Clinical investigations to evaluate high-risk orthopaedic devices: systematic review and meta-analysis

James A Smith[1,2,3], Christophe Combescure[4], Christophe Barea[4], Anne Lübbeke[1,4]

1. Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, UK
2. National Institute for Health Research Oxford Biomedical Research Centre, John Radcliffe Hospital, Oxford, UK
3. Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, UK
4. Division of Orthopaedic Surgery and Traumatology, Geneva University Hospitals, Rue Gabrielle-Perret-Gentil 4, 1205 Geneva, Switzerland

Introduction

Little is known about the clinical evidence used to establish the safety and performance of medical devices pre and post market access in Europe. The new EU Medical device regulation (MDR Article 2(45)) defines 'clinical investigation' as a systematic investigation involving one or more human subjects, undertaken to assess the safety or performance of a device. Unlike drugs in Europe and the US, and devices in the US, no public summaries of clinical investigations supporting device marketing authorisations and post-market requirements are available due to a requirement under medical device Directive 93/42/EEC for data supporting device CE-marking to remain confidential. Perhaps because of this, few detailed analyses have been attempted. The Medical Device Regulation (Regulation (EU) 2017/745; MDR) is changing the requirements for stakeholders in Europe and will increase transparency of the clinical investigations supporting device CE-marking, and 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 European Commission has funded the Co-ordinating Research and Evidence for Medical Devices (CORE-MD) group to review and recommend methodologies for the improved clinical investigation and evaluation of high-risk medical devices[1]. An important component for recommending how devices should be evaluated is understanding how they have been evaluated historically and the strengths and limitations of previous approaches. In this project, we will systematically review the published clinical literature and registry reports for high-risk orthopaedic devices. There are other groups undertaking similar tasks for cardiovascular and diabetes devices.

Despite changes to the clinical evidence requirements for medical devices under the MDR, a systematic review of studies supporting CE marking under the medical device directives (MDD) is useful for several reasons. First, it will provide better understanding of the public

[1] *This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 965246*

availability of clinical results for devices that were available under the MDD, and therefore the evidence basis available for clinicians and healthcare systems. Second, it will provide a useful base against which to evaluate the impact of the MDR on clinical investigations and resulting evidence for devices in the future. Third, it will allow comparison to evidence available for devices from another regulatory environment, specifically in this project to those with FDA market clearance or approval (hereafter clearance). Finally, it will be useful in determining whether the availability and quality of clinical evidence from published studies relates to its real-world device performance and safety as documented in registries.

The MDR specifically mentions registries for the first time in medical device legislation in Europe. Orthopaedics is unique amongst device specialties in the extent to which high quality (inter)national registries have been established and adopted. Particularly among joint replacements, this is the case because the most important measure of device performance is the revision rate, differences in which are only observable after considerable time, making assessment in traditional prospective clinical trials before regulatory approval challenging. As such, it can be argued that registry data often represent the source of evidence of most relevance for performance evaluation of joint replacement devices, after regulatory approval as part of routine clinical practice.

This review focuses specifically on knee and hip implants and describes the clinical investigations that have taken place over the full device life cycle of hip and knee devices.

The specific objectives of this systematic review are:

1. To identify and describe methodologies and outcomes that were available prior to regulatory approval (CE-marking and/or FDA [if applicable])) in the published literature
2. To identify and describe methodologies and outcomes that became available subsequent to regulatory approval (CE-marking and/or FDA [if applicable])) in the published literature
3. To describe and combine for each device the all-cause revision rates in the published literature and in the most recent registry reports and compare the findings between the two data sources.
4. To describe whether safety concerns for a specific device expressed in published studies are concordant to its real-world safety reported by registries

The findings of this investigation will also be qualitatively compared to the findings of the similar investigations in cardiovascular and diabetes (this will probably not be done as part of the research paper(s) reporting the results of this investigation, and will more likely be done in a separate paper).

Methods

In summary, we selected 30 (in total) hip and knee devices used for primary hip/knee replacement for inclusion in this study. For each device, we will attempt to identify the date of first CE-marking. We will then conduct systematic literature searches to identify all published literature available for each device 10 years before and 20 years after CE-marking,

and will summarise studies available before and after CE-marking (and FDA clearance, if applicable).

We will conduct meta-analysis of all cause revision for devices with sufficient data available: We will select devices for meta-analysis on the basis of availability of literature once the literature searches have been done, taking into account practical constraints (for some older devices there may be very large numbers of studies, making additional data collection from all articles too resource intensive). We will aim to include at least 2 devices from each device group (see below).

The protocol is reported according to the relevant items of the preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P)[1].

The study is summarised in the PICO framework as follows (more detail is provided below):

- Participants: Any patients that would receive the device under its typical intended use
- Interventions: hip or knee implants listed in Appendix I.
- Comparators: any
- Outcomes: descriptive summaries of available evidence and revision rate estimates from published studies and registries.

Selection of devices (implants) for inclusion in the study

This study aims to assess a representative sample of CE-marked devices. However, complete lists of CE-marked devices are not available. Through consultation with CORE-MD members, including regulatory agencies, we identified Orthopaedic Data Evaluation Panel (ODEP, https://www.odep.org.uk/, accessed 8th June 2021) as one of the most complete lists of hip and knee implants currently available on the European market. ODEP rates orthopaedic devices against performance benchmarks and their ratings are used throughout Europe. We compared the devices contained in the ODEP lists to those included in several national registries, another source of devices known to be used in Europe (and therefore CE-marked) to assess the coverage. The specific registries used were Denmark, Finland, Germany, Netherlands, Norway, Sweden, Switzerland and UK. The majority of devices included in these registries were already included in ODEP.

We combined devices identified from ODEP and registries into lists of hip cups, hip stems, and knee implants and then randomly sampled each list to identify devices to include in the study (N = 10 in each group). This sample size was chosen to generate a manageable workload for the study team while being informative. The generated datasets do not represent a truly random sample of all CE-marked hip and knee implants but we think are as close as possible to this given the available data. In particular, it seems plausible that the devices included in ODEP and/or registries will be those that are CE-marked and subsequently used, rather than those that are CE-marked but not used substantially in practice.

Until recently, devices used in Europe have not had unique identifying numbers. We therefore have to make a decision about how to define a unit of analysis that is somewhat subjective and will inevitably be imprecise. For example, the QUADRA system (one of the devices we have selected for inclusion) consists of four different products: QUADRA-S, QUADRA-H, QUADRA-C, and QUADRA-R. It is unlikely that we will be able to find the CE-marking date for

all of these individual products (see next section), and that they will be differentiable in papers and registry reports. We will separate products where possible, but will treat them as a single unit of analysis if this is not possible, and use the earliest CE-marking date for the overall unit. Using the earliest date of CE-marking reduces the time available before CE-marking for evidence to be published and therefore may bias the before and after comparison (see below).

CE-marking date

There is no database with CE-marking dates of all devices, and identifying the date of first issue is not straightforward. We will try to identify these dates through several methods:

1. Asking ODEP for the CE-mark date, which manufacturers may have provided to them
2. Searching the internet for press releases or mentions in academic papers that state the date (or indicate approximately the date) of CE-marking device of the product.
3. Contacting the device manufacturers via email to request the date.
4. Contacting notified bodies, where possible: CORE-MD includes representatives of notified bodies, the organisations responsible for conducting the conformity assessments for CE-marking.

For those devices for which these strategies do not find a date, we will try to find the date that devices entered the market in key European countries via national databases, and look at the dates in the Australian Register of Therapeutic Goods. We do not know which of these will best approximate the CE-mark date of the devices, so we will collect this information for the devices with known CE-mark dates from methods 1-3 above, and compare them against the dates from the other databases, selecting the best method.

FDA clearance date

We will search for the devices in the FDA medical device databases to establish whether they have 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 will search 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, will be searched. We will use the general structure of ("Device name" AND "Hip" and "Humans") for all searches. Search results will be combined and deduplicated in Endnote web.

Searches will be limited to 10 years before the CE-marking date and 20 years after. If the CE-marking date is not available, we will try to find a date of first use from e.g. manufacturer brochures and will use that as an approximation.

Example searches from each database are shown in Appendix II.

Registry annual reports will be identified and all regional and national publicly available registry reports (latest update 2019 or 2020) worldwide will be searched.

Searches will be conducted in English but registry reports in any language will be included (and not identified through the searches).

Inclusion and exclusion criteria

We will include studies that report clinical investigations of the devices of interest. The new EU Medical device regulation (MDR article 2(45)) defines 'clinical investigation' as any systematic investigation involving one or more human subjects, undertaken to assess the safety or performance of a device.

In practice we will include:

- Case reports and series
- Case-control studies
- Registry-based cohorts
- Cohort studies
- Randomised controlled trials

We will interpret "undertaken to assess the safety or performance of a device" as follows. The study must specifically aim to assess the device in question using at least one of the safety and performance outcomes of interest (defined below) in the context of usual use of the device. The outcomes must be presented by device. Reports of e.g. registries that provide performance or safety of many devices will be considered includable, but studies that are testing something other than the device, but happen to use the device for all patients in the study, will be excluded (for example, they are testing different wound dressings in two groups that happen to receive the implant of interest).

The outcomes of interest are:

- 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 migration or osteolysis (recognized surrogate markers for implant failure)
- Assessment of the patient reported outcomes (PROMs) listed below
- Frequency of postoperative orthopaedic complications relevant to arthroplasty (must clearly be an outcome in the study)

We will only include studies describing the results of the selected implants in the context of primary total joint replacement. Studies describing the results in the context of revision surgery, after hip fracture only, or any other unusual subpopulation, will be excluded.

If we find meta-analyses or systematic reviews, we will individually assess each study included in the paper rather than extracting data from the meta-analysis directly.

If more than one paper describes the findings of a study the most comprehensively reported paper will be included to avoid duplicate data. This would be the case if e.g. there are two papers in different languages from the same study group around the same time.

Studies written in a language spoken by one of the investigators (English, French, and German) will be included.

Cadaver studies and conference abstracts will be explicitly excluded.

Core data collection items

Devices

The following information will be included for each device included in the study.

- Name of device
- Manufacturer
- Implant reference number, if available
- Date of CE approval
- Date of FDA approval
- Date of first use (ODEP)

Procedure:

For total knee arthroplasty:

- Stability: Cruciate ligament preserved (yes/no), Medial Pivot design, other or not recorded
- Mobility: fixed bearing /mobile bearing/ or not recorded
- Fixation (all cemented/all uncemented/other) or not recorded
- Patella resurfaced (yes/no) or not recorded

For total hip arthroplasty:

- Fixation of stem/cup (cemented yes/no) or not recorded
- Type of bearing surface or not recorded

Associated stem/cup (free text)

Fixation associated (free text)

- Type of bearing or not recorded (yes/no)

Papers

This section details the data that will be extracted from each paper identified in our literature search. We will not contact authors of papers for additional information not present in the papers because we are interested in assessing the published evidence rather than evidence that may have been generated but is not published.

Meta-data:

- First author
- Date of publication (first available online if available)
- Date publication first available
- Submission (or publication) before /after CE mark date
- Submission (or publication) before /after FDA approval date
- Journal
- Study location(s) (continent)
- First and last year of recruitment

Objective (free text, copied from paper)
Key finding (free text, copied from paper)

Study characteristics

- Study type (cohort, registry-based cohort, case control, randomised controlled trial, case series or reports)
- Retrospective, prospective, both elements
- Population-based or specific population (e.g. young patients only)
- Real-world or experimental setting
- Comparative study (yes/no)
- Which comparison implant/group (e.g. established vs. new)
- Study aim (superiority/non-inferiority)
- Randomisation (yes/no)
- Blinding (select from: participant, investigator, outcome assessor)
- Type of RCT (registry-nested, other)
- Clinical trial registration ID provided (yes/no)

Patient characteristics

- Number in study
- Number in device in question arm
- Age (mean/median)
- Women (%)
- Diagnostic (% primary OA)

Investigators and sponsors:

- Author affiliations (academic, industry, mix)
- Is one of the (co-)authors developer of the device (yes/no)
- Industry sponsored (yes/no)
- Sponsored/funded by device manufacturer (yes/no)

Outcomes reported:

- All-cause revision as outcome (yes/no)
    - Revision rate at x years (upper CI, lower CI)
- Imaging (yes/no) if yes, which method
        - Radiograph
        - CT
        - MRI
        - EOS
        - RSA
    - Migration (yes/no)
    - Osteolysis (yes/no)
    - Other (yes/no)
- Patient reported outcome measures (yes/no)
    - Oxford knee score (yes/no)
    - Knee Injury and Osteoarthritis Outcome Score (KOOS) (yes/no)
    - Oxford hip score (yes/no)
    - Hip disability and osteoarthritis outcome score (HOOS) (yes/no)

    - WOMAC (yes/no)
    - EQ-5D (yes/no)
    - SF-36/SF-12 (yes/no)
- Performance (yes/no)
    - Gait (yes/no)
    - Flexion (yes/no)
    - Posterior stability (yes/no)
    - Other (yes/no)

Are analyses stratified or outcomes presented by gender (yes/no)
Are analyses stratified or outcomes presented by age (yes/no)

Safety:
- Did paper report safety concerns? (yes/maybe/no) if yes, which:
    - Higher revision rate
    - Imaging abnormality
    - Inferior clinical results
    - PROMS
    - Biomechanical

Safety concern reported in which section of paper (e.g. abstract, discussion)

We will also record
- Mean and max. length of follow-up
- Adverse events/complications
    - Infection (N and %)
    - Dislocation (N and %)
    - Fracture (N and %)
    - Thromboembolic event (N and %)
    - Myocardial infarction (N and %)
- Mortality (N and %)

Risk of bias

Attrition
- Lost to follow-up (N and % [per group if comparative])
    - Reasons for loss mentioned yes/no

Information bias
- Exposure identification = Procedure (see above) details provided yes/no
- Outcome definition provided yes/no
- Response rate PROs (see above)

Selection bias in observational comparative studies (for RCTs bias assessed above)
- Measures used to reduce bias yes/no and which: Adjustment/Restriction/Matching

Registries

We will include all registries that provide in their annual report the cumulative failure /survival rate with 95% confidence intervals at specific points in time for one or more of the implants selected.

Preliminary list of registries corresponding to these criteria: Australia, Finland, Germany, Netherlands, Norway, RIPO Italy, (Sweden), Switzerland, United Kingdom

Meta-data that will be extracted from registries:

- Registry Country/Region
- National or regional
- Year of latest annual report publication (2019 or 2020)
- Year of first annual report

All-cause revision

The following will be extracted per device:

- Number of devices included
- Total number of observed events
- Timing of measurement (all time points available)
- Point estimates (cumulative incidence of revision or cumulative survival)
- Confidence interval
- First year the use of the device is mentioned in an annual report of the registry
- Pre CE-mark registry data available yes/no

Data management

Data extracted from published literature will be documented in a database created for this project in REDCap. Every device will have a separate folder. Two reviewers will screen all records and extract all data in duplicate and will discuss and consolidate any differences, involving a third reviewer if needed. Non-english language studies will only be extracted by the reviewer that speaks that language (i.e. they will not be done in duplicate).

Data extracted from registry reports will be documented in excel according to a pre-specified format.

Data repository

All data will be made publicly available, including the list of devices and publications identified in our search. Data will be made available in the central CORE-MD database, or, if this is not possible, on the Open Science Framework or a similar repository.

Analysis

First and second objective (To identify and describe methodologies and outcomes that were available before vs. after regulatory approval (CE-marking and/or FDA) in the published literature)

Characteristics of clinical investigations (published literature) will be described separately pre- and post-CE marking date. To allow for a delay in publication of studies after conduct,

we will consider studies published up to 2 years after CE marking or FDA clearance to be pre CE mark or FDA clearance. We will qualitatively compare the data available before and after CE marking across devices. We will do the same stratification for pre and post FDA clearance and qualitatively assess differences.

Third objective (To describe and combine for each device all-cause revision rates in the published literature and in the most recent registry reports and compare the findings between the two data sources)

For each device with available data, the revision-free survival rates at time points where sufficient data are available according to reported revision rates (one time point for short-, one for mid-term and one for long-term if applicable) will be combined across registries using models with random effects (Der Simonian & Laird's approach)[3]. For meta-analyses, a complementary log-log transformation will be applied. Amount of between registries heterogeneity will be assessed with I^2 statistics[4] and Cochran Q test for heterogeneity. In sensitivity analyses, the logit transformation will be applied to check the robustness of the findings. In addition, leave-one-out sensitivity analyses will be conducted to investigate if findings are driven by specific studies. The revision-free survival rates at 5 years will be combined across published studies by using similar methods. The pooled revision-free survival rates will be compared between published studies and registries with a two-sided interaction test with a risk alpha of 5%. Similar analyses will be conducted to combine the survival rates at 10, 15 and 20 years, where sufficient data is available.

To assess the similarity between the revision-free survival rates combined across registries and combined across published studies, the difference between the revision-free survival rates will be reported. The 95% confidence interval of the difference will be obtained by a parametric bootstrap approach. In addition, a comparison of the combined rates will be conducted with a two-sided interaction test with a risk alpha of 5%.

For published studies, a potential publication bias will be investigated by a visual inspection of funnel plots. If more than 10 published studies report revision-free survival rates, Egger's test and the trim and fill method[5] will used. Publication bias is not applicable to registries since the reporting is independent of the revision-free survival rates and it will not be investigated.

Fourth objective (To describe whether safety concerns for a specific device expressed in published studies are concordant to its real-world safety reported by registries)

This objective will be addressed as part of a different work package, and we do not detail the methods for it here. We include it for completeness because we are extracting data relevant to it.

Pilot

Before submitting this protocol, we piloted the full process of searching for papers, including and excluding them, and extracting data into RedCap for three hip stems: Avenir, Alloclassic Zweymuller SL and Quadra. This resulted in some changes to our initial protocol. No data analysis was conducted. We have also collected data from registries.

Presenting results

It is possible that we will split the results of this investigation across multiple papers, since we are collecting a lot of data and it may be clearer to present results across separate write-ups: for example, one for objectives one and two, and one for three and one for objective four.

Acknowledgements

We are grateful for review and input from all CORE-MD consortium members, in particular: Stephan Windecker, André Frenk, Tom Melvin, Gearoid McGauran, Rob Nelissen, Perla J. Marang-van de Mheen, Per Kjærsgaard-Andersen and Alan Fraser.

References

1. Moher, D. *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews* 4, 1 (2015).
2. Taylor, MV. Clinical Evidence. *J Med Libr Assoc* 94, 478–480 (2006).
3. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 7, 177-188 (1986)
4. Higgins JPT, Thompson S G, Deeks J J, Altman D G. Measuring inconsistency in meta-analyses. *BMJ* 327, 557-60 (2003)
5. Duval, S, Tweedie, R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56, 455-63 (2000)

Appendix I: list of devices for inclusion

Hip stems:

Accolade II	Stryker
Stelia stem	Stemcup medical product
QUADRA	Medacta
MiniHip	Corin
Filler 3ND	Biotechni
COLLO-MIS	LimaCorporate
C-Stem AMT Total Hip System	DePuy Synthes
BiContact Cementless	Braun
Avenir	Zimmer Biomet
Alloclassic Zweymuller SL	Zimmer Biomet

Hip cups:

Versafit CC Trio	Medacta
Stanmore Cup	Zimmer Biomet
POLARCUP™ Cemented	Smith & Nephew
Plasmacup SC	Braun
IP X-LINKed acetabular cup	Waldemar Link
Exceed ABT Cup	Zimmer Biomet
EcoFit Cementless	Implantcast
Cenator	Corin

aneXys	Mathys
ANA.NOVA cup	ArtiQO

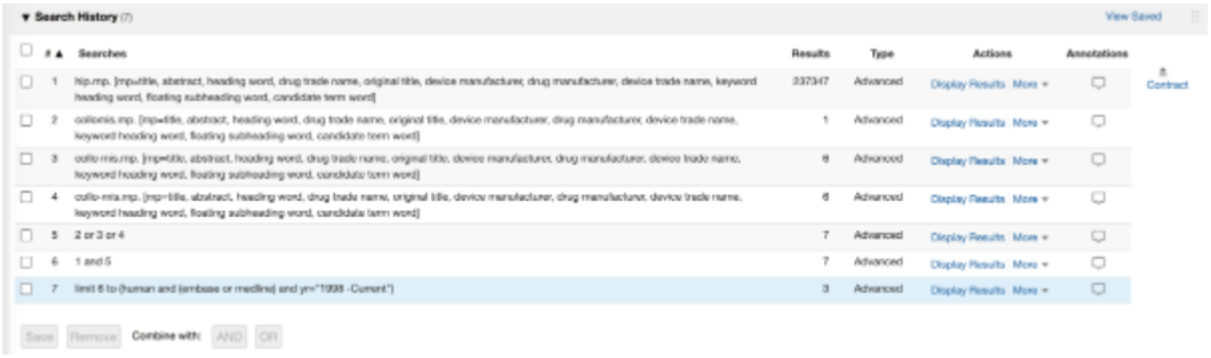
Knee systems:

LCS Complete	DePuy Synthes
NexGen CR	Zimmer Biomet
ACS Unc, Unicondylar	Implantcast
balanSys CR	Mathys
Logic RBK	Exactech
Optetrak CR	Exactech
Sigma High Performance Partial Knee	DePuy Synthes
TREKKING CR	Samo
Vanguard CR	Zimmer Biomet
Innex Gender	Zimmer Biomet

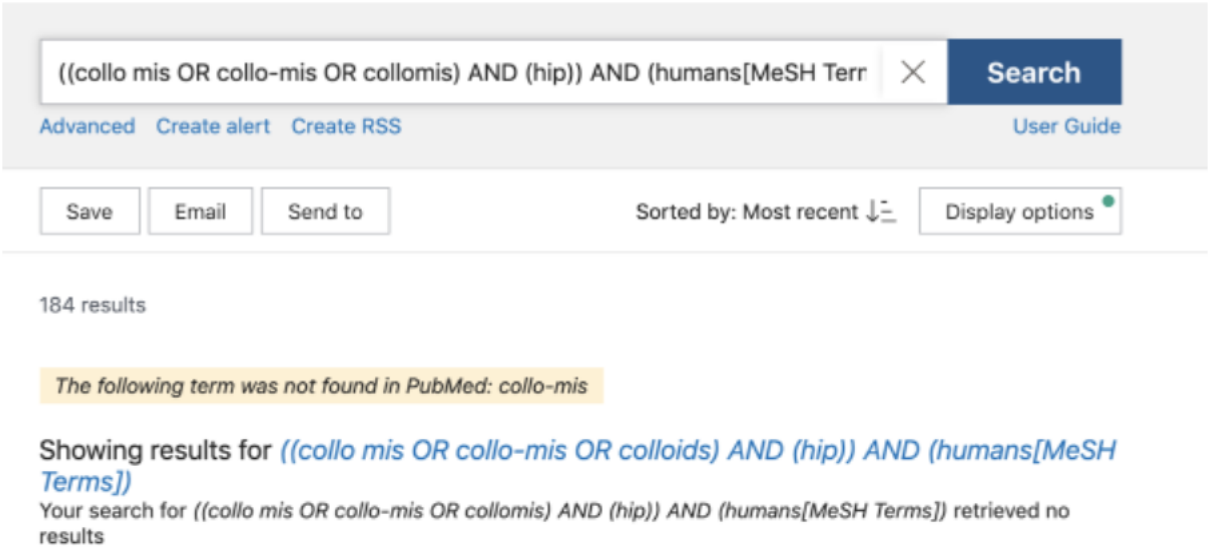
Appendix II: example searches

Example searches for the "Collo-mis" device are below:

Embase:



PubMed:



Web of science: