The what, when and how of orthopaedic registers
An introduction into register-based research
Varnum, Claus; Pedersen, Alma Bečić; Gundtoft, Per Hviid; Overgaard, Søren

Published in:
EFORT Open Reviews

DOI:
10.1302/2058-5241.4.180097

Publication date:
2019

Document version
Final published version

Document license
CC BY-NC

Citation for published version (APA):

Terms of use
This work is brought to you by the University of Southern Denmark through the SDU Research Portal. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:
• You may download this work for personal use only.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk
Establishment of orthopaedic registers started in 1975 and many registers have been initiated since. The main purpose of registers is to collect information on patients, implants and procedures in order to monitor and improve the outcome of the specific procedure.

Data validity reflects the quality of the registered data and consists of four major aspects: coverage of the register, registration completeness of procedures/patients, registration completeness of variables included in the register and accuracy of registered variables.

Survival analysis is often used in register studies to estimate the incidence of an outcome. The most commonly used survival analysis is the Kaplan–Meier survival curves, which present the proportion of patients who have not experienced the defined event (e.g. death or revision of a prosthesis) in relation to the time. Depending on the research question, competing events can be taken into account by using the cumulative incidence function. Cox regression analysis is used to compare survival data for different groups taking differences between groups into account.

When interpreting the results from observational register-based studies a number of factors including selection bias, information bias, chance and confounding have to be taken into account. In observational register-based studies selection bias is related to, for example, absence of complete follow-up of the patients, whereas information bias is related to, for example, misclassification of exposure (e.g. risk factor of interest) or/and outcome.

The REporting of studies Conducted using Observational Routinely-collected Data guidelines should be used for studies based on routinely-collected health data including orthopaedic registers.

Linkage between orthopaedic registers, other clinical quality databases and administrative health registers may be of value when performing orthopaedic register-based research.

**Keywords:** orthopaedic registers; data validity; survival analyses; bias and confounding; guidelines

Cite this article: **EFORT Open Rev** 2019;4 DOI: 10.1302/2058-5241.4.180097

**Introduction**

Established in 1975, the Swedish Knee Arthroplasty Register was the first nationwide orthopaedic register and during the following 20 years, arthroplasty registers were established in other Nordic countries. The National Joint Registry for England, Wales, Northern Ireland and the Isle of Man (NJR) and the Dutch Arthroplasty Register were established in 2003 and 2007, respectively. Outside of Europe, other important arthroplasty registers include the Australian Orthopaedic Association National Joint Replacement Registry, which was established in 1999. Some arthroplasty registers contain information on primary and revision procedures for several joints, e.g. the Dutch Arthroplasty Register and the NJR, whereas other registers are specific for a given joint, e.g. the Swedish Hip Arthroplasty Register and the Danish Shoulder Arthroplasty Register. Moreover, there are other nationwide orthopaedic registers than arthroplasty registers. For instance in Norway, data on hip fractures are reported to the Norwegian Hip Fracture Register. The main purpose of registers is to collect information on patients, implants and procedures in order to monitor and improve the patient course and outcome of the specific procedure. During the last ten years we have seen development and implementation of disease- and procedure-specific, evidence-based quality indicator sets in registers, and outcomes have been subjected to national, regional and department specific clinical auditing. Annual reports from the registers give an overview of the data and quality indicators, disclosed with health professional interpretations and recommendations for improvement of quality. The quality indicators are used for hospital benchmarking.
In addition, data in orthopaedic registers have the potential to answer a number of clinical research questions which are very important when choosing the optimal study design. The ‘benchmark’ design for investigating the prognostic effect of surgery, implants or other treatment factors is randomized clinical trials (RCTs). Unfortunately, well-designed and -powered RCTs are often costly in both time and money. In addition, often RCTs are too small and include healthy and younger patients, which decrease generalizability. Moreover, RCTs focus on short-term follow-up, are performed in ideal conditions not necessarily applicable in everyday clinical practice, are not able to study rare outcomes and are only able to study one to two interventions at the same time. All of these issues reduce the reliability of data from RCTs to inform clinical decision-making in a practical setting. Studies conducted from existing registers are an alternative to RCTs with the advantages of already collected data with a large study population, making it possible to study rare interventions and rare outcomes at low cost and quickly. In addition, observational register studies are suitable for studying non-modifiable risk factors such as underlying hip disease, body mass index (BMI), age, etc since patients cannot be randomized to these factors. Therefore, nationwide, population-based cohort studies are often used to investigate outcomes after arthroplasty surgery.9,10 However, it is possible to conduct register-based RCTs which are better powered studies than traditional RCTs and with higher cost-effectiveness.11

In several countries, every citizen is given a unique personal identification number that allows unambiguous linkage between arthroplasty registers and other administrative health registers and almost complete follow-up of each patient until a defined outcome, end of study, emigration or death. However, observational register-based studies also have important limitations, which may not be neglected. Limitations are related to data quality, possibility to control for confounding, missing and erroneous registration of data and methods of ascertainment of outcome. Further, observational studies are not able to distinguish causal associations from associations that are derived from bias or random error.

The aim of this paper is to focus on some basic epidemiological concepts that researchers should be aware of when starting a research project using register data. These concepts include basic knowledge on the data validity, basic statistical methods and basic issues that have to be taken into consideration when interpreting the results from observational register-based studies in addition to how to report the study.

Data validity

Before deciding on which register to use for a research question, it is important to get familiar with validity of the register. The validity consists of four major aspects including coverage of the register, registration completeness of procedures/patients, registration completeness of variables included in the register and accuracy of registered variables. In several countries, data are reported to more than one register, for example, in the Nordic countries orthopaedic departments are reimbursed from the authorities when reporting to national patient registers which can be considered ‘gold standard’ as the reimbursement is a motivation for reporting for the departments. It is possible to compare some data in the orthopaedic registers with that of the national patient registers by use of the unique personal identification number, and the comparison is used to estimate the coverage and completeness of the orthopaedic register. The coverage is defined as the number of departments reporting to the register out of the total number of departments within a given region or country. The completeness is defined as the proportion of procedures – either primaries or revisions – reported to the orthopaedic register compared with registered procedures in the national patient register and/or the orthopaedic register, for example, in Denmark, it is compulsory to report to the Danish Hip Arthroplasty Register (DHR) which is why both its coverage and completeness are high.7

Not only completeness is important but also the completeness of registration of all other information (variables), which has to be reported to the specific register; like diagnosis, type of components and their fixation, duration of surgery and perioperative complications. Further, the registration of all these variables have to be correct and in accordance with the clinical information on the patient and the surgical procedure. In 2004, the positive predictive value (PPV) of the registered primary diagnoses in patients undergoing primary total hip arthroplasty (THA) in the DHR was assessed using the review of medical records and radiographs as a reference, which is often referred to as the ‘gold standard’. The PPV was calculated as the probability of the registered diagnosis in the register that could be confirmed after review of medical records and preoperative radiographs. For instance, the primary osteoarthritis diagnosis in the DHR could be confirmed in 85% of patients undergoing primary THA.12 However, the PPV of the fresh hip fracture diagnosis in DHR was only 30%. If the study population included osteoarthritis patients from the DHR, we can be quite sure that we only have osteoarthritis patients. On the other hand, if we identify hip fracture study population in the DHR, we will include about 70% patients with diagnosis other than fresh hip fracture such as sequelae after hip fracture. In the Norwegian Arthroplasty Register, the unique catalogue numbers of implants are used to ascertain information on the components used during primary or revision hip arthroplasty. Regarding specific causes of revision, only
revision of a THA due to prosthetic joint infection (PJI) has been validated in the DHR and the PPV was 77%, but this value increased to 98% when combining data from the DHR and Danish microbiology databases.\textsuperscript{13} Completeness of registration of revision due to PJI in the DHR is as low as 60% compared with the ‘true’ infection risk defined from several data sources.\textsuperscript{13,14} To our knowledge, no other revision causes of a THA have been validated.

### Statistical considerations

In register studies, survival analysis is often used either to describe the incidence of an outcome by calculating absolute risk estimates, for example, absolute risk of revision or infection following the primary operation, or to compare the risk of an outcome between two different groups of patients.

The most commonly used survival analysis is the Kaplan–Meier estimator.\textsuperscript{15} The Kaplan–Meier estimator is functional as it allows for incomplete follow-up of patients, which occurs when patients are either lost to follow-up or are alive when the study ends without having experienced the outcome. In analyses of data from arthroplasty registers, the time intervals analyzed may represent the survival of implants, where the starting point is the date of the primary operation and the endpoint is the date of revision. Kaplan–Meier survival curves or plots present the proportion of patients who have not experienced the defined event (e.g. death or revision of prosthesis) in relation to the time.

In some cases a different event can preclude the outcome that is being studied; this is often referred to as a competing event, e.g. if the patients dies this precludes the revision of a THA.\textsuperscript{15} In cases of a competing event it can sometimes be advisable to use the cumulative incidence function. Whether to use the Kaplan–Meier estimator or the cumulative incidence function depends on the research question.\textsuperscript{16}

Both the Kaplan–Meier estimator and the cumulative incidence function are based on a number of assumptions, which have to be fulfilled in order to get unbiased estimates. One of the assumptions is independence of observations. In orthopaedic registers, patients can be recorded twice as both sides of the patients can be operated on, i.e. bilateral THA. Hence, the assumption of independence can be violated. Several advantageous methods of dealing with this violation of the independence assumption exist, but studies have shown that the results are only marginally different from results obtained when bilateral observations are treated as independent observations, especially if the outcome is revision surgery and study populations are large.\textsuperscript{15} Another issue with bilateral observations is the risk of erroneous reporting of the side of operation on either a primary operation or the subsequent revisions render a linkage between the two impossible. This issue can be difficult to account for unless several registers and data sources are available.

In observational studies, such as register studies, there may be systematic differences between groups of patients with different types of implants, and these systematic differences may affect the validity of the results by confounding. Cox regression analysis is a statistical model, which is used to analyze survival data taking these differences between comparison groups into account. For instance, the model can compare two types of prosthesis, calculating the hazard ratios with 95% confidence intervals (CIs). The 95% CI means that if we repeat the data collection and analyses many times, the 95% CI will include the correct value of measurements in 95% of the cases. CIs indicate to which extent random variation can explain the registered survival and is closely connected with the number of operations being part of the analysis. A wide CI indicates that there is a considerable uncertainty about the real prosthesis survival, while, on the contrary, to a lesser extent, a narrow interval indicates that the prosthesis survival can be interpreted as a result of random variation. Hazard ratio expresses the effect of each variable included in the Cox model in relation to the reference group, adjusted for other variables (confounders) in the model. In case of two implant groups in relation to revision as outcome, we calculate hazard ratio to compare the survival for patients included in two prosthesis groups. If the hazard ratio is 1.00 there is no difference in the incidence of revision when the two patient categories are compared. On the other hand, a hazard ratio < 1 will indicate that the incidence of revision in a given patient category is lower than the incidence in the reference category. In cases where the stated 95% CIs for hazard ratio do not include 1.00 it can be concluded that the given category of patients has an incidence of revision which differs from the reference category and that this difference probably cannot be explained by random variation. In other words, there exist statistically significant differences at the 0.05-level. On the other hand, if the 95% CIs include 1.00 it is not possible to determine whether the incidence is different in the two categories. In the interpretation of a study’s results it is of importance not only to consider whether a difference is statistically significant, but also the clinically relevant difference. A statistically significant difference might be so small that it has no clinical importance or the cost or risk of other complications exceed the benefits.

Besides erroneous reporting as the example stated above, studies using register data often have to address the problem of missing values. In a register study, missing values are the result of either incomplete registration or loss to follow-up, and they have a high impact on the estimates of incidences of, for example, revision burden. Missing data on account of incomplete registration are
from childhood hip disorders. These differences might be reserved for young and active patients as recommended by some authors, or some departments might have CoC as their ‘standard’ bearings, whereas other departments might reserve these bearings for only very young patients. Different methods have been developed to handle missing data. The simplest way of dealing with missing data is deletion of observations with missing data. This deletion can be either pairwise deletion or list-wise deletion. A more advanced and probably also more correct way is by either single or multiple imputation.

**Interpretation of data**

Several factors have to be taken into consideration when interpreting the results from observational register-based studies before inferring a causal association. These factors include selection problems potentially leading to selection bias, information problems potentially leading to information bias, chance and confounding. In this section, these considerable factors will be illustrated based on a study by Varnum et al published in 2015: the aim was to compare cementless THA with ceramic-on-ceramic (CoC) bearings with metal-on-polyethylene (MoP) bearings in patients identified from the DHR. This was a nationwide population-based cohort study including 1773 patients with CoC THA and 9323 patients with MoP THA which were followed until revision, end of study, emigration or death (whichever came first). The outcome was revision of the THA. No statistically significant difference in the risk of revision for any reason was found for CoC and MoP bearings after nine years of follow-up.

**Selection bias**

In general, selection problems in a cohort study can occur due to loss to follow-up. However, in the study example there was complete follow-up of all patients included. Thus, selection bias was not likely. Another selection problem could have occurred because the use of CoC bearings might be reserved for young and active patients as recommended by some authors, or some departments might have CoC as their ‘standard’ bearings, whereas other departments might reserve these bearings for only very rare cases, for example, very young patients suffering from childhood hip disorders. These differences might reflect surgeons’ preferences, the ‘culture’ in a department for using these bearings and socioeconomic circumstances, and might result in better outcomes for patients treated in hospitals and by surgeons with greater experience with the specific bearings. These selection problems are referred to as confounding by indication.

**Information bias**

In register-based cohort studies, information problems can occur due to misclassification of exposure (CoC or MoP bearings) or outcome (revision). If misclassification of exposure is dependent of misclassification of outcome (differential misclassification), the results may be influenced by information bias. We may have misclassification of both exposure and outcome, but if these were independent of each other (non-differential misclassification), the relative risk (RR) estimates would go towards the null hypothesis (no difference in RR of revision in the above-mentioned study).

Misclassification of bearings can occur, if data are missing or registered incorrectly. In the DHR, data on bearings, implant design, femoral head size and other causes of revision other than PJI are not validated which might give rise to concerns related to the quality of these data. In the above-mentioned study, misclassification can be related to the registration of a couple of bearings. However, due to the prospective registration of data in DHR, the misclassification of cause of revision was unlikely to be related to the registration of the type of bearings for primary THAs. The resulting non-differential misclassification might produce bias towards the null hypothesis.

**Chance**

Chance, or random error, is inherent in all observations and is related to the statistical precision of an estimate. This is expressed as a CI representing the range of values that is likely to include the true value. Statistical precision increases with the statistical power of the study, which is dependent of the sample size. The above-mentioned study included a large cohort of patients resulting in increased precision of the estimates.

**Confounding**

For confounding to occur, the following three conditions must be present:

- the confounding factor must be associated with both the exposure (CoC or MoP bearings) and the outcome (revision);
- the confounding factor must be distributed unequally among the groups being compared;
- a confounder cannot be an intermediary step in the causal pathway from exposure to outcome.

In a study by Johnsen et al based on DHR data, male patients had a 20% higher RR of any revision compared with female patients, and patients younger than 60 years had increased RR of revision after 0.5-year follow-up. Diagnosis was found to be a time-dependent predictor, although no difference in RR of revision was found for any diagnosis...
after 0.5-year follow-up. As sex, age and diagnosis was also distributed unequally between the CoC and MoP groups the definition of confounding was fulfilled, and adjustments in the statistical analyses were made for these three patient-related confounders in order to eliminate the confounding effect on the results. Among the surgery-related factors, the fixation technique has been shown to influence the risk of revision. The confounding effect of fixation is eliminated in the above-mentioned study, since only cementless fixation was used in the included patients.

Although adjusting for several patient- and treatment-related confounders, our study can still be biased by unmeasured confounding. BMI and THA due to osteoarthritis may be associated, and BMI > 35 kg/m² has been found to be a predictor for revision due to PJI. However, information on height and weight was not registered in the DHR at that time period and therefore not adjusted for the analyses, leading to potential unmeasured confounding. Treatment-related prognostic factors potentially leading to confounding include surgical approach (worse scores on Hip Disability and Osteoarthritis Outcome Score were reported after lateral approach than after posterior approach, and lateral approach was shown to increase the risk of revision due to aseptic loosening and decrease the risk of revision due to dislocation); and type of polyethylene as both conventional and highly crosslinked polyethylene have been included (the use of highly crosslinked polyethylene reduces polyethylene wear substantially). The organization-related prognostic factors, which may result in confounding, include: hospital volume (hospitals operating ≤ 50 procedures per year had an increased risk of revision after two-, five-, ten- and 15-year follow-up); set-up including fast-track; and surgeon’s skills including learning-curve and positioning of components.

Another type of confounding to take into account is residual confounding. In our study, we adjusted for diagnosis, and the most common diagnosis was osteoarthritis. It would have been possible to make several groups of patients with osteoarthritis according to the severity of the disease – e.g. mild, moderate or severe – by stratification. However, patients with osteoarthritis were categorized as one group, which may have resulted in residual confounding. Similarly, we adjusted for comorbidity using the Charlson comorbidity index, but we did not include information on comorbidities from general practitioners, or information on the severity of several diseases included in the index.

Further research perspectives

When defining a research question it is often that other types of data than those captured in orthopaedic registers are wanted or needed. However, linkage between orthopaedic registers, other clinical quality databases, and other administrative health registers is possible: Gundtoft et al linked data from the DHR to microbiology databases, the Danish National Patient Register and the Civil Registration System to obtain data on microbiology, comorbidity and vital status on all patients in a study on mortality after PJI in primary THA and Thillemann et al combined data from the DHR and the Danish Prescription Database in order to investigate the impact of statin use on the risk of revision after primary THA. In addition, both the DHR and the Danish Multidisciplinary Hip Fracture Register have been linked to the Danish Transfusion Database in order to study the association of transfusion with outcome after THA, as well as association between BMI and transfusion risk. Biochemistry data (including creatinine measurements) have been linked to the Danish Multidisciplinary Hip Fracture Register in order to study the risk of acute kidney injury after hip fracture and subsequent mortality. Moreover, when an intervention or outcome of interest is rare, data from one country may be insufficient, requiring combination of data from several countries. Such collaboration between registers is found in the Nordic Arthroplasty Register Association. The possibilities of data linkage and using data from collaborations shall be remembered and used when designing and conducting studies based on observational data from orthopaedic registers.

REporting of studies Conducted using Observational Routinely-collected Data (RECORD) guidelines

Many of the national orthopaedic registers are partly based on administrative routinely-collected data which is why the use RECORD guidelines should be used when reporting from these registers. These guidelines are an extension of the STrengthening the Reporting of Observational studies in Epidemiology (STROBE) guidelines, which still can be used when non-administrative databases are used. The goal of using RECORD guidelines is to enhance the transparency of the research project in order to understand the quality of the study. A number of requirements regarding a variety of items for all parts in a manuscript are presented in a checklist. The guidelines have been shown to enhance the quality of the manuscripts. Thus it is recommended to use either STROBE or RECORD guidelines.

AUTHOR INFORMATION

1Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Odense, Denmark.
2Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark.
3Department of Orthopaedic Surgery and Traumatology, Odense University Hospital, Odense, Denmark.
4Department of Orthopaedic Surgery, Vejle Hospital, Vejle, Denmark.
England and Wales.

metal-on-metal hip replacements: analysis of data from the National Joint Registry of


