EFORT open reviews

GENERAL ORTHOPAEDICS | 8:11 | 781–791

Clinical investigations to evaluate high-risk orthopaedic devices: a systematic review of the peer-reviewed medical literature

Anne Lübbeke[1,2], Christophe Combescure[3], Christophe Barea[1], Amanda Inez Gonzalez[1], Keith Tucker[4], Per Kjærsgaard-Andersen[5], Tom Melvin[6], Alan G Fraser[7], Rob Nelissen[8] and James A Smith[9,10]

[1]Division of Orthopaedic Surgery & Traumatology, Geneva University Hospitals and University of Geneva, Switzerland
[2]Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, UK
[3]Division of Clinical Epidemiology, Geneva University Hospitals and University of Geneva, Switzerland
[4]Chair Orthopaedic Data Evaluation Panel, UK
[5]Center for Adult Hip and Knee Reconstruction, Department of Orthopaedics, South Danish University, Vejle Hospital, Denmark
[6]School of Medicine, Trinity College Dublin, Ireland
[7]Department of Cardiology, University Hospital of Wales, Cardiff, UK
[8]Department of Orthopaedics, Leiden University Medical Center, Leiden, The Netherlands
[9]Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, UK
[10]National Institute for Health Research Oxford Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK

Correspondence should be addressed to A Lübbeke
Email
anne.lubbekewolff@hcuge.ch

- *Purpose:* The objective of this systematic review was to give an overview of clinical investigations regarding hip and knee arthroplasty implants published in peer-reviewed scientific medical journals before entry into force of the EU Medical Device Regulation in May 2021.
- *Methods:* We systematically reviewed the medical literature for a random selection of hip and knee implants to identify all peer-reviewed clinical investigations published within 10 years before and up to 20 years after regulatory approval. We report study characteristics, methodologies, outcomes, measures to prevent bias, and timing of clinical investigations of 30 current implants. The review process was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.
- *Results:* We identified 2912 publications and finally included 151 papers published between 1995 and 2021 (63 on hip stems, 34 on hip cups, and 54 on knee systems). We identified no clinical studies published before Conformité Européene (CE)-marking for any selected device, and no studies even up to 20 years after CE-marking in one-quarter of devices. There were very few randomized controlled trials, and registry-based studies generally had larger sample sizes and better methodology.
- *Conclusion:* The peer-reviewed literature alone is insufficient as a source of clinical investigations of these high-risk devices intended for life-long use. A more systematic, efficient, and faster way to evaluate safety and performance is necessary. Using a phased introduction approach, nesting comparative studies of observational and experimental design in existing registries, increasing the use of benefit measures, and accelerating surrogate outcomes research will help to minimize risks and maximize benefits.

Keywords

- systematic review
- hip arthroplasty
- knee arthroplasty
- clinical evidence
- registries
- medical device regulation

EFORT Open Reviews
(2023) **8**, 781–791

Introduction

Little is known about the clinical evidence used to establish the safety and performance of medical devices before and after market access in Europe. Unlike medicines in Europe and in the USA, and medical devices that are subject to pre-market authorization in the USA, there has been no requirement for summaries of clinical evidence to be made publicly available. Under the Medical Device Directive 93/42/EEC (MDD) system, which is still the legal basis for the marketing of the vast majority of medical devices today, it is not possible to identify the clinical evidence

© 2023 the author(s)

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

supporting device CE-marking (Conformité Européene) as this is considered to be commercially confidential (Article 20 of the MDD). This might be the reason for the very few detailed analyses on evidence for medical devices being published.

The Medical Device Regulation (MDR) ((EU) 2017/745) is changing the requirements for certification (CE-marking) of implantable medical devices in Europe. The MDR will increase transparency of the clinical investigations supporting device CE-marking, by requiring the publication of clinical investigation reports (MDR, Article 77), and it may increase the clinical evidence requirements for some devices. For example, a clinical investigation is required for Class III devices, unless the use of existing clinical data is sufficiently justified. The MDR has also introduced restrictions with respect to the use of data from equivalent devices for the purpose of market entry, with a contract required between manufacturers for high-risk devices (MDR Article 61(5)).

The peer-reviewed medical literature is an established major source of clinical evidence regarding medical devices (1). In orthopaedic surgery, information derived from the published literature is complemented by annual reports from registries, which monitor real-world safety and performance of implants at the national or regional level over the long term (2). EU regulatory and health technology assessment bodies have recognized the importance of high-quality registries and wish to optimize their use to generate evidence to support decision-making in clinical practice (3).

The European Commission has funded the Coordinating Research and Evidence for Medical Devices (CORE-MD) consortium to review and recommend methodologies for the improved clinical investigation and evaluation of high-risk medical devices (4). An important component for recommending how devices should be evaluated in the future is understanding how they have been assessed as well as addressing the strengths and limitations of previous evaluation approaches. The aim of the current project is to review the evidence for high-risk orthopaedic devices; the quality and validity of registries are covered elsewhere by the CORE-MD consortium (5).

Despite changes to the clinical evidence requirements for medical devices under the MDR, a systematic review of studies supporting CE-marking under the MDD is useful for several reasons. First, it will provide a better understanding of the availability of published evidence for clinicians and healthcare systems. Secondly, it will provide a useful baseline against which to evaluate the impact of the MDR on clinical investigations and the evidence available in the future. Thirdly, it will allow comparison to evidence available for devices in other regulatory environments, which in our project refers specifically to those devices, which have received US FDA market clearance or approval (hereafter clearance).

The objective of this systematic review was to give an overview of clinical investigations regarding hip and knee arthroplasty published in peer-reviewed scientific medical journals, with a focus on methodology and clinically relevant outcomes, before and after regulatory approval (CE-marking).

Methods

We selected for inclusion a total of 30 hip and knee devices used for primary hip or knee replacement. For each device, we attempted to discover the date of the first CE-marking, and we conducted a systematic literature search to identify all published literature available 10 years before and 20 years after the introduction of these implants. We identified studies assessing patients who would receive the hip or knee implant under its typical intended use, and we described evidence reported in the studies.

The systematic review is reported according to the relevant items of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (6) statement, and it was registered on the open science framework (https://osf.io/6gmyx)

Selection of devices (implants) for inclusion in the review

This review aimed to assess a representative sample of CE-marked medical devices. Since a complete list is not available, two sources were used: the Orthopaedic Data Evaluation Panel (ODEP, https://www.odep.org.uk/, accessed 8 June 2021) and European national registries. Consultation with CORE-MD members including regulatory agencies identified ODEP as having one of the most complete lists of hip and knee implants available on the European market. National registries from Denmark, Finland, Germany, the Netherlands, Norway, Sweden, Switzerland, and the UK were also searched. Merging these two sources, we obtained lists of hip cups (n = 138), hip stems (n = 165), and knee (n = 97) implants. From that pool of CE-marked implants, ten devices were then randomly selected from each of the three lists.

The unit of analysis used was determined for the hip by the implant name and the type of fixation (i.e. cemented or cementless) and for the knee by the implant name and the type of stability in accordance with International Society of Arthroplasty Registries (ISAR) Benchmarking recommendations (7).

CE-marking and FDA clearance dates

We identified CE-marking dates by asking ODEP, to which manufacturers often provide them. If unsuccessful,

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

we then searched the internet for press releases, manufacturers' brochures, or mentions in academic papers that stated the date or that indicated the approximate date.

We searched for the selected medical devices in the FDA medical device databases to establish whether they had FDA clearance and, if so, to record the date of clearance (https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/search/default.cfm).

Search strategy

For the published literature, we searched Embase through Ovid, PubMed, and Web of Science. All Web of Science core collection editions, apart from Conference Proceedings Citation Index – Science (CPCI-S) – 1990–present and Conference Proceedings Citation Index – Social Science & Humanities (CPCI-SSH) – 1990–present, were searched. We used the general structure of 'Device name' AND 'Hip' (or 'Knee')] AND 'Humans' for all searches. Search results were combined and automatically de-duplicated in Endnote web, and one author (JAS) manually de-duplicated the results before screening for inclusion and exclusion was done. Full details of searches are provided in Appendix II (see section on supplementary materials given at the end of this article). Searches were limited to 10 years before the CE-marking date and 20 years afterwards. References of relevant systematic reviews were reviewed to identify additional clinical investigations.

Inclusion and exclusion criteria

We included studies that reported clinical investigations (defined by MDR Article 2(45)) of the devices of interest. We operationalized 'undertaken to assess the safety or performance of a device' as (i) the study specifically aimed to assess the device in question using at least one of the safety and performance outcomes of interest (defined further) in the context of the usual use of the device and (ii) the outcomes were presented by the device. Studies that tested something other than the device were excluded (e.g. testing of different wound dressings in two groups that both received the implant of interest).

We included case reports and series, case–control studies, registry-based cohorts, cohort studies, and randomized controlled trials (RCTs).

The outcomes of interest were as follows:

- All-cause revision, assessed at a specific time point (a count of events without any information about when those events occurred would not be included).
- Assessment of implant migration or periprosthetic osteolysis (recognized surrogate markers for implant failure).
- Assessment of the patient-reported outcomes (PROs).
- Frequency of postoperative orthopaedic complications relevant to arthroplasty (if these were defined as a distinct outcome in the study).

We only included studies describing the results of the selected implants in the context of primary total joint replacement. Studies describing results in the context of revision surgery, after hip fracture only, or in any other unusual subpopulation or in cadavers, and conference abstracts, were excluded.

If more than one paper described the findings of a study, then the most comprehensively reported paper was included to avoid duplicate data. Studies written in a language spoken by one of the investigators (English, French, and German) were included.

Data collection and management

Details are provided in Appendix III. We collected information such as CE-mark date, manufacturer, and FDA clearance date in Microsoft Excel. Data extracted from published literature were documented in a database created for this project in REDCap. Two reviewers (AL and JAS) screened all records, and two reviewers (hip stems and knees: AL and JAS; hip cups: AL and AIG) extracted all data in duplicate and discussed and consolidated any differences, with the exception of non-English language studies, which were only extracted by the reviewer (AL) who spoke that language.

Analysis

Characteristics of clinical investigations in the published literature were described in terms of study location, year, study design, methodology, and outcomes. We intended to describe investigations performed before and after CE-mark dates separately, but no investigations before CE-marking were identified. We considered studies published up to 2 years after the date of CE-marking or FDA clearance to have been performed pre-CE-mark or FDA clearance. Where information was available, we compared studies available pre- and post- FDA clearance.

Results

Through the literature search, 2901 peer-reviewed publications were identified, and 11 additional papers were found via their references. After de-duplication, in most cases using the full text, we finally included 151 published between 1995 and 2021, of which 63 were for the 10 hip stems, 34 for the 10 hip cups, and 54 for the 10 knee systems (Appendix I). Table 1 summarizes the number of studies identified and included at each stage of the systematic review.

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
www.efortopenreviews.org Attribution-NonCommercial License.
https://creativecommons.org/licenses/by-nc/4.0/

Table 1 Literature search results. The number of articles is presented in the table.

	Implant type			Total
	Hip stem	Hip cup	Knee	
Embase	408	199	825	1432
PubMed	238	50	399	687
Web of Science	293	137	352	782
Before de-duplication	939	386	1576	2901
After de-duplication	751	302	1078	2131
Other sources	9	1	1	11
Studies included	63	34	54	151

Information on the CE-mark year was found for 28 of the 30 implants (Table 2 and Appendix IV). For those 28, all publications dated after their CE-mark (median: 9 years later, range: 3–13 years). No peer-reviewed publication was found for eight implants (27%), of which one was a hip stem, four were hip cups, and three were knee systems.

Table 2 Device names and corresponding pre- and post- market publications.

Device name	Pre-market publications, *n*	CE-mark year found	Post-market publications, *n*
Hip stem			
Accolade II	0	Yes	12
Alloclassic Zweymuller SL	0	Yes	19
Avenir	0	Yes	4
BiContact Cementless	0	Yes	8
COLLO-MIS	0	Yes	2
C-Stem AMT Total Hip System	0	Yes	2
Filler 3ND	0	Yes	1
MiniHip	0	Yes	8
QUADRA	0	Yes	7
Stelia stem	0	Yes	0
Hip cup			
ANA.NOVA cup	0	Yes	2
aneXys	0	Yes	0
Cenator	0	Yes	0
EcoFit Cementless	0	Yes	0
Exceed ABT Cup	0	Yes	4
IP X-LINKed acetabular cup	0	Yes	0
Plasmacup SC	0	Yes	9
POLARCUP™ Cemented	0	Yes	3
RM pressfit Vitamys	0	Yes	8
Versafit CC Trio	0	Yes	8
Knee system			
ACS Unc, Unicondylar	0	No	0
balanSys CR	0	Yes	4
Innex Gender	0	Yes	0
LCS Complete	0	Yes	10
Logic PS	0	Yes	4
NexGen CR	0	Yes	18
Optetrak CR	0	Yes	0
Sigma high-performance partial knee	0	Yes	3
TREKKING CR	0	No	2*
Vanguard CR	0	Yes	13

*No CE-mark year information for trekking; identified publications were from 2012 and 2018.

Study characteristics, methodology, and outcomes overall and by device group

The majority of studies had been conducted in Europe (64%) (Table 3). This proportion was similar for hip stems, cups, and knee systems. On average there were five publications (range 0–19) per implant within the period up to 20 years after the CE-mark year. The median time between inclusion of the first patient into a study and the publication of the results was 10 years (range: 2–22 years).

The FDA had approved 16 of the 30 randomly selected implants for use in the USA (Appendix IV). Overall, devices had been approved by CE-marking earlier in the EU, at a median interval before approval by the FDA in the USA of 4.6 years (range: −1 year to +17.8 years). In six cases, regulatory approval was obtained around the same year (within a period of 1 year). On average, the first publication for those hip and knee devices appeared 5 years after approval by the FDA (median interval: 5.0 years, range: 8 years before to 10 years afterwards).

The median duration of follow-up in the selected studies was 4.6 years, ranging from 0.1 to 17.8 years, and the mean duration was 5.2 years (S.D. ± 3.7). Median follow-up was 1.7 years longer in studies evaluating hip prostheses compared to knee implants (Mann–Whitney *U* test $P = 0.033$). More than half of the hip studies (56% of cup and 52% of stem studies) reported follow-up times between 5 and 17 years, while 37% of the knee studies reported follow-up times between 5 and 13 years (Fig. 1 and Table 3).

The median number of implants evaluated in a study (counting only the selected implant, not its comparators) was 139, ranging from 1 to 27,193. Forty-four per cent of studies included a comparator group, which was more common for knee than for hip implant studies (59% vs 35%, Pearson chi-square $P = 0.004$). Regarding study design, the majority were cohort studies (72%), which were mostly retrospective and conducted in one or more academic institutions/hospitals. Adjustment for baseline imbalances in prognostic factors was performed in 26% of studies. Cohort studies based on prospectively collected national or regional registry data made up 13% of the studies. RCTs constituted 9%. In 6 of the 14 RCTs (43%), blinding of the assessor or the patient was indicated. Knee arthroplasty tended to be more frequently assessed by registry-based cohort studies and RCTs than were hip arthroplasty devices (Fisher's exact test $P = 0.085$ and Pearson chi-square $P = 0.08$, respectively; Table 3).

The mean age of subjects (in all studies taken together) was 63 years (range: 24–88 years). Women represented 55%, and in 80% of the participants, the diagnosis was primary osteoarthritis (OA). Demographics differed between hip and knee arthroplasty patients (Table 4).

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

Table 3 Study characteristics and study methodology.

	Hip stems	Hip cups	Knees	All
Number of articles	63	34	54	151
Publication period	1995–2021	2007–2021	2002–2021	1995–2021
Location (%)				
Europe	66.7	70.6	61.1	63.6
Americas	23.8	0	29.6	19.9
Asia	1.6	23.5	9.3	9.3
Other	7.9	11.8	1.9	5.3
Study type				
Case report	3.2%	11.8%	1.9%	4.6%
Case–control	–	–	5.6%	2%
Cohort registry based	7.9%	11.8%	18.5%	12.6%
Other cohorts	84.1%	67.6%	59.3%	71.5%
Retrospective*	83.0%	56.5%	62.5%	72.2%
RCT	4.8%	8.8%	14.8%	9.3%
Comparator group(s), yes	41.3%	23.5%	59.3%	43.7%
Adjusted† analysis, yes	25.4%	5.9%	38.9%	25.8%
Number of prostheses included				
Mean	615	613	1460	917
Median (range)	139 (1–14'147)	95 (1–14'147)	180 (1–27'193)	139 (1–27'193)
Median inclusion period (years)	3	2	3	3
Follow-up‡ (years)	5.5 (0.1–17.8)	5.0 (0.3–15.0)	3.4 (1–13.4)	4.6 (0.1–17.8)
First inclusion date to publication in years‡	10 (4–22)	9 (2–21)	11 (3–20)	10 (2–22)
FDA approval to first publication in years‡	5 ((–8)–10)	2 (1–3)	5 ((–3)–8)	5 ((–8)–10)
CE-mark date to first publication in years‡	9 (3–13)	10 (7–12)	7 (5–10)	9 (3–13)

*Percentage of other cohorts; †matching instead of adjustment was used in 1 study; ‡values are median (range).

Complete information on the devices used – including cup–stem combination, fixation of the combination, and bearing surface for the hip and stability, mobility, fixation, and patella resurfacing for the knee – was found in 32% of the publications. Information was incomplete in 52%, and no information other than the device name was reported in 16%.

The most frequently reported outcome was all-cause revision (74% of studies), followed by orthopaedic complications (73%) and by imaging results (72%) (Table 5). Complications recorded were prosthetic joint infection, dislocation, or periprosthetic fracture or else a thromboembolic event or myocardial infarction. The occurrence of these complications overall and by device group is detailed in Table 5. PROs were assessed in 36% of the studies. There were fewer imaging results reported in knee as compared to hip (stem and cup combined) studies (56% vs 80%, Pearson chi-square $P = 0.001$) and more functional outcomes in knee studies (59% vs 2%, Pearson chi-square $P < 0.001$).

A safety concern or an inferior result as compared to another group on one of the outcomes was clearly expressed in 5%, and a potential concern in another 7% of the studies (Table 5). In hip arthroplasty studies, it was most frequently based on imaging results (especially radiographs), whereas in knee arthroplasty, it was based mostly on revision rates and PROs.

Study methodology and outcomes by device name

There were large variations between implants in sample size, follow-up period, study methodology, and outcomes for the published studies (Table 6). For 10 of the 30 implants, we found no comparative study and for 12 no prospective study. For 11 implants, no study reporting on PROs was found. Comparative PRO information was published for 12 implants (40%). Information on revision rates was missing for the eight implants with no post-market publication. Comparative revision rates including reporting of cumulative failure or survival and 95% CIs were available for 11 implants (37%).

Comparison of study methodology and outcomes in studies that were registry-based vs those that were not

There were large differences in sample size, reported methodology, and outcomes between cohort studies that

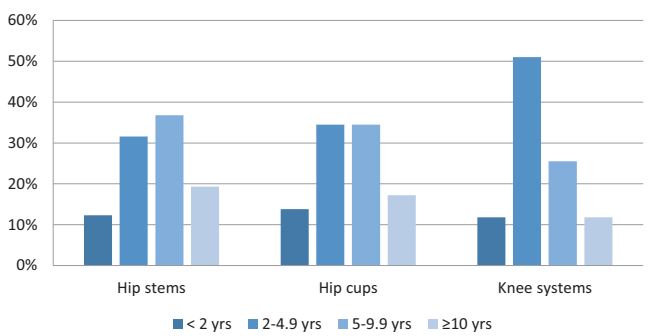

Figure 1
Length of study follow-up by device group.

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
www.efortopenreviews.org
https://creativecommons.org/licenses/by-nc/4.0/

Table 4 Patient characteristics.

Characteristics	Hip stems	Hip cups	Knees	All
n	63	34	54	151
Age,* years (weighted)	62.1 (41–88)	67.8 (24–75)	69.3 (54–77)	68.1 (24–88)
Women (%)*	50.1 (0–100)	46.9 (0–100)	66.8 (10–100)	55.3 (0–100)
Primary OA (%)*	73.8 (15–100)	68.3 (0–100)	94.5 (78–100)	79.5 (0–100)
Mortality (%)				
Mean	11.2	6	7.7	8.7
Median (range)	7.8 (0–42.8)	2.2 (0–30.3)	0.9 (0–44.2)	3 (0–44.2)
Lost-to-follow-up (%)				
Mean	6.4	5.8	6.2	6.3
Median (range)	5 (0–22.1)	4.7 (0–21.1)	5.5 (0–23.4)	5 (0–23.4)

*Values are mean (range).

were based on registries and those that were not (Fig. 2). The median numbers of prostheses were 3341 and 149, respectively, and the median numbers of revision events were 102 and 3. Studies based in registries more often were prospective, had a comparison group, had more precise reporting of all-cause revision reporting, and more often adjusted analyses. The variety of outcomes assessed was lower in registry based than in other types of studies.

Trends in study methodology and outcomes

Temporal trends in selected characteristics and outcomes are shown in Figs. 3 and 4, combining data from hip and knee arthroplasty studies. There was an increase in comparative, prospective, and registry-based RCTs and radiostereometric analysis (RSA) studies, in particular, between the first period (1995–2003) and the second period (2004–2012). The largest increase was in the

Table 5 Outcomes reported. Data are presented as median (range) or as reported.

Outcomes reported	Hip stems	Hip cups	Knees	All
n	63	34	54	151
All-cause revision	81%	67.6%	70.4%	74.2%
n revisions reported	4.5 (0–440)	1.5 (0–440)	3 (0–437)	4 (0–440)
Time-to-event analysis (95% CI)	25.4%	29.4%	33.3%	29.1%
PROs	23.8%	44.1%	46.3%	36.4%
Imaging	77.8%	85.3%	55.6%	71.5%
RSA study	8.3%	5.9%	9.3%	7.3%
Functional measures	1.6%	2.9%	59.3%	22.5%
Complications (excl. revision)	79.4%	73.5%	66.7%	73.5%
Reported complications %				
Prosthesis infection (%)				
Mean	0.9	0.5	0.6	0.7
Median (range)	0.8 (0–2.4)	0.1 (0–2.9)	0.4 (0–2.1)	0.7 (0–2.9)
Dislocation (%)				
Mean	1.5	5.5	0	2.2
Median (range)	0.9 (0–8.2)	1.0 (0–100)	0 (0–0.4)	0 (0–100)
Fracture (%)				
Mean	8.3	1.3	0.2	4.4
Median (range)	1.6 (0–100)	1.1 (0–4.4)	0 (0–1.8)	0.5 (0–100)
Thromboembolic event (%)				
Mean	2.5	1.6	1.0	1.9
Median (range)	1.9 (0–8)	0 (0–4.7)	0.2 (0–4.8)	1.4 (0–8)
Myocardial infarction (%)				
Mean	0.3	0	0	0.2
Median (range)	0 (0–1.2)	0 (0)	0 (0)	0 (0–1.2)
'Concern' reported in study				
No concern expressed	87.3%	82.4%	90.7%	87.4%
Potential	4.8%	11.7%	7.4%	7.3%
Yes	7.9%	5.9%	1.9%	5.3%
'Concern' yes/potential based on				
Imaging	38%	78%	0%	46%
Revision	25%	11%	60%	27%
PROs	0%	0%	40%	9%
Other	38%	11%	0%	18%

PROs, patient-reported outcomes; RSA, radiostereometric analysis.

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

Table 6 Sample size, follow-up, study methodology, and outcomes by implant.

Device name	Sample size (mean)	Follow-up max. (years)	Comparative study (%)	Prospective study (%)	PROs (%)	Revision (%)
Hip stem						
Accolade II	933	5.5	75	33	25	75
Alloclassic Zweymuller SL	159.1	15.5	15.8	0	5.3	89.5
Avenir	294	7	25	0	25	75
BiContact Cementless	182.3	17.8	37.5	37.5	12.5	75
COLLO-MIS	145	5.2	0	0	0	100
C-Stem AMT total hip system	225	6.3	50	50	100	100
Filler 3ND	1313	7	100	100	0	100
MiniHip	68.9	9.4	50	62.5	50	62.5
QUADRA	2755	11.2	57.1	28.6	42.9	85.7
Stelia stem	–	–	–	–	–	–
Hip cup						
ANA.NOVA cup	60	2	0	50	0	100
aneXys	–	–	–	–	–	–
Cenator	–	–	–	–	–	–
EcoFit Cementless	–	–	–	–	–	–
Exceed ABT Cup	547.8	6	50	75	75	75
IP X-LINKed acetabular cup	–	–	–	–	–	–
Plasmacup SC	92.2	15	11.1	44.4	33.3	77.8
POLARCUP™ Cemented	352	11	33.3	0	33.3	100
RM pressfit Vitamys	156.5	5	37.5	75	25	50
Versafit CC Trio	1926	11.2	12.5	25	75	50
Knee system						
ACS Unc, Unicondylar	–	–	–	–	–	–
balanSys CR	98.5	10.7	25	25	50	25
Innex Gender	–	–	–	–	–	–
LCS Complete	1989.5	10.3	60	70	40	100
Logic PS	940	3	75	75	75	50
NexGen CR	1939.2	11.7	66.7	61.1	27.8	77.8
Optetrak CR	–	–	–	–	–	–
Sigma high perf. partial knee	37.3	2.3	66.7	66.7	100	33.3
TREKKING CR	164.5	13.4	100	50	100	50
Vanguard CR	1496.5	10.3	46.2	38.5	46.2	69.2

reporting of PROs, from 0 in the first to 46% in the third period (2013–2021). There was a substantial decrease (from 94% to 64%) in the reporting of radiographic results.

Discussion

This systematic review reports study characteristics, methodologies, outcomes, and timing of clinical investigations in relationship to the CE-marking of high-risk medical devices in orthopaedics (hip and knee implants) before entry into force of the EU MDR in May 2021. We identified no clinical studies published before CE-marking for any selected device and no studies, even up to 20 years after CE-marking, in one-quarter of devices. There were very few RCTs, and registry-based studies generally had larger sample sizes and better methodology.

Previous systematic reviews of hip and knee arthroplasty implants largely corroborate our findings. The lack of evidence in 27% of the hip and knee implants in our review is very similar to the proportions reported in publications from the UK (24%), Norway (30%), and Catalonia (23%) (8, 9, 10). The absence of clinical studies published before CE-marking reflects the regulatory situation under the former MDD (93/42/EEC) and confirms literature focusing on medical devices in general (11, 12). Our finding that RCTs were done to assess only 9% of these hip and knee implants is identical to results from an evaluation of evidence available for implants used in Norway between 1996 and 2000 (9) and to the review of levels of evidence of studies published in major orthopaedic journals (13). The observed absolute and proportional increases in reporting of PROs in our study are in accordance with Siljander *et al.* (14) who found an increase from 21% in 2004 to 48% in 2016 in arthroplasty publications in four major orthopaedic journals.

Lack of premarket evidence

The lack of evidence published before CE-marking that was observed in this review is consistent with other studies (12). Several calls have been made for more evidence to be available before regulatory approval, and

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
www.efortopenreviews.org
https://creativecommons.org/licenses/by-nc/4.0/

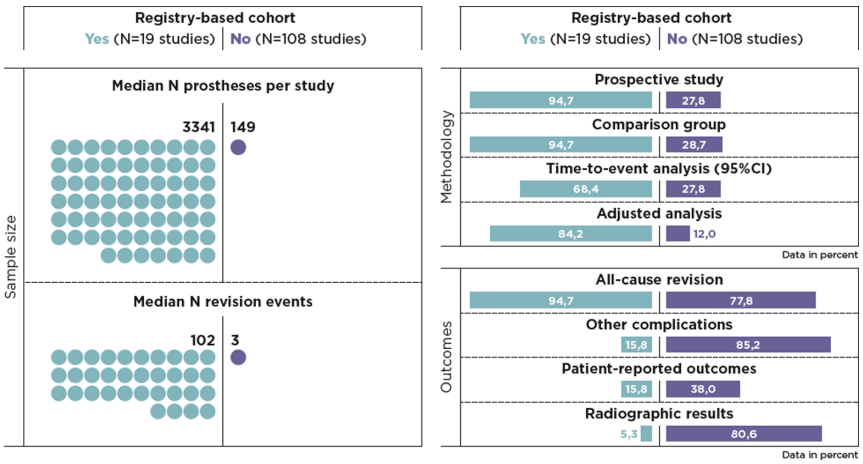

Figure 2
Comparison of cohort studies according to whether or not they were conducted within a registry (median follow-up 5 years for both).

particularly for high-risk devices for which alternatives are available, higher evidence requirements would inform better clinical decisions. Limited pre-market evidence might sometimes be acceptable, if complemented by appropriate post-market studies for similar devices but that should not be a commonplace as in the past. For several devices, however, we found neither pre- nor post-market published studies. Considering the high revision rates of some devices, a phased introduction of new implants is paramount to assure optimal safety (15).

Post-market evidence and its adequacy

Post-market clinical follow-up (PMCF) studies must resolve questions that are unanswered at the time of regulatory approval, regarding clinical benefit throughout the expected lifetime of a device, its safety under widespread use, the generalizability of pre-market findings, and the continuing acceptability of its benefit–risk ratio. Under the MDR, post-market surveillance is expected to be proactive and continuous, with clinically meaningful comparator(s) and clinically relevant endpoints (risks and benefits). The evidence identified in this review would often not have met those expectations. For 27% of implants, we found no published post-market evidence. Comparative studies reporting on PROs were missing for 60% of the implants, and comparative studies reporting cumulative failures or survival rates (with 95% CI) were missing for about two-thirds of the implants.

Of the outcomes included in our review, all-cause revision is the main performance indicator (and risk) of hip and knee arthroplasty in published PMCF studies. Unless a study was nested in a registry (which was the case in 13% of those in our review), the number of revisions in the evaluated publications was generally low. A challenge with revision as a clinically relevant outcome is that the evaluation of implant longevity requires at least 5 years of follow-up, followed by re-evaluations at regular intervals (16). This explains the long follow-up times of about half of the studies in this review.

To reduce the duration of follow-up needed before a new implant can be marketed, an alternative clinically meaningful endpoint should be used in early clinical evaluations. Recognized surrogate outcomes that predict the effect of a therapeutic intervention for long-term implant failure are based on imaging, using RSA, Einzel-Bild-Röntgen-Analyse, or another similar validated radiographic analysis of implant migration (17, 18). A majority of the reviewed hip studies (>85%) and two-thirds of the knee studies reported either radiographic or RSA results. This confirms that there is an important role for academic institutions to evaluate new implants compared to a standard legacy device, before their market approval. Studies to estimate the risk of revision will assess implant migration and osteolysis on radiographs and other surrogate markers.

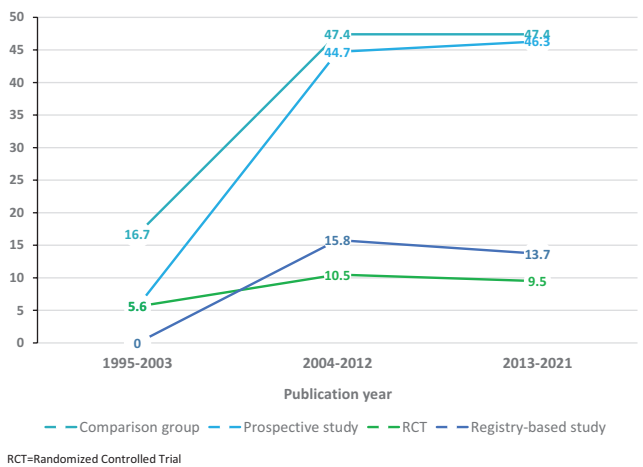

Figure 3
Trends in reported study methodologies for hip and knee implants 1995–2021 (% per period).

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

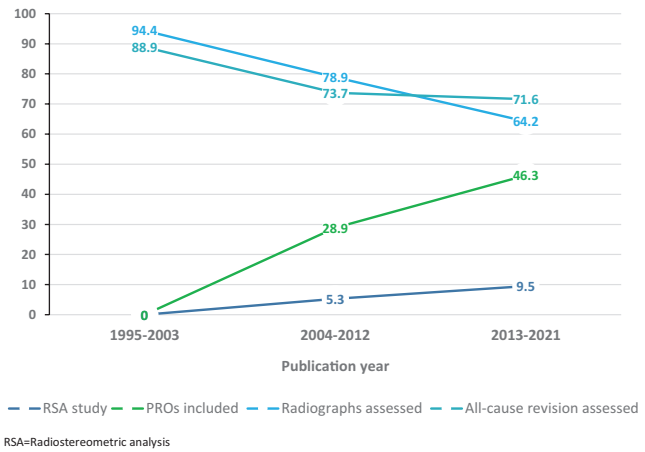

Figure 4
Trends in reported study outcomes on hip and knee implants 1995–2021 (% per period).

Recognized measures of benefit include PROs, which were assessed in half of the more recent studies selected for this review. Another way of measuring benefit is clinician-reported scores, which most of the earlier studies reported. Collection of PROs was more frequent in non-registry-based than in registry-based studies, but collection of PROs in registries has greatly increased over the past decade. Currently, 16 out of 25 arthroplasty registries worldwide record PROs (19).

'Traditional' (non-registry-based) follow-up studies alone are unable to document either clinical benefit throughout the expected lifetime of an implant or its safety under widespread use because those tasks require much larger sample sizes, more comparators, longer follow-up, and real-world results. Registries or large observational population-based cohorts are better because they generate high-quality post-market clinical evidence for legacy and new devices faster and more efficiently (2, 20, 21, 22, 23). They are now recognized by regulators as a preferred source and platform for post-market surveillance and clinical studies (24). Randomized trials using highly accurate methodology such as implant migration analysis in small studies of up to 50 patients per group and observational studies including large numbers of patients should both be nested within registries (25, 26). These studies should be independent and transparent and of high quality (27). This will require more resources for registries or alternative funding schemes.

Limitations

There are several limitations of this study. First, we focused on the peer-reviewed medical literature as the source of information about clinical investigations of 30 selected hip and knee implants. There are other publicly available sources, such as annual registry reports, so our findings likely under-represent the total available evidence for the studied implants. Secondly, we limited the outcomes that were included, so the identified papers do not necessarily represent all those investigating a given implant. Thirdly, we constructed and sampled from a list of medical devices that is unlikely to be exhaustive, because there is no list of CE-marked devices currently available. This means that the reported averages in our study refer to the random sample of our list of hip cups and stems and knee implants but not to all CE-marked hip and knee implants or to other devices such as shoulders. The sources that we used to identify devices (ODEP and registries) preferentially include those that are used in practice, and we would expect such devices generally to have more evidence available for them than is available for those that are used less often. If so, then the included sample may have had more evidence available than would be found for a sample of all CE-marked devices.

Conclusions

There is a common perception that more clinical evidence is needed for high-risk medical devices before they are approved for implantation in patients within the EU and one of the goals of the new EU regulation is to achieve that. An objective of the CORE-MD project is to identify if that will require more clinical studies, better-designed clinical trials, better use of real-world data from high-quality registries, and/or more transparency of the results of clinical investigations. Our systematic review suggests that all those measures will be required.

Publication on the EUDAMED portal of a summary of the safety and clinical performance for each new high-risk device will make clinical evidence available at the time that it is approved, instead of many years later when the first paper appears. As the peer-reviewed literature provides insufficient evidence from clinical investigations of high-risk devices, a more systematic, efficient and faster approach to evaluating safety and performance is necessary. Performing randomized studies in small groups of patients using imaging should detect badly or underperforming orthopaedic implants before CE-marking. After market approval, nesting studies of observational and experimental design within existing registries or cohorts, increasing the use of benefit measures, and accelerating surrogate outcomes research would optimize an implant's benefit–risk ratio.

Supplementary materials
This is linked to the online version of the paper at https://doi.org/10.1530/EOR-23-0024.

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
www.efortopenreviews.org
https://creativecommons.org/licenses/by-nc/4.0/

Declaration of interest
AL declares no conflicts of interest. AL is the current president elect of the ISAR. CC declares no conflict of interest that could be perceived as prejudicing the impartiality of the research reported. CB declares no conflict of interest that could be perceived as prejudicing the impartiality of the research reported. AIG declares no conflict of interest that could be perceived as prejudicing the impartiality of the research reported. KT declares no conflict of interest that could be perceived as prejudicing the impartiality of the research reported. PKA declares no conflict of interest that could be perceived as prejudicing the impartiality of the research reported. AGF declares no conflict of interest that could be perceived as prejudicing the impartiality of the research reported. TM declares no Competing Financial Interests and declares the following Non-Financial Interest: he is an unpaid advisory board member of Pumpinheart Ltd.; previously a senior medical officer in medical devices at the Health Products Regulatory Authority, Ireland; previous co-chair of the Clinical Investigation and Evaluation Working Group of the European Commission. RN declares no conflict of interest that could be perceived as prejudicing the impartiality of the research reported. JAS became a consultant and subsequently employee of Alvea LLC beginning in January 2022.

Funding
This study was supported by a Horizon 2020 grant from the European Union (project number 965246).

Data availability
All data are publicly available on Open Science Framework (https://osf.io/6gmyx).

Acknowledgements
We are grateful for review and input to the study protocol from CORE-MD consortium members, in particular, Stephan Windecker, André Frenk, Gearoid McGauran, and Perla J. Marang-van de Mheen. We would like to thank Olga Taylor from ODEP for her assistance.

References

1. **Medical Device Coordination Group Document** MDCG 2020-6. Regulation (EU) 2017/745: Clinical evidence needed for medical devices previously CE marked under Directives 93/42/EEC or 90/385/EEC. A guide for manufacturers and notified bodies. April 2020. Available at: https://health.ec.europa.eu/system/files/2020-09/md_mdcg_2020_6_guidance_sufficient_clinical_evidence_en_0.pdf

2. **Lübbeke A, Silman AJ, Prieto-Alhambra D, Adler AI, Barea C & Carr AJ**. The role of national registries in improving patient safety for hip and knee replacements. *BMC Musculoskeletal Disorders* 2017 **18** 414. (https://doi.org/10.1186/s12891-017-1773-0)

3. **European Medicines Agency** 2022 Development of a joint work plan (2021-2023) between EMA and European HTA bodies facilitated through EUnetHTA21. Available at: https://www.ema.europa.eu/en/documents/work-programme/european-collaboration-between-regulators-health-technology-assessment-bodies-joint-work-plan-2021_en.pdf

4. **Fraser AG, Nelissen RGHH, Kjaersgaard-Andersen P, Szymanski P, Melvin T & Piscoi P**. Improved clinical investigation and evaluation of high-risk medical devices: the rationale and objectives of CORE-MD (Coordinating Research and Evidence for Medical Devices). *EFORT Open Reviews* 2021 **6** 839–849. (https://doi.org/10.1302/2058-5241.6.210081)

5. **Hoogervorst LA, Geurkink TH, Lübbeke A, Buccheri S, Schoones JW, Torre M, Laricchiuta P, Piscoi P, Pedersen AB, Gale CP**, ***et al.*** Quality and reliability of clinical registries for the regulatory evaluation of medical device safety and performance across the implant lifecycle: a systematic review of European cardiovascular and orthopaedic registries. *International Journal of Health Policy and Management* 2023 **12** 7648. (https://doi.org/10.2106/JBJS.K.00907)

6. **Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE**, ***et al.*** The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Journal of Clinical Epidemiology* 2021 **134** 178–189. (https://doi.org/10.1016/j.jclinepi.2021.03.001)

7. International Society of Arthroplasty Registries (ISAR). International Prosthesis Benchmarking Working Group guidance document 2018. Available at: https://www.isarhome.org/publications (accessed 19 May 2021)

8. **Kynaston-Pearson F, Ashmore AM, Malak TT, Rombach I, Taylor A, Beard D, Arden NK, Price A, Prieto-Alhambra D, Judge A**, ***et al.*** Primary hip replacement prostheses and their evidence base: systematic review of literature. *BMJ* 2013 **347** f6956. (https://doi.org/10.1136/bmj.f6956)

9. **Aamodt A, Nordsletten L, Havelin LI, Indrekvam K, Utvåg SE & Sundberg KH**. Documentation of hip prostheses used in Norway. A critical review of the literature from 1996–2000. *Acta Orthopaedica Scandinavica* 2004 **75** 663–676. (https://doi.org/10.1080/00016470410004021)

10. **Chaverri-Fierro D, Lobo-Escolar L, Espallargues M, Martínez-Cruz O, Domingo L & Pons-Cabrafiga M**. Primary total hip arthroplasty in Catalonia: what is the clinical evidence that supports our prosthesis? *Revista Española de Cirugía Ortopédica y Traumatología* 2017 **61** 139–145. (https://doi.org/10.1016/j.recot.2016.10.001)

11. **Naci H, Salcher-Konrad M, Kesselheim AS, Wieseler B, Rochaix L, Redberg RF, Salanti G, Jackson E, Garner S, Stroup TS**, ***et al.*** Generating comparative evidence on new drugs and devices before approval. *Lancet* 2020 **395** 986–997. (https://doi.org/10.1016/S0140-6736(19)33178-2)

12. **Hulstaert F, Neyt M, Vinck I, Stordeur S, Huić M, Sauerland S, Kuijpers MR, Abrishami P, Vondeling H, Flamion B**, ***et al.*** Pre-market clinical evaluations of innovative high-risk medical devices in Europe. *International Journal of Technology Assessment in Health Care* 2012 **28** 278–284. (https://doi.org/10.1017/S0266462312000335)

13. **Cunningham BP, Harmsen S, Kweon C, Patterson J, Waldrop R, McLaren A & McLemore R**. Have levels of evidence improved the quality of orthopaedic research? *Clinical Orthopaedics and Related Research* 2013 **471** 3679–3686. (https://doi.org/10.1007/s11999-013-3159-4)

14. **Siljander MP, McQuivey KS, Fahs AM, Galasso LA, Serdahely KJ & Karadsheh MS**. Current Trends in patient-reported outcome measures in total joint arthroplasty: a study of 4 major orthopaedic journals. *Journal of Arthroplasty* 2018 **33** 3416–3421. (https://doi.org/10.1016/j.arth.2018.06.034)

15. **Nelissen RG, Pijls BG, Kärrholm J, Malchau H & Nieuwenhuijse MJ**. Valstar ER. RSA and registries: the quest for phased introduction of new implants. *Journal of Bone and Joint Surgery. Am.* 2011 **93** (Supplement 3) 62–65.

16. Available at: https://journals.lww.com/jbjsjournal/Pages/Concise-Format-Guidelines.aspx

17. **Malak TT, Broomfield JAJ, Palmer AJR, Hopewell S, Carr A, Brown C, Prieto-Alhambra D & Glyn-Jones S**. Surrogate markers of long-term outcome in primary total hip arthroplasty: a systematic review. *Bone and Joint Research* 2016 **5** 206–214. (https://doi.org/10.1302/2046-3758.56.2000568)

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
Attribution-NonCommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

18. Kärrholm J, Gill RH & Valstar ER. The history and future of radiostereometric analysis. *Clinical Orthopaedics and Related Research* 2006 **448** 10–21. (https://doi.org/10.1097/01.blo.0000224001.95141.fe)

19. Bohm ER, Kirby S, Trepman E, Hallstrom BR, Rolfson O, Wilkinson JM, Sayers A, Overgaard S, Lyman S, Franklin PD, ***et al.*** Collection and reporting of patient-reported outcome measures in arthroplasty registries: multinational survey and recommendations. *Clinical Orthopaedics and Related Research* 2021 **479** 2151–2166. (https://doi.org/10.1097/CORR.0000000000001852)

20. Fraser AG, Byrne RA, Kautzner J, Butchart EG, Szymanski P, Leggeri I, de Boer RA, Caiani EG, Van de Werf F, Vardas PE, ***et al.*** Implementing the new European Regulations on medical devices—clinical responsibilities for evidence-based practice. *European Heart Journal* 2020 **41** 2589–2596. (https://doi.org/10.1093/eurheartj/ehaa382)

21. Wilkinson J & Crosbie A. A UK medical devices regulator's perspective on registries. *Biomedizinische Technik. Biomedical Engineering* 2016 **61** 233–237. (https://doi.org/10.1515/bmt-2015-0142)

22. Sedrakyan A, Campbell B, Merino JG, Kuntz R, Hirst A & McCulloch P. IDEAL-D: a rational framework for evaluating and regulating the use of medical devices. *BMJ* 2016 **353** i2372. (https://doi.org/10.1136/bmj.i2372)

23. Malchau H, Garellick G, Berry D, Harris WH, Robertson O, Kärrholm J, Lewallen D, Bragdon CR, Lidgren L & Herberts P. Arthroplasty implant registries over the past five decades: development, current, and future impact. *Journal of Orthopaedic Research* 2018 **36** 2319–2330. (https://doi.org/10.1002/jor.24014)

24. Medical Device Clinical Evaluation Working Group. *Post-market clinical follow-up studies*. International Medical Device Regulator Forum (IMDRF) 2021. Available at: https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-210325-wng65.pdf

25. Derbyshire B, Prescott RJ & Porter ML. Notes on the use and interpretation of radiostereometric analysis. *Acta Orthopaedica* 2009 **80** 124–130. (https://doi.org/10.1080/17453670902807474)

26. Ochen Y, Gademan MG, Nelissen RG, Poolman RW, Leenen LP, Houwert RM & Groenwold RH. The potential value of observational studies of elective surgical interventions using routinely collected data. *Annals of Epidemiology* 2022 **76** 13–19. (https://doi.org/10.1016/j.annepidem.2022.10.004)

27. Fraser AG. Post-market surveillance of high-risk medical devices needs transparent, comprehensive and independent registries. *BMJ Surgery, Interventions, and Health Technologies* 2020 **2** e000065. (https://doi.org/10.1136/bmjsit-2020-000065)

Downloaded from Bioscientifica.com at 01/03/2025 01:37:07PM
via Open Access. This work is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivatives 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/