 1*	Figure 1.
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). *Comment:* *Two of the authors (LvB and TP) first independently screened the titles and abstracts of all the articles, using the above mentioned selection criteria. Both reviewers screened the full-text articles of the articles found eligible in the first round. A third author (LDB) compared these results and in case of different opinions, consensus was reached. The study selection procedure is schematically presented in figure 1.*	Page 5
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. *Comment:* *One of the authors (LDB) extracted the data, double checked by a second author. The extracted data of all included studies are summarized in table 1. These results are categorized by predictor variable. From all the articles, the following information was extracted: (1) predictor variable; (2) author (3) year of publication; (4) level of evidence; (5)*	Page 5

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

		number of patients; (6) measurement tools that are used; (7) follow-up period; (8) significance level; (9) association between predictor variable and outcome measure; (10) predictor level of measurement.	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. *Comment:* *See methods for the complete search strategy.*	Page 4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. *Comment:* *Risk of bias and our attempt to reduce the risk of bias in the individual studies was described in both the Subjects and Methods section and in the Results.*	Page 4,5,7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). *Comment:* *Best-evidence synthesis were described in the Subjects and Methods section. As this systematic review was a qualitative synthesis of the available evidence. In view of the heterogeneity of the target population, the variability of study objectives and differences in methodological quality, a meta-analysis could not be performed.* *In the Results we described in detail our findings with regard to the predictors of functional outcome after THA. All predictors evaluated in literature are reported.*	Page 6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis. *Comment:* *The performed best-evidence syntheses were described in the Methods section.* *In the results and Table 2-7 we described in detail our findings of the predictors of functional outcome after THA. This systematic review was a qualitative synthesis of the available evidence, a meta-analysis could not be performed. The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up..*	

Section/topic	#	Checklist item	Reported on page #

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). *Comment:* *Risk of bias in the individual studies was determined by the GRADE approach and displayed in Table 8.*	Table 1
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. *Comment:* *Not applicable.*	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. *Comment:* *Figure 1 shows the flow of information through the different phases of the systematic review.*	Figure 1.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. *Comments:* *See Table 2-7 for all extracted data.*	Table 2-7
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). *Comment:* *See Table 1 for the GRADE rating scheme.*	Table 1
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. *Comment:* *Details about the individual studies are described in the Results and in detail in Table 2-7*	Table 2-7
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency. *Comment:* *The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up.*	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15).	Table 1

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

		Comment: *See Table 1 for the GRADE rating scheme.*	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). *Comment:* *Not applicable.*	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). *Comment:* *The main findings and their implications are described in the Discussion section..*	Page 20
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). *Comment:* *Limitations of the review were described in detail in the discussion section.*	Page 20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research. *Comment:* *We described that methodological well-conducted, randomized, controlled trials in larger groups of subjects with more equal distribution and extensive measurements methods are necessary to investigate the pain sensitivity and pain perception in obese subjects vs non-obese subjects.* *In addition we advised to study the unknown variables of influence to pain sensitivity and pain perception in obese subjects.*	Page 23
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. *Comment:* *This systematic review was performed without any funding and the authors have no disclosure of conflicts of interest.* *The authors have no disclosure of conflicts of interest regarding the systematic review.*	24

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Page 2 of 2

BMJ Open

Predictors of physical functioning after total hip arthroplasty: a systematic review.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-010725.R2
Article Type:	Research
Date Submitted by the Author:	30-May-2016
Complete List of Authors:	Buirs, Leon; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery van Beers, Loes; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Scholtes, Vanessa; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Pastoors, Tom; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery Sprague, Sheila; McMaster University, Department of Clinical Epidemiology and Biostatistics Poolman, Rudolf; Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery
Primary Subject Heading:	Surgery
Secondary Subject Heading:	Rheumatology
Keywords:	total hip, arthroplasty, functional outcome, systematic review, predictors

Predictors of physical functioning after total hip arthroplasty: a systematic review

Buirs LD[1], MD; Van Beers LWAH[1], MSc; Scholtes VAB[1], PHD; Pastoors T[1]; Sprague S[2], PHD; Poolman RW[1], MD PHD

[1]Department of Orthopaedic Surgery, OLVG, Amsterdam, the Netherlands.

[2]Division of Orthopedic Surgery, Centre for Evidence-Based Orthopedics, Department of Surgery, McMaster University, Hamilton, Ontario, Canada.

Corresponding author:

Leon Delmore Buirs, MD

Onze Lieve Vrouwe Gasthuis, Orthopaedic Surgery

Oosterpark 9

1090 HM Amsterdam

Amsterdam, Noord-Holland, NL 1090HM

0031 650120521, ldbuirs@gmail.com

Abstract

Objective: The objective of this systematic review of the literature was to identify the predictors of functional outcome after total hip arthroplasty.

Method: A systematic literature search in Web of Science, CINAHL, Embase and PubMed was conducted on June 23, 2015. The articles were selected based upon their quality, relevance and measurement of the predictive factor. The level of evidence of all studies was determined using the GRADE rating scheme.

Results: The initial search resulted in 1,092 citations. After application of the inclusion and exclusion criteria, 33 articles met our eligibility criteria and were graded. Included studies were classified as level of evidence low (11), moderate (17) or high (5). Of the included studies, 18 evaluated body mass index (BMI), 17 evaluated pre-operative physical function, 15 evaluated age, 15 evaluated gender, and 13 evaluated co-morbidity. There was strong evidence suggesting an association between BMI, age, comorbidity, pre-operative physical function, and mental health with functional outcome after THA. There was weak evidence suggesting an association between quadriceps strength and education with functional outcome after THA. The evidence was inconsistent for associations with gender and socio-economic status and functional outcome following THA. We found limited evidence suggesting that alcohol consumption, vitamin-D insufficiency and allergies were predictors of functional outcome following THA.

Conclusion: We have identified multiple predictors of functional outcome after THA, which will enable general practitioners and orthopedic surgeons to better predict the improvement in physical functioning for their THA patients. They can use this information to provide patient specific advice regarding the referral for THA and the expected outcomes after THA. Further research with consistent measurement tools, outcomes, and duration of follow-up across studies is needed to confirm the influence of these factors.

Keywords: total hip arthroplasty - functional outcome - systematic review - predictors

Strengths and limitations of this study

- We have carried out a comprehensive and robust systematic review in accordance with the PRISMA guidelines.
- We included a range of patient related predictors and did not limit ourselves to the most common predictors. This led to a broad overview of predictors evaluated.
- We screened a large number of literature sources, and all reviewing and data extraction was carried out by one author (LDB) and double checked by a second author (LWAHB).
- Because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up we could not apply a meta-analysis.
- The predictors quadriceps strength, education, socioeconomic status and alcohol consumption were reported only a few times and therefore conclusions cannot be reached.

Introduction

Total hip arthroplasty (THA) is a surgical procedure performed to reduce pain and improve function in patients with osteoarthritis (OA) of the hip. According to the Agency for Healthcare Research and Quality, more than 305,000 total hip replacements are performed each year in the United States[1]. Following THA, the majority of patients experience reductions in pain, improvements in function, and better health related quality of life [2]. However, not all patients achieve the same level of functional improvement after THA. Specifically, greater than 30 percent of patients undergoing THA report moderate to severe activity limitations two years post THA [3]. It is unclear which factors are associated with these limitations in function [4;5].

In the last decade, many studies have been published investigating the predictors of functional outcome after THA. Young et al. published a systematic review on this topic in 1998. Since then considerable research has been published on predictors of functional outcome which justifies a new systematic review[6]. Therefore, we conducted a systematic review of predictors of mid-term and long-term functional outcome after THA.

Methods

Registration

This systematic review is registered at Prospero (http://www.crd.york.ac.uk/PROSPERO/) with registry number CRD42015016929.

Selection criteria

Studies that met the following criteria were included in our review: (1) included patients undergoing a THA; (2) included physical functioning was an outcome measure; (3) had at least one variable that was considered as a predictor of physical functioning; and (4) was written In English. We did not select a time period.

Search strategy

With the guidance of an independent medical librarian we conducted a literature search through four medical databases: Web of Science; CINAHL; Embase , and PubMed. This literature search was performed on June 23, 2015. In Web of Science we used the following search terms: TOPIC: (total hip arthroplasty) *AND* TOPIC: (predictor*). In CINAHL we searched for: (MM "Arthroplasty, Replacement, Hip") AND predictor*. In Embase we searched for: exp hip arthroplasty/ exp prediction/ or exp predictor variable/ exp prognosis/ or exp functional assessment/ or exp treatment outcome/ or exp daily life activity/. In PubMed we searched for ("Arthroplasty, Replacement, Hip"[Majr] OR "Hip Prosthesis"[Majr]) AND (predictor* OR risk Factor* OR risk assessment OR predictive value of tests OR prognostic factor* OR Prognostic*) AND (HOOS OR "hip disability and osteoarthritis outcome score " OR WOMAC OR "Western Ontario and McMaster Universities Arthritis Index" OR "Harris hip score" OR HHS OR SF-12 OR short form 12 OR SF 36 OR "short form 36" OR Trendelenburg OR TUG OR "timed up and go" OR "Oxford hip score" OR "IOWA hip score" OR "Functional recovery score" OR FRS OR AFI OR "Hospital for special surgery" OR AAOS OR "Charnley hip score" OR HSS OR LEGS OR "Mayo clinical hip score"). The results of these four different searches were combined in Reference Manager and duplicates were removed.

Study selection

Two of the authors (LWAHB and TP) independently screened the titles and abstracts of all the articles using the above mentioned selection criteria. Both reviewers screened the full-text articles of the articles found eligible in the first round. A third author (LDB) compared these results and in case of different opinions, consensus was reached. The study selection procedure is schematically presented in Figure 1.

Data extraction

One of the authors (LDB) extracted the data, double checked by a second author (LWAHB). From each articles, the following information was extracted: (1) predictor variable; (2) author; (3) year of publication; (4) level of evidence; (5) number of patients; (6) measurement tools used; (7) follow-up period; (8) significance level; (9) association between predictor variable and outcome measure; (10) predictor level of measurement (Table 1) . The results were categorized by predictor variable.

Methodological quality assessment

The level of evidence of all studies was determined by one of the authors (LDB) with the GRADE rating scheme (http://www.gradeworkinggroup.org).

Measurement tools

We aimed to include all predictors mentioned in previous studies, and not limit ourselves to the most common predictors. Widely used measurement tools to define functional outcome are the Harris Hip Score (HHS)[7], Oxford Hip Score (OHS)[8;9], Short Form-36 (SF-36)[10], LEFS (Lower Extremity Functional Scale)[11], Timed Up and Go test (TUG)[12;13] and the Western Ontario and McMaster Universities OA Index (WOMAC)[14]. We used all these measurement tools as outcome in this study.

Best evidence synthesis

A follow-up period up to 24 months was considered as 'short term' and a follow-up period of more than 24 months was considered as 'long term'. Results were divided into four categories of evidence: Strong evidence: at least 60% of the studies, with a minimum of 3 studies, describing the same significant (p-value <.05) association. Weak evidence: a) only 2 studies describe the same significant association; b) 3 studies describe the same association of which 2 are significant and 1 is not significant (p-value >.05). Limited evidence: a) only 1 study available; b) more studies available of which none found a significant association. Inconsistent evidence: all other scenarios[15]. No conclusions can be drawn in this literature review when no or inconsistent evidence is available.

This systematic review conforms to the PRISMA statement[16].

Results

Selection and methodological quality

The initial search resulted in 1,092 citations (Figure 1) and 33 articles met our eligibility criteria. The articles included were designated as level of evidence low (11), moderate (17) or high (5) (Table 1)

Table 1 Methodological quality of included studies

Study	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	GRADE
Kessler, 2007	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Villalobos, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Nankaku, 2013	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Slaven, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Moran, 2005	observational study	n.a.	not serious	not serious	not serious	strong association	moderate
Stevens, 2012	observational study	not serious	not serious	not serious	not serious	strong association	moderate

Wang, 2010	observational study	not serious	not serious	not serious	not serious	none	moderate
Dowsey, 2010	observational study	serious	not serious	not serious	not serious	strong association	low
Judge, 2014	observational study	not serious	not serious	not serious	not serious	very strong association	high
Bergschmidt, 2010	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Jones, 2012	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Smith, 2012	observational study	not serious	not serious	serious	not serious	strong association	moderate
Judge, 2013	observational study	not serious	not serious	not serious	not serious	very strong association	high
Bischoff, 2004	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Gandhi, 2010	observational study	serious	not serious	not serious	not serious	none	low
Nilsdotter, 2003	observational study	not serious	serious	not serious	not serious	strong association	low
Davis, 2012	observational study	not serious	not serious	not serious	not serious	very strong association	high
Hamilton, 2012	observational study	not serious	not serious	not serious	not serious	none	low
Quintana, 2009	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Nilsdotter, 2002	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Dowsey, 2014	observational study	not serious	not serious	not serious	not serious	very strong association	high
Lavernia, 2010	observational study	serious	not serious	not serious	not serious	strong association	low
Mahomed, 2002	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Vogl, 2014	observational study	not serious	serious	not serious	not serious	n.a.	low
Clement, 2011	observational study	not serious	not serious	not serious	not serious	very strong association	high
Johansson, 2010	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Fortin, 2002	observational study	not serious	not serious	not serious	serious	strong association	low
Badura-Brzoza, 2009	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Holstege, 2011	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Schafer, 2010	observational study	not serious	not serious	not serious	n.a.	strong association	low
Graves, 2014	observational study	not serious	not serious	not serious	not serious	strong association	moderate
Lavernia, 2012	observational study	not serious	not serious	not serious	n.a.	none	low

Lavernia, 2013	observational study	not serious	not serious	not serious	not serious	strong association	moderate

GRADE: Grading recommendations assessment development and evaluation

High: true effect lies close to the estimate of the effect

Moderate: true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

Low: true effect may be substantially different from the estimate of effect

Very low: true effect is likely to be substantially different from the estimate of effect

Measures of functional outcome

Multiple outcome measures were used across these studies including the HHS, OHS, SF-36 PF, LEFS, TUG and the WOMAC score. The follow-up period ranged from 3 to 72 months with an average of 18 (SD 17) months (Table 2).

TABLE 2: Characteristics of all included studies.

author, year, nr	age baseline	n of pts	female (n, %)	inclusion criteria	follow-up time	measurement tool
Badura-Brzoza, 2009, nr 42	61(54-75)	156	59, (58%)	prim THA, OA	6 months	SF-36 PF
Bergschmidt, 2010 nr 113	66(58- 74)	100	48 (50%)	prim THA, OA	6-12-24 months	HHS WOMAC SF-12
Bischoff, 2004, nr 51	73,1(65-93)	922	60%	OA, prim THA >65 y	3 years	WOMAC PF
Clement, 2011, nr 101	68,1(65-74)	1312	n.a	Prim OA, THR	12 months	OHS SF-12
Davis, 2011, nr 100	69(34-96)	1617	994	cemented THA	5 years	HHS SF-36 PF
Dowsey, 2010, nr 32	68,6/67/65,6	471	60,70%	prim THA OA	12 months	HHS SF-12 PF

Dowsey, 2014, nr 15	68,4	835	60,10%	prim THA	12 months	SF-12
Fortin, 2002, nr 145	65,7	222	59%	prim THA OA	2 years	WOMAC SF-36
Hamilton, 2012, nr 17	68,1	1410	57,20%	prim THA OA	6-12 months	OHS SF-12
Gandhi, 2010, nr 30	63,2(13.7)	636	53,50%	<18y, prim OA,	3.3 years	WOMAC SF-36 PF
Graves, 2014, 29	59,5	459	61,00%	THA OA	10,4 months	WOMAC SF-36
Holstege, 2011, nr 102	72,7(6,8)	55	41 (74,5)	THA OA	3 months	WOMAC PF
Johansson, 2010, nr 114	67(7)	75	36(48%)	THA OA	6-12-24 months	HHS WOMAC SF-36
Jones, 2012, nr 90	68,2(10,9)	231	138 (60%)	prim THA	6-36 months	WOMAC
Judge, 2013, nr 14	70	1431	887(62%)	OA	1-6 years	OHS
Kessler, 2007, nr 131	63,6	76	44,8 (59%)	THA OA	3 months	WOMAC
Lavernia, 2012, nr 73	70	60	48(80%)	prim THA	3-24 months	QWB-7 SF-36 PF WOMAC HHS
Lavernia, 2013, nr 81	62	191	70	prim THA	12 months	WOMAC SF-36
Lavernia, 2010, nr 103	61(15)	532	59%	THA	6-7 years	SF-26 HHS WOMAC
Mahomed, 2002, nr 149	66(9)	103	57(55%)	THA OA	6 months	WOMAC PF SF-36 pcs
Moran, 2005, nr 136	68	749	61%	prim THA	6, 18 months	HHS
Nankaku, 2013, nr 83	60,4	204	173	THA OA	6 months	ambulatory status
Nilsdotter, 2002, nr 147	71	148	83	THA OA	3-6-12 months	WOMAC SF-36
Nilsdotter, 2003, nr 52	71	211	106	prim THA	3,6 years	WOMAC PF
Quintana, 2009, nr 35	69,1	788	381(48%)	prim THA OA	6-24 months	SF-36 PF WOMAC
Schafer, 2010, nr 110	61	1007	55%	prim THA	6 months	WOMAC
Slaven, 2012, nr 15	68,2(8,2)	40	22 (55%)	prim THA	6 months.	LEFS
Smith, 2012, nr 92	68,5(9,9)	1683	n.a	prim THA	3 years	HHS
Stevens, 2012, nr 22	70,3(8,2)	653	74,20%	prim THA, OA	52,4 weeks	WOMAC
Villalobos, 2012, nr 80	62,39(13,6)	63	35(55,55%)	prim THA	3 months	HHS OHS WOMAC SF-12 PF
Vogle, 2014, nr 108	68	321	58%	prim THA	6 months	WOMAC
Wang, 2010, nr 107	61,65	97	62,40%	OA/osteonecrosis	3-12-24 months	WOMAC

SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THA= total hip arthroplasty; PF= physical function; HHS= Harris Hip Score; N of pts= number of patients;

Predictive factors of functional outcome

BMI

Eighteen studies evaluated BMI as a potential predictor of functional outcome after THA [17-34] (Table 3). A total of 14,432 patients were included in all articles concerning the impact of BMI, with a mean follow-up time of 22 months. The applied levels of measurement of BMI were continuous, dichotomous or categorical.

The measurement tools used to determine the functional outcome were WOMAC score, HHS, OHS, LEFS, SF-12 PF and the ambulatory status. The classification of a high BMI ranged from >28kg/m^2 to >35kg/m^2.

Of the 18 studies, 13 found a significant association [17-19;22;23;25;27-31;33;34]. Twelve studies evaluated the short-term functional outcome of which eight studies[17;20;22;25;28;30;33;34] found a significant negative association and one article a significant positive association[31]. Of the seven studies evaluating the long-term functional outcome, five articles found a significant negative association [18;19;23;27;29]. Studies were designated as level of evidence low (5), moderate (9) or high (4).

Because more than 60% of the studies report a significant negative association, there is strong evidence of a negative association between BMI and short-term and long-term functional outcome after THA. These results were consistent when we only considered the studies with high or moderate level of evidence according to GRADE.

TABLE 3: Studies reporting BMI as possible predictor of funtional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	0.49	no	*cont (BMI)*
Villalobos, 2012	low	63	SF-12 PCS WOMAC HHS OHS	ST (3m)	**0.004*** **0.041*** 0.793* 0.428*	pos pos no no	*dich* *(1: BMI >28 2: BMI ?28)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	0.06	no	*cont (BMI)*
Slaven, 2012	low	40	LEFS	ST (6m)	**n.a.**	neg	*dich* *(1: BMI >34 2: BMI ?34)*
Moran, 2005	moderate	749	HHS	ST (6m) ST (18m)	**0.02** **0.001**	neg neg	*cont (BMI)*
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.001**	neg	*cont (BMI)*
Wang, 2010	moderate	97	WOMAC	ST (12m)	0.11	no	*cont (BMI)*
Dowsey, 2010	low	471	HHS SF-12 PCS	ST (12m)	**<0.01** **0.05**	neg neg	*cat (3)* *(1: BMI <30 2: BMI 30-39 3: BMI >=40*
Dowsey, 2014	high	835	HHS	ST (12m)	**<0.0001**	neg	*cont (BMI)*
Judge, 2014	high	4413	OHS	ST (12m)	**0.003**	neg	*cat (5)* *(1: BMI 18.5-25 2: BMI 25-30 3: BMI 30-35 4: BMI 35-40 5: BMI >40*
Bergschmidt, 2010	moderate	100	HHS	ST (24m)	**0.007**	neg	*cat (3)* *(1: BMI <26 2: BMI 26-29 3: BMI >29*
Jones, 2012	moderate	231	WOMAC	ST (6m) LT (36m)	**0.001** no	neg no	*dich* *(1: BMI >35 2: BMI ?35)*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.01**	neg	*cont (BMI)*
Judge, 2013	high	1431	OHS	LT (36m)	**<0.001**	neg	*cont (BMI)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	neg	*cont (BMI)*
Gandhi, 2010	low	636	WOMAC	LT (39m)	0.06	no	*cont (BMI)*
Nilsdotter,2003	low	211	WOMAC PF	LT (42m)	**0.03**	neg	*cont (BMI)*
Davis,2011	high	1617	HHS	LT (60m)	**<0.001**	neg	*cont (BMI)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dich=dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Age

Fifteen studies evaluated age as a possible predictor of functional outcome after THA [17;18;21;23;24;26-30;32;34-37] (Table 4). A total of 9,234 patients were included in all studies that identified age as a possible predictor, with a mean follow-up time of 19 months. The applied levels of measurement of age were continuous, dichotomous or categorical.

The measurements tools used to determine the functional outcome were WOMAC score, HHS, OHS, SF-36 PF, SF-12 PF and the ambulatory status. Different classifications of greater age were used, ranging from >60 to >75 years.

Of the 15 studies, 10 found a significant association[21;23;24;26;27;29;30;34;36;37]. Ten studies evaluated the short-term functional outcome of which six studies found a significant negative association [24;26;30;34;36;37]. The other four studies did not find a significant association. Of the six studies evaluating the long-term functional outcome, five studies found a significant negative association [21;23;29;36;37]. Studies were designated as level of evidence low (4), moderate (9) or high (2).

Because more than 60% of the studies report a significant negative association, there is strong evidence of a negative association between high age and short-term and long-term functional outcome after THA. These results were consistent when we only considered the studies with high or moderate level of evidence according to GRADE.

TABLE 4: Studies reporting age as possible predictor of functional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	**0.03**	neg	*cont (age)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	yes	neg	*dich (1: age >67.5 2: age ?67.5)*
Slaven, 2012	low	40	LEFS	ST (6m)	no	no	*dich (1: age >68.5 2: age ?68.5)*
Hamilton, 2012	low	1410	OHS SF-12	ST (6m) ST (12m)	x x	no no	*cont (age)*
Quintana, 2009	moderate	788	WOMAC PF	ST (6m) ST (24m)	0.41 **0.001**	no neg	*dich (1: age >70 2: age ?70)*
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.01**	neg	*cont (age)*
Wang, 2010	moderate	97	WOMAC	ST (12m)	no	no	*cont (age)*
Dowsey, 2014	high	835	HHS SF-12 PCS	ST (12m)	**<0.0001** **0.003**	neg neg	*cont (age)*
Nilsdotter, 2002	moderate	148	WOMAC PF SF-36	ST (12m)	**0.004** **0.002**	neg neg	*dich (1: age >72 2: age ?72)*
Bergschmidt, 2010	moderate	100	HHS WOMAC SF-12	ST (12m)	>0.097 >0.097 >0.097	no no no	*cat (3) (1: age <60 2: age 60-69 3: age >69*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	x	no	*dich (1: age >75 2: age ?75)*
Judge, 2013	high	1431	OHS	LT (36m)	**n.a.**	neg	*cat (3) (1: age <50 2: age 50-60 3: age >60*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	neg	*cont (age)*
Nilsdotter, 2003	low	211	WOMAC PF	LT (43m)	**0.002**	neg	*cont (age)*
Gandhi, 2010	low	636	WOMAC SF-36	LT (39m)	**<0.05** **<0.05**	neg	*cont (age)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dich= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Gender

Fifteen studies evaluated gender as a possible predictor of functional outcome after THA [17;18;21;22;24;26-30;32;34;36-38](Table 5). A total of 7,156 patients were included in all studies that evaluated gender as a possible predictor, with a mean follow-up time of 23.3 months. The measurement tools used to determine the functional outcome included the WOMAC score HHS, LEFS, SF-36 and the ambulatory status.

Of the 15 studies, seven found a statistically significant association between preoperative physical function and functional outcome [21;28-30;32;37;38]. Nine studies evaluated the short-term functional outcome of which four studies found a significant association [28;30;32;37]. Six studies evaluated the long-term functional outcome of which three found a significant association [21;29;38]. All studies were designated as level of evidence low (5), moderate (9) or high (1).

In four of the seven studies with a significant outcome, being male predicted a better outcome [29;30;32;37] whereas three studies reported being female as a predictor of better functional outcome [21;28;38]. This demonstrates inconsistent evidence for an association between gender and functional outcome after THA.

TABLE 5: Studies reporting gender as possible predictor of functional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Kessler, 2007	moderate	76	WOMAC	ST (3m)	n.a.	no	*dich (1: men 2: woman)*
Nilsdotter, 2002	moderate	148	WOMAC SF-36	ST (3m) ST (12m)	0.7	no	*dich (1: men 2: woman)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	0.10	no	*dich (1: men 2: woman)*
Slaven, 2012	low	40	LEFS	ST (6m)	**0.039**	pos, woman	*dich (1: men 2: woman)*
Quintana, 2009	moderate	788	SF-36 PF	ST (6m) ST (24m)	**n.a.** n.a.	pos, men no	*dich (1: men 2: woman)*
Bergschmidt, 2010	moderate	100	HHS	ST (12m)	n.a.	no	*dich (1: men 2: woman)*
Stevens, 2012	low	653	WOMAC	ST (12m)	**0.002**	pos, men	*dich (1: men 2: woman)*
Dowsey, 2014	high	835	HHS	ST (12m)	0.06	no	*dich (1: men 2: woman)*
Wang, 2010	moderate	97	WOMAC	ST (16.8m)	**0.0001**	pos, men	*dich (1: men 2: woman)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	no	no	*dich (1: men 2: woman)*
Jones, 2012	moderate	231	WOMAC	LT (36m)	0.118	no	*dich (1: men 2: woman)*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	pos, men	*dich (1: men 2: woman)*
Gandhi, 2010	low	636	WOMAC SF-36 PF	LT (39m)	no **<0.05**	no pos, woman	*dich (1: men 2: woman)*
Lavernia, 2010	low	532	WOMAC PF	LT (42m)	**<0.001***	pos, woman	*dich (1: men 2: woman)*
Nilsdotter,2003	low	211	WOMAC PF	LT (66m)	0.37	no	*dich (1: men 2: woman)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; dich= dichotomous; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable pos= positive; neg= negative**

Pre-operative physical function

Seventeen studies evaluated pre-operative physical function as a possible predictor of functional outcome after THA [17;23;25-29;32;34-37;39-43] (Table 6). A total of 9,689 patients were included in all studies that evaluated pre-operative physical function, with a mean follow-up time of 16 months. The applied levels of measurement of preoperative physical function were continuous, dichotomous or categorical.

The WOMAC score [14] was the measurement tool most used to determine the preoperative physical function [17;27;32;36;37;39-41;43]. Other measurement tools used were the HHS, TUG, OHS, SF-36, SF-12 and the ambulatory status.

Of the 17 studies, 16 found a statistically significant correlation between pre-operative physical function and functional outcome. Fourteen studies evaluated the short-term outcome of which 13 reported a significant association. Three studies evaluated the long-term outcome; all three found a significant association. The only study that did not report a significant association, was a study with a small patient group that used the TUG to determine the preoperative physical function[28]. Studies were designated as level of evidence low (5), moderate (9) or high (3).

As more than 60% of the studies report a significant negative association, there is strong evidence of a short-term and long-term association between the preoperative physical function and the functional outcome after THA.

TABLE 6: Studies reporting pre-operative physical function as possible predictor of functional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Quintana, 2009	moderate	788	WOMAC PF SF-36 PF	ST (6m)	**<0.001**	yes	*cont (WOMAC + SF-36*
Slaven, 2012	low	40	TUG	ST (6m)	n.a.	no	*dich (successful/unsuccessful)*
Mahomed, 2002	moderate	103	WOMAC PF+P SF36 PF	ST (6m)	**<0.05** **<0.05**	yes	*cont (WOMAC + SF-36)*
Hamilton, 2012	low	1410	OHS SF-12	ST (6m) ST (12m)	yes	yes	*cont (OHS)*
Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	**n.a.**	yes	*dich (TUG score 10)*
Vogl, 2014	low	281	WOMAC	ST (6m)	**n.a.**	yes	*cont (WOMAC)*
Bergschmidt, 2010	moderate	100	WOMAC SF-36	ST (12m)	<0.022 0.003	yes	*cat (3) 1: HHS ,<48 2: HHS 48-59 3: HHS >59*
Clement, 2010	high	1312	OHS SF-12	ST (12m)	**0.001***	yes	*cont (OHS)*
Johansson, 2010	moderate	75	HHS WOMAC SF-36	ST (12m)	**?0.006** **<0.001** **?0.005**	yes yes yes	*cat (3) 1: HHS ,<45 2: HHS 45-55 3: HHS >55*
Nilsdotter, 2002	moderate	148	WOMAC SF-36	ST (12m)	**<0.0001**	yes	*dich low quartile vs high quartile WOMAC*
Dowsey, 2014	high	835	HHS	ST (12m)	**<0.0001**	yes	*cont (HHS)*
Wang, 2010	moderate	97	WOMAC	ST (16.8m)	**0.0001**	yes	*cont (WOMAC PF)*
Moran, 2005	moderate	749	HHS	ST (18m)	**n.a.**	yes	*cont*
Fortin, 2002	low	222	WOMAC SF-36	ST (24m)	**n.a.** **n.a.**	yes yes	*dig (1: high WOMAC 2. low WOMAC*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	yes	*cont (HHS)*
Nilsdotter,2003	low	211	WOMAC PF	LT (42m)	**0.007**	yes	*dich low quartile vs high quartile SF-36 PF*
Judge, 2013	high	1431	OHS	LT (60m)	**<0.001**	yes	*cont (OHS)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dich= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index; LoE: level of evidence; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Comorbidity

Thirteen studies evaluated comorbidity as a possible predictor of functional outcome after THA (Table 7). A total of 9,363 patients were included in all studies that evaluated comorbidity as a possible predictor, with a mean follow-up time of 23.3 months. The applied levels of measurement of preoperative status were continuous, dichotomous or categorical.

The measurements tools used to determine the functional outcome were the WOMAC score, HHS, LEFS, SF-36 and the ambulatory status. Most studies used the number of comorbidities or ASA grade as predictor of functional outcome. Other studies used the presence of a specific comorbidity as a predictor like cardiac disease, coronary heart disease and thrombo-embolism.

Of the 13 studies, 11 found a significant negative association [18;21;22;25;27;29;30;32-34;37;39;42]. Seven studies evaluated the short-term outcome of which six reported a significant negative association[22;22;23;25;30;32;34;39;42]. Six studies evaluated the long-term outcome, of which five found a significant negative association[18;21-23;29]. All articles were designated as level of evidence low (2), moderate (8) or high (3).

Because more than 60% of the studies report a significant negative association, there was strong evidence of a negative association between comorbidities and short-term and long-term functional outcome after THA.

TABLE 7: Studies reporting comorbidity status as possible predictor of functional outcome after THA.

Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Quintana, 2009	moderate	788	WOMAC PF SF-36 PF	ST (6m)	n.a. n.a.	no	*cat(3)* *1: 0 comorb 2: 1-2 comorb 3: >2 comorb*
Mahomed, 2002	moderate	103	WOMAC PF+P	ST (6m)	**<0.05**	neg	*cont* *(number of comorbidities)*
Moran, 2005	moderate	749	HHS	ST (6m) ST (18m)	**n.a.**	neg	*dich* *(presence of coronary heart disease and previous thromo-embolism)*
Stevens, 2012	moderate	653	WOMAC	ST (12m)	**0.01**	neg	*cat(3)* *1: 0 comorb 2: 1-2 comorb 3: >2 comorb*
Clement, 2010	high	1312	OHS SF-12	ST (12m)	**0.01**	neg	*cont* *(number of comorbidities)*
Dowsey, 2014	high	835	HHS	ST (12m)	**0.0001**	neg	*cont* *(age adjusted CCI)*
Wang, 2010	moderate	97	WOMAC	ST (16.8m)	**0.0246**	neg	*dich* *(1: >0 comorbidities 2: 0 comorbidities)*
Jones, 2012	moderate	231	WOMAC	LT (36m)	**0.012**	neg	*dich* *(1; 0 cardiac diseases* *2: >0 cardiac diseases)*
Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	neg	*dich* *(1; >2 chron diseases* *2. 0-1 chronic diseases)*
Smith, 2012	moderate	1683	HHS	LT (36m)	**<0.001**	neg	*cont* *(asa grade)*
Gandhi, 2010	low	636	WOMAC SF-36 PF	LT (39m)	**<0.05**	neg	*cont* *(number of comorbidities)*
Nilsdotter,2003	low	211	WOMAC PF	LT (42m)	0.08	no	*dich* *(1: >1 comorbidities 2: 0-1 comorbidities)*
Judge, 2013	high	1431	OHS	LT (60m)	**0.001**	neg	*cont* *(number of comorbidities)*

*** All significant results are bold; studies that used change in function as outcome are marked with *; cont= continuous; dich= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Other predictors

The predictors that were evaluated in five studies or less are displayed in Table 8.

Five studies evaluated *mental health* as a possible predictor of functional outcome after THA, with a total of 3,563 patients [18;23;34;37;44]. All four studies evaluating the short-term functional outcome found a significant positive association [23;34;37;44]. Both studies that evaluated the long-term outcome found a significant positive association. Because more than 60% of the studies report a significant positive association, there is strong evidence of an association between good mental health and better short-term physical function outcome after THA. Because only two studies evaluated the long-term outcome, this evidence is weak.

Two studies evaluated *alcohol consumption* as a predictor of functional outcome [18;45]. Neither of them found a significant result and therefore none show evidence of an association. The two studies evaluating quadriceps strength as a possible predictor [26;46] looked at the short-term functional outcome and both found a significant association. Therefore the evidence for an association is weak.

All three studies that evaluated educational level as a possible predictor, found a significant association [18;39;47]. Two studies evaluated the short-term outcome and both found a significant association [39;47]. One study evaluated the long-term effect and found a significant association [18]. All three studies used the WOMAC score to measure the functional outcome. These results show weak evidence for a short-term association, and incomplete evidence for a long-term association.

One study reported *socio-economic status* as a predictor, using the socio-economic status score (SES) as measurement tool[34]. They did not find a significant result and therefore show limited evidence of an association.

The influence of having more than 3 *allergies* on the short-term functional outcome was reported in one study [48]. Patients with allergies had diminished improvements on SF-36 PCS and

WOMAC scores 6.5 months after THA. There was limited evidence of an association between having more than 3 allergies and functional outcome.

Vitamin-D insufficiency as a predictor of functional outcome after THA was evaluated in one study [49]. A preoperative 25-hydroxyvitamin-D3 plasma level of under 30 ng/ml, predicted a worse HHS 11 months postoperative. Because no other studies evaluated vitamin-D insufficiency as a possible predictor, this result shows limited evidence of an association.

TABLE 8: All predictors that are evaluated in five studies or less

predictor	Author & Year	GRADE	N of pts	Measurement tool	FU-period (months)	Significance level (p-value)	Association	*Predictor Level of Measurement*
Mental health	Badura-Brzoza, 2009	moderate	102	SF-36 PCS	ST (6m)	**0.005**	neg	*cont (anxiety as a trait)*
	Quintana, 2009	moderate	788	SF-36 PF WOMAC P	ST (6m) ST (24m)	**<0.001** **0.002**	yes	*cont (SF-36 MH score)*
	Dowsey, 2014	high	835	HSS	ST (12m)	**<0.0001**	yes	*cont (SF-12 MH score)*
	Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	yes	*dich (1: >60 pts on the SF-36 MH score 2: ?60pts on SF-36 MH score)*
	Judge, 2013	high	916	OHS	ST (12m) LT (60m)	**0.045**	yes	*cont (SF-36 MH score)*
Alcohol consumption	Bischoff, 2004	moderate	914	WOMAC PF	LT (36m)	**n.a.**	no	*dich (1: >1 alcoholic drinks per day 2: 0-1 alcoholic drinks per day)*
	Lavernia, 2012	low	191	WOMAC	LT (72m)	**n.a.**	no	*cat (3) (1: nondrinkers 2: occasional drinkers 3: moderate drinkers)*
Quadriceps strength	Holstege, 2011	moderate	55	WOMAC PF	ST (3m)	**0.004**	pos	*cont (knee extensor strength)*
	Nankaku, 2013	moderate	204	ambulatory status	ST (6m)	**n.a.**	pos	*dich (1: > 1.25 N m/kg 2: ?1.25 m/kg knee extensor strength)*
Education	Schafer, 2010	low	1007	WOMAC	ST (6m)	**n.a.**	pos	*dich (1; >12 years school 2: <9 years school)*
	Mahomed, 2002	moderate	103	WOMAC PF+P	ST (6m)	**0.007**	pos	*cont (level of education)*
	Bischoff, 2004	moderate	922	WOMAC PF	LT (36m)	**n.a.**	pos	*dich (1:college education 2:less than college education)*
Socio eco-nomic status	Dowsey, 2014	high	835	HHS	LT (12m)	0.63	no	*cont (SES score)*
Allergies	Graves, 2014	moderate	459	WOMAC PF SF-36 PCS	ST (6.5m)	**0.04** **0.0002**	neg	*dich (>3 allergies)*
Vitamin-D insufficiency	Lavernia, 2013	moderate	60	HHS WOMAC	ST (11m)	**0.002** 0.478	neg	*dich (25-hydroxyvitamin-D3) (1; >30 ng/ml 2: <30 ng/ml)*

*** All significant results are bold; cont= continuous; dich= dichotomous; cat= categorical; SF-36 PF= Short Form 36 physical function; WOMAC= Western Ontario and Mcmaster universities Osteoarthritis Index; LoE: level of evidence; LEFS=lower extremity functional scale; OHS= Oxford Hip Score; THR= total hip replacement; ST=short-term; LT=long-term; BMI= Body Mass Index; HHS= Harris Hip Score; N of pts= number of patients; FU= follow-up; n.a.= not applicable; pos= positive; neg= negative**

Discussion

In this systematic literature review we sought to provide a clear overview of a range of patient related predictors of functional outcome after THA.

Key findings

Our review found strong evidence of an association of BMI, age, comorbidity, preoperative physical function, and mental health with functional outcome after THA. Weak evidence was found for the predictors quadriceps strength and education. Inconsistent evidence was found for the predictors gender and socio-economic status. Limited evidence was found for the predictors alcohol consumption, vitamin-D insufficiency, and allergies.

In our review, thirteen studies found a significant negative association between BMI and functional outcome after THA. A prior review of Young et al. [6] found the same significant negative association. Although the review of Young et al and our current review come to the same conclusion, the clinical impact of this outcome is still questionable. A large study by Judge et al., showed a small significant correlation between a high BMI and a worse functional outcome, but concluded that the total improvement in function outweighs the small lack of improvement caused by a high BMI[33].

Although our review shows strong evidence of an association between BMI and functional outcome, different classifications of high BMI were used. Because of these different classifications, it is difficult to define a specific BMI that predicts who will do well after THA. We could not conduct a meta-analysis since different classifications of BMI were used and there was heterogeneity in outcome instruments. Therefore future research on the impact of BMI should use clearly defined outcomes that are consistent across studies.

In our review, eight of the 14 studies found an association between higher age and poorer functional outcome, therefore age is an important factor predicting functional outcome. Some articles used a linear regression analysis for age. When looking at age, it is not only interesting to see the effect of

high age, but also of low age. Therefore linear regression analysis might not be the best statistical analysis with variables as age or BMI. There is no consensus among studies about what specific age limit is recommended for THA. This current review shows inconclusive evidence of an association between gender and functional outcome because six out of 14 studies found a statistically significant result.

Three studies reported being female led to a better functional outcome [21;28;38]. The other four significant articles found the opposite result where being male had a positive association with functional outcome after THA [29;30;32;37]. The results are contradictory and the differences may be attributable to confounding factors.

Pre-operative physical function was found to be a strong predictor of long-term functional outcome. With the exception of one study reporting the timed up and go test as an outcome, better pre-operative physical function was consistently associated with better long-term physical function [28]. This might be due to the use of TUG score as measurement tool[28]. The WOMAC score was the measurement tool most used to define the pre-operative status (9 times)[17;27;32;36;37;39-41;43]. Other pre-operative measurement tools that were good predictors of functional outcome were the HHS, OHS, SF-12 PF, SF-36 PF, and ambulatory status.

Of the 13 studies that evaluated comorbidity as a possible predictor of functional outcome, 11 found a significant negative association [18;21-23;25;29;30;32;34;37;39;42]. Comorbidity can be measured in several ways, for example: the number of comorbidities, the presence of a specific comorbidity, the Charlson index [50] and the Elixhauser comorbidity measure [51]. Comorbidities can affect the true functional outcome after THA but can also affect the score on the measurement tool. For example: if a patient is unable to walk to the grocery store after a THA due to a lung disease, his functional outcome score will be lower despite a possible good functioning total hip. Except for one article, all studies found a significant negative effect. Therefore having comorbidities can be seen as a predictor of negative functional outcome.

All five studies that evaluated mental health as a predictor of functional outcome found a statistically significant positive association. Four of these studies used SF-36 MH [52] as the measurement tool to measure mental health[18;23;34;37]. These results show strong evidence of a positive association between mental health and short-term functional outcome after THA. The two studies reporting quadriceps strength as a predictor had both small sample sizes which can affect the external validity of the studies[26;46]. Therefore this evidence is weak and more research must be done on the effect of quadriceps strength.

Three studies evaluated education as predictor of functional outcome. Mahomed et al[39] and Bischoff et al[18] used the level of school education as a predictor, and Schafer et al[47] used years of education as a predictor. Because education is in part a surrogate of socioeconomic status, this might also indicate that low socioeconomic status is a factor associated with poor functional outcome. Dowsey et al however did not find a correlation between socioeconomic status and functional outcome[34]. Future research is needed on various components of socioeconomic status to specify the impact on functional outcome. As only one study evaluated each of allergies[48] and vitamin-D insufficiency[49] as possible predictors of functional outcome, no conclusions can be drawn.

Previous systematic reviews

The previous systematic review of Young et al. concluded that important research remained to be done to examine the magnitude and interaction of patient factors on the outcome of THA[6]. The review of Young et al. used only one database (MEDLINE) and is more than 15 years old. Young et al. also looked at implant survivorship. In our systematic review we used multiple databases (Web of Science, CINAHL; Embase and PubMed) and reported only patient related predictors evaluated in the literature.

Strengths and Limitations

We included a range of patient related predictors and did not limit ourselves to the most common predictors. This led to a broad overview of predictors evaluated. The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tools, predictors and duration of follow-up. Not all studies used in this review adjusted their outcomes for potential confounders. Therefore some outcomes may be due to confounding factors. A limitation of our review is that we looked at functional outcome without including pain. Some patients will not see an improvement in their function after THA, but will lose the hip related pain. For this reason especially people with a high BMI and older age can benefit from THA, without improving the function of the hip. Some predictors such as quadriceps strength, education, socioeconomic status and alcohol consumption are reported only a few times and therefore conclusions cannot be reached. More research in large datasets is needed to draw definitive conclusions on these predictors.

Implications for practice

Our review provides a clear overview of the current literature on the predictors for physical functioning after THA. Orthopedic surgeons and general practitioners can use this information to predict the improvement in physical functioning for their patients and it enables them to provide patient specific advice on THA surgery.

Implications for future research

In the future, we suggest studies that evaluate possible predictors of functional outcome after THA to use similar measurement tools, outcomes and durations of follow-up. In that way a meta-analysis can be applied and the influence of these factors can be specified.

Conclusion

This review shows that several patient related characteristics can predict the functional outcome after THA. It shows strong evidence of an association between BMI, age, comorbidity, preoperative physical function and mental health with functional outcome after THA. Weak evidence suggested that quadriceps strength and education were predictive of functional outcomes after THA. Inconsistent evidence was found for the predictors gender and socio-economic status. Alcohol consumption, vitamin-D insufficiency and allergies showed limited evidence predicting functional outcome after THA. Understanding predictors will help orthopedic surgeons and general practitioners predict the outcomes in physical functioning after THA they can use this information to provide patient specific advice and target care for patients with THA. Further well-conducted cohort studies are necessary to confirm these findings.

Acknowledgement

We would like to thank the medical librarian Bert Berenschot at OLVG for his help with the literature search.

Contributors: LDB, LWAHB, VABS, TP, SS, RWP, made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data. LDB, LWAHB, VABS, SS and RWP

have been involved in drafting the manuscript or revising it critically for important intellectual content. All authors read and approved the final manuscript.

Funding: The study was not funded and none of the contributors received funding for this study.

Competing interests:"All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; L.W.A.H. Beers reports grants from Link / Lima, grants from Stryker, outside the submitted work. Dr. Sprague reports employment/salary from McMaster University, employment/salary from Global Research Solutions, outside the submitted work. Dr. Scholtes reports grants from Link / Lima, grants from Stryker, grants from NuVasive, grants from Zonmw, grants from Achmea and grants from Tornier, outside the submitted work. Dr. Poolman reports grants from Link / Lima, grants from Stryker, grants from NuVasive, grants from Zonmw, grants from Achmea and grants from Tornier, outside the submitted work; no other relationships or activities that could appear to have influenced the submitted work**."**

Ethical approval: Not required.

Transparency: The lead author (LDB) affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned and registered have been explained.

Data sharing: No additional data are available, though details on statistical analysis are available from the corresponding author (LDB) on request.

Reference List

(1) *Bernstein J, Derman P. Dramatic increase in total knee replacement utilization rates cannot be fully explained by a disproportionate increase among younger patients. Orthopedics 2014; 37(7):e656-e659.*

(2) *Judge A, Cooper C, Williams S, Dreinhoefer K, Dieppe P. Patient-reported outcomes one year after primary hip replacement in a European Collaborative Cohort. Arthritis Care Res (Hoboken) 2010; 62(4):480-488.*

(3) *Singh JA, Lewallen DG. Predictors of Activity Limitation and Dependence on Walking Aids After Primary Total Hip Arthroplasty. Journal of the American Geriatrics Society 2010; 58(12):2387-2393.*

(4) *Williams O, Fitzpatrick R, Hajat S, Reeves BC, Stimpson A, Morris RW et al. Mortality, morbidity, and 1-year outcomes of primary elective total hip arthroplasty. J Arthroplasty 2002; 17(2):165-171.*

(5) *MacWilliam CH, Yood MU, Verner JJ, McCarthy BD, Ward RE. Patient-related risk factors that predict poor outcome after total hip replacement. Health Serv Res 1996; 31(5):623-638.*

(6) Young NL, Cheah D, Waddell JP, Wright JG. Patient characteristics that affect the outcome of total hip arthroplasty: A review. Canadian Journal of Surgery 1998; 41(3):188-195.

(7) Nilsdotter A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. Arthritis Care Res (Hoboken) 2011; 63 Suppl 11:S200-S207.

(8) Gosens T, Hoefnagels NHM, de Vet RCW, Dhert WJA, van Langelaan EJ, Bulstra SK et al. The "Oxford Heup Score": the translation and validation of a questionnaire into Dutch to evaluate the results of total hip arthroplasty. Acta Orthop 2005; 76(2):204-211.

(9) Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996; 78(2):185-190.

(10) Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992; 305(6846):160-164.

(11) Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. Phys Ther 1999; 79(4):371-383.

(12) Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39(2):142-148.

(13) Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the "get-up and go" test. Arch Phys Med Rehabil 1986; 67(6):387-389.

(14) Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988; 15(12):1833-1840.

(15) Singh AS, Mulder C, Twisk JW, van MW, Chinapaw MJ. Tracking of childhood overweight into adulthood: a systematic review of the literature. Obes Rev 2008; 9(5):474-488.

(16) Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009; 151(4):264-9, W64.

(17) Bergschmidt P, Bader R, Finze S, Tokar I, Kundt G, Mittelmeier W. Impact of preoperative variables on the functional and radiological outcome of an uncemented femoral stem: a prospective two-year follow-up. Hip Int 2010; 20(2):187-197.

(18) Bischoff-Ferrari HA, Lingard EA, Losina E, Baron JA, Roos EM, Phillips CB et al. Psychosocial and geriatric correlates of functional status after total hip replacement. Arthritis & Rheumatism-Arthritis Care & Research 2004; 51(5):829-835.

(19) Davis AM, Wood AM, Keenan AC, Brenkel IJ, Ballantyne JA. Does body mass index affect clinical outcome post-operatively and at five years after primary unilateral total hip replacement performed for osteoarthritis? A multivariate analysis of prospective data. J Bone Joint Surg Br 2011; 93(9):1178-1182.

(20) Dowsey MM, Liew D, Stoney JD, Choong PFM. The impact of obesity on weight change and outcomes at 12 months in patients undergoing total hip arthroplasty. Medical Journal of Australia 2010; 193(1):17-21.

(21) Gandhi R, Dhotar H, Davey JR, Mahomed NN. Predicting the Longer-term Outcomes of Total Hip Replacement. J RHEUMATOL 2010; 37(12):2573-2577.

(22) Jones CA, Cox V, Jhangri GS, Suarez-Almazor ME. Delineating the impact of obesity and its relationship on recovery after total joint arthroplasties. Osteoarthritis Cartilage 2012; 20(6):511-518.

(23) Judge A, Arden NK, Batra RN, Thomas G, Beard D, Javaid MK et al. The association of patient characteristics and surgical variables on symptoms of pain and function over 5 years following primary hip-replacement surgery: a prospective cohort study. Bmj Open 2013; 3(3).

(24) Kessler S, Kafer W. Overweight and obesity: two predictors for worse early outcome in total hip replacement? Obesity (Silver Spring) 2007; 15(11):2840-2845.

(25) Moran M, Walmsley P, Gray A, Brenkel IJ. Does body mass index affect the early outcome of primary total hip arthroplasty? J Arthroplasty 2005; 20(7):866-869.

(26) Nankaku M, Tsuboyama T, Akiyama H, Kakinoki R, Fujita Y, Nishimura J et al. Preoperative prediction of ambulatory status at 6 months after total hip arthroplasty. Phys Ther 2013; 93(1):88-93.

(27) Nilsdotter AK, Petersson IF, Roos EM, Lohmander LS. Predictors of patient relevant outcome after total hip replacement for osteoarthritis: a prospective study. Annals of the Rheumatic Diseases 2003; 62(10):923-930.

(28) Slaven EJ. Prediction of Functional Outcome at Six Months Following Total Hip Arthroplasty. PHYS THER 2012; 92(11):1386-1394.

(29) Smith GH, Johnson S, Ballantyne JA, Dunstan E, Brenkel IJ. Predictors of excellent early outcome after total hip arthroplasty. J Orthop Surg Res 2012; 7:13.

(30) Stevens M, Paans N, Wagenmakers R, van Beveren J, van Raay JJAM, van der Meer K et al. The Influence of Overweight/Obesity on Patient-Perceived Physical Functioning and Health-Related Quality of Life After Primary Total Hip Arthroplasty. Obesity Surgery 2012; 22(4):523-529.

(31) Villalobos PA, Navarro-Espigares JL, Hernandez-Torres E, Martinez-Montes JL, Villalobos M, Arroyo-Morales M. Body Mass Index as Predictor of Health-Related Quality-of-Life Changes After Total Hip Arthroplasty: A Cross-Over Study. Journal of Arthroplasty 2013; 28(4):666-670.

(32) Wang W, Morrison TA, Geller JA, Yoon RS, Macaulay W. Predicting short-term outcome of primary total hip arthroplasty:a prospective multivariate regression analysis of 12 independent factors. J Arthroplasty 2010; 25(6):858-864.

(33) Judge A, Batra RN, Thomas GE, Beard D, Javaid MK, Murray DW et al. Body mass index is not a clinically meaningful predictor of patient reported outcomes of primary hip replacement surgery: prospective cohort study. Osteoarthritis Cartilage 2014; 22(3):431-439.

(34) Dowsey MM, Nikpour M, Choong PFM. Outcomes following large joint arthroplasty: does socio-economic status matter? BMC Musculoskelet Disord 2014; 15:148.

(35) Hamilton D, Henderson GR, Gaston P, MacDonald D, Howie C, Simpson AH. Comparative outcomes of total hip and knee arthroplasty: a prospective cohort study. Postgrad Med J 2012; 88(1045):627-631.

(36) Nilsdotter AK, Lohmander LS. Age and waiting time as predictors of outcome after total hip replacement for osteoarthritis. Rheumatology (Oxford) 2002; 41(11):1261-1267.

(37) Quintana JM, Escobar A, Aguirre U, Lafuente I, Arenaza JC. Predictors of Health-related Quality-of-life Change after Total Hip Arthroplasty. Clinical Orthopaedics and Related Research 2009; 467(11):2886-2894.

(38) Lavernia CJ, Alcerro JC, Contreras JS, Rossi MD. Patient perceived outcomes after primary hip arthroplasty: does gender matter? Clin Orthop Relat Res 2011; 469(2):348-354.

(39) Mahomed NN, Liang MH, Cook EF, Daltroy LH, Fortin PR, Fossel AH et al. The importance of patient expectations in predicting functional outcomes after total joint arthroplasty. J RHEUMATOL 2002; 29(6):1273-1279.

(40) Johansson HR, Bergschmidt P, Skripitz R, Finze S, Bader R, Mittelmeier W. Impact of preoperative function on early postoperative outcome after total hip arthroplasty. J Orthop Surg (Hong Kong) 2010; 18(1):6-10.

(41) Fortin PR, Penrod JR, Clarke AE, St Pierre Y, Joseph L, Belisle P et al. Timing of total joint replacement affects clinical outcomes among patients with osteoarthritis of the hip or knee. Arthritis Rheum 2002; 46(12):3327-3330.

(42) Clement ND, Muzammil A, MacDonald D, Howie CR, Biant LC. Socioeconomic status affects the early outcome of total hip replacement. J Bone Joint Surg Br 2011; 93(4):464-469.

(43) Vogl M, Wilkesmann R, Lausmann C, Hunger M, Plotz W. The impact of preoperative patient characteristics on health states after total hip replacement and related satisfaction thresholds: a cohort study. Health Qual Life Outcomes 2014; 12:108.

(44) Badura-Brzoza K, Zajac P, Kasperska-Zajac A, Brzoza Z, Matysiakiewicz J, Piegza M et al. Anxiety and depression and their influence on the quality of life after total hip replacement: preliminary report. International Journal of Psychiatry in Clinical Practice 2008; 12(4):280-284.

(45) Lavernia CJ, Villa JM, Contreras JS. Alcohol use in elective total hip arthroplasty: risk or benefit? Clin Orthop Relat Res 2013; 471(2):504-509.

(46) Holstege MS, Lindeboom R, Lucas C. Preoperative quadriceps strength as a predictor for short-term functional outcome after total hip replacement. Arch Phys Med Rehabil 2011; 92(2):236-241.

(47) Schafer T, Krummenauer F, Mettelsiefen J, Kirschner S, Gunther KP. Social, educational, and occupational predictors of total hip replacement outcome. Osteoarthritis Cartilage 2010; 18(8):1036-1042.

(48) Graves CM, Otero JE, Gao Y, Goetz DD, Willenborg MD, Callaghan JJ. Patient reported allergies are a risk factor for poor outcomes in total hip and knee arthroplasty. J Arthroplasty 2014; 29(9 Suppl):147-149.

(49) Lavernia CJ, Villa JM, Iacobelli DA, Rossi MD. Vitamin D insufficiency in patients with THA: prevalence and effects on outcome. Clin Orthop Relat Res 2014; 472(2):681-686.

(50) Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40(5):373-383.

(51) van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care 2009; 47(6):626-633.

(52) McHorney CA, Ware JEJ, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31(3):247-263.

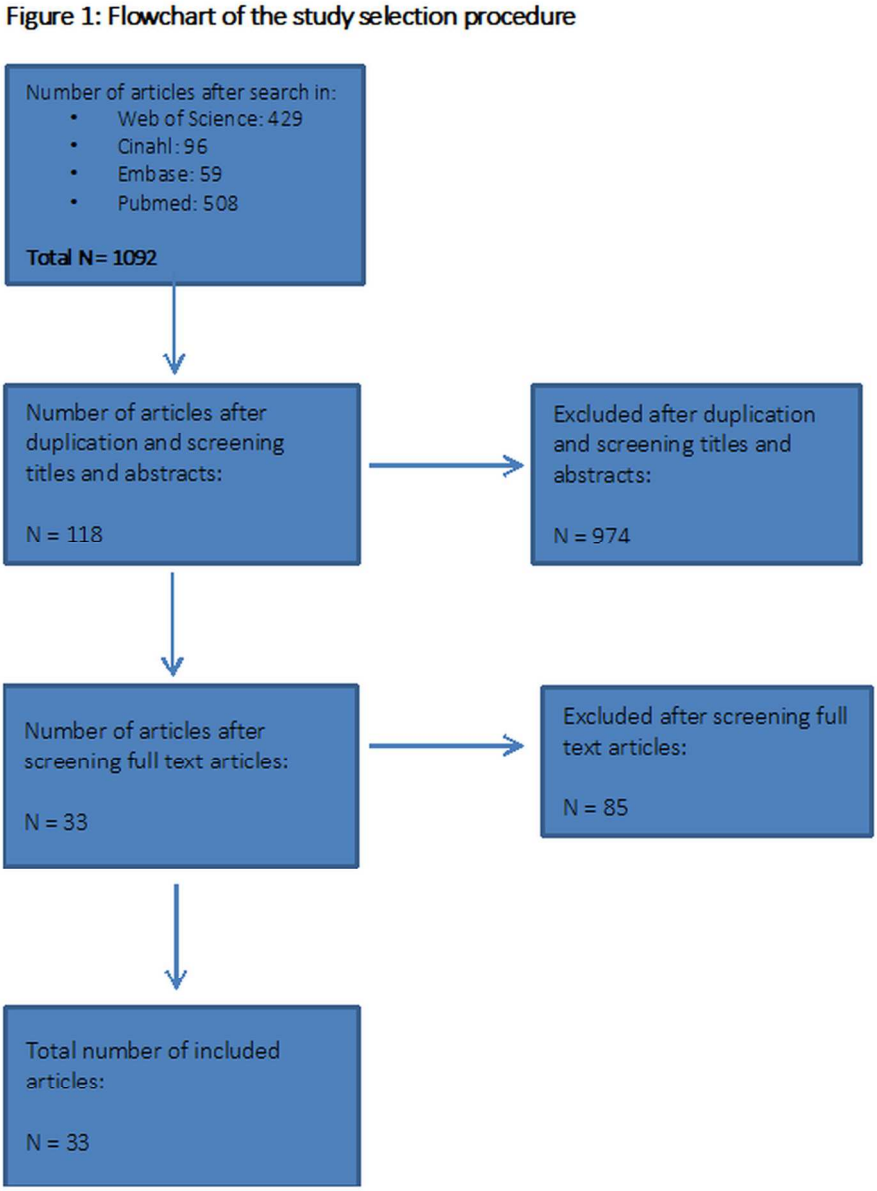


Figure 1: The study selection procedure.

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both. *Comment:* *Tittle*: *Predictors of physical functioning after total hip arthroplasty: a systematic review.*	Page 1.
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. *Comment:* *We performed a structured summary including all mentioned aspects.*	Page 2.
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. *Comment:* *Rationale is described in detail in the Introduction.*	Page 3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). *Comment:* *This systematic review aims to evaluate the predictors of functional outcome after total hip arthroplasty.* *P = subjects with osteoarthritis* *I = total hip arhtroplasty* *C = -* *O = functional outcome after THA* (HHS, OHS, SF-36, LEFS, TUG, WOMAC) *S= systematic review*	
METHODS			

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. *Comment:* *The protocol can be retrieved electronically through the first author.*	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. *Comment:* *All follow-up lengths and languages were included. The study characteristics can be found at the selection criteria.*	Page 4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. *Comment:* *With the help of an independent medical librarian we conducted a literature search through four different databases: Web of Scienel; Cinahl; Embase and PubMed. This literature search was performed on June 23, 2015. The exact search strategy can be found in the methods chapter.*	Page 4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. *Comment:* *Details of the flowchart and entire search strategy are described in Figure 1*	Figure 1.
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). *Comment:* *Two of the authors (LvB and TP) first independently screened the titles and abstracts of all the articles, using the above mentioned selection criteria. Both reviewers screened the full-text articles of the articles found eligible in the first round. A third author (LDB) compared these results and in case of different opinions, consensus was reached. The study selection procedure is schematically presented in figure 1.*	Page 5
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. *Comment:* *One of the authors (LDB) extracted the data, double checked by a second author. The extracted data of all included studies are summarized in table 1. These results are categorized by predictor variable. From all the articles, the following information was extracted: (1) predictor variable; (2) author (3) year of publication; (4) level of evidence; (5)*	Page 5

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

		number of patients; (6) measurement tools that are used; (7) follow-up period; (8) significance level; (9) association between predictor variable and outcome measure; (10) predictor level of measurement.	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. *Comment:* *See methods for the complete search strategy.*	Page 4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. *Comment:* *Risk of bias and our attempt to reduce the risk of bias in the individual studies was described in both the Subjects and Methods section and in the Results.*	Page 4,5,7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). *Comment:* *Best-evidence synthesis were described in the Subjects and Methods section. As this systematic review was a qualitative synthesis of the available evidence. In view of the heterogeneity of the target population, the variability of study objectives and differences in methodological quality, a meta-analysis could not be performed.* *In the Results we described in detail our findings with regard to the predictors of functional outcome after THA. All predictors evaluated in literature are reported.*	Page 6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis. *Comment:* *The performed best-evidence syntheses were described in the Methods section.* *In the results and Table 2-7 we described in detail our findings of the predictors of functional outcome after THA. This systematic review was a qualitative synthesis of the available evidence, a meta-analysis could not be performed. The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up..*	

Section/topic	#	Checklist item	Reported on page #

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). *Comment:* *Risk of bias in the individual studies was determined by the GRADE approach and displayed in Table 8.*	Table 1
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. *Comment:* *Not applicable.*	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. *Comment:* *Figure 1 shows the flow of information through the different phases of the systematic review.*	Figure 1.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. *Comments:* *See Table 2-7 for all extracted data.*	Table 2-7
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). *Comment:* *See Table 1 for the GRADE rating scheme.*	Table 1
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. *Comment:* *Details about the individual studies are described in the Results and in detail in Table 2-7*	Table 2-7
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency. *Comment:* *The reason we could not apply a meta-analysis is because of the heterogeneity across studies regarding measurement tool, predictor and duration of follow-up.*	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15).	Table 1

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

		Comment: *See Table 1 for the GRADE rating scheme.*	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). *Comment:* *Not applicable.*	
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). *Comment:* *The main findings and their implications are described in the Discussion section..*	Page 20
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). *Comment:* *Limitations of the review were described in detail in the discussion section.*	Page 20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research. *Comment:* *We described that methodological well-conducted, randomized, controlled trials in larger groups of subjects with more equal distribution and extensive measurements methods are necessary to investigate the pain sensitivity and pain perception in obese subjects vs non-obese subjects.* *In addition we advised to study the unknown variables of influence to pain sensitivity and pain perception in obese subjects.*	Page 23
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. *Comment:* *This systematic review was performed without any funding and the authors have no disclosure of conflicts of interest.* *The authors have no disclosure of conflicts of interest regarding the systematic review.*	24

PRISMA 2009 Checklist Predictors of functional outcome after total hip arthroplasty: a systematic review.

Buirs et al. 2015

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.