Feasibility of Using Common Data Model for Orthopedic Research: Analysis of Risk Factors for Periprosthetic Joint Infection after Total Joint Arthroplasty

Young Ju Chae
Seoul National University Bundang Hospital
Jin Ho Shin
Seoul National University Bundang Hospital
Sung Jae Jung
Seoul National University Bundang Hospital
Hyun Sik Gong (hsgong@snu.ac.kr)
Seoul National University Bundang Hospital, Seoul National University College of Medicine
https://orcid.org/0000-0003-4028-1559

Research article

Keywords: Total joint arthroplasty, Revision, Periprosthetic joint infection, Risk factors, Common data model

DOI: https://doi.org/10.21203/rs.3.rs-35258/v1

Abstract

Background: Common data model (CDM) is a standardized data structure defined to efficiently use different sources in hospitals. A study using the CDM is scarce for orthopedic outcome researches due to the complexity of variables. We aimed to test the feasibility of applying CDM in the orthopedic field and analyzed risk factors for periprosthetic joint infection (PJI) after total joint arthroplasty (TJA) using CDM.

Methods: We undertook a retrospective cohort study of all primary and revision hip and knee TJAs at our institution from January 2003 to October 2017. We identified potential risk factors for PJI after TJAs in the literatures, which included preoperative demographic/social factors, previous medical history, intraoperative factors, laboratory results and others. The data sourced from EMR was extracted, transformed, and loaded into CDM.

Results: Variables such as demographic/social factors, medical history and laboratory results could be converted into CDM, but the other known risk factors could not. In total, 12,320 primary hip and knee TJAs and 120 revision arthroplasties were identified. Among them, 34 revisions were done because of PJI. Risk factors of PJI were hypertension and urinary tract infection after total hip arthroplasty, and age (70-79 years), male sex, anemia, steroid use, and urinary tract infection after total knee arthroplasty.

Conclusions: This study demonstrates that orthopedic outcome researches using CDM is feasible although data converting to CDM was possible for limited factors. Further data transforming technologies need to be developed to analyze more factors relevant to orthopedic area, such as intraoperative factors and imaging findings.

Background

Research using medical information has been actively carried out through the development of the information technology (IT). Commonly, electronic medical records (EMR) or administrative claims databases have been widely used for observational studies of clinical data. However, inconsistent data formats make large-scale clinical research collaboration between hospitals difficult and take a lot of time and effort. Thus, the need for standardization of EMR data is considered important in the medical field. The development of standard clinical information models is an attempt to tackle the storage and exchange of clinical data. Some researchers have shown that analyzing EMR data using standard-based methods is economical and improves efficiency [21, 25]. Common Data Model (CDM) allows for the systematic analysis of disparate observational databases. The concept is to transform different data into a standardized common data format by coding schemes and terminologies.

Total joint arthroplasty (TJA) is a commonly performed orthopedic procedure that can improve quality of life in patients with advanced arthritis. Over the past two decades, the number of TJAs has increased exponentially [7]. However, periprosthetic joint infection (PJI), which is the most serious complication of TJA, can result in severely limited joint function and increased mortality. Many studies have attempted to identify risk factors for PJI, which include rheumatoid arthritis, diabetes, renal disease, depression, hypercholesterolemia, anemia, urinary tract infection, hypertension, age, male, obesity, smoking, steroid use, blood transfusion, prolonged operative time, wound problem, and malnutrition [1, 2, 4, 6, 8–10, 15–17, 20, 24, 26–28, 30, 34]. However, only a few of them have considered multiple risk factors [2, 4, 6, 8, 27, 34]. Furthermore, results obtained from different studies examining the same risk factor have reported conflicting results [29].

Although several studies have been performed for orthopedic outcome researches using EMR or administrative claims databases, studies using CDM is scarce due to the complexity of variables in the orthopedic field. Therefore, we wanted to test the feasibility of applying CDM in the orthopedic research, especially for evaluation of risk factors of PJI. As such, the purposes of this study were, 1) to apply standard CDM methods and algorithms to an observational orthopedic research, and to identify problems in converting EMR parameters into CDM, and 2) to evaluate risk factors of PJI when analyzed by CDM.

Methods

Patients

We obtained approval from Institutional Review Board (IRB) for this retrospective review of medical records. We included patients who underwent primary TJAs (hip and knee) from January 2003 to October 2017 at our institution, which is a referral, training hospital located in an urban area in South Korea. We first identified cases of revision after TJA by using the International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) diagnosis codes in EMR. The reasons for revision were periprosthetic joint infection (PJI), loosening, prosthesis failure, periprosthetic fracture, osteolysis, and dislocation. Then we identified cases of PJI from the revision cases and focused on the risk factors of PJI (Fig. 1). The patients with prior primary or revision surgery at outside hospitals were excluded. Staged revision procedures for infection were counted only once.

Risk factors for PJI

We searched for all possible risk factors for PJI in the literatures[2, 3, 13, 29, 34] and collected the following clinical data: preoperative demographic/social factors (age and sex), previous medical history (comorbidities and drug history), intraoperative factors (operative time, oxygenation, preparation, skin closure, and blood transfusion), laboratory results (albumin, cholesterol, blood cell counts, inflammatory markers, etc.) and others (admission days, observation duration, and venous thromboembolism prophylaxis). Among these data, we identified items that could be converted to CDM to test the feasibility. All clinical data abstracted were at the time of primary TJA.

Conversion of EMR parameters to CDM

The algorithm that EMR data converted to CDM is as follows; mapping EMR to standard concepts, extraction-transformation-loading (ETL) of patient data into CDM, and evaluation of the CDM-based results [11](Fig. 2). The coding system used for diagnosis in the EMR is an ICD-10 code, whereas the standard concept of CDM for diagnosis is based on the Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT) [31]. Although most codes was mapped to CDM through the SNOMED-CT, the standard concepts of CDM for drug exposure was based on the RxNorm from the US National Library of Medicine for medications [19]. However, all of codes in EMR were not corresponded to mapping in the CDM version available at the time of this research. Therefore, we conducted to find the corresponding CDM code with the mapped code, which was then re-grouped to obtain the desired value from the code in the CDM. To ensure minimal grouping errors and minimal information loss, four authors reviewed the concept mappings and achieved agreement by consensus.

To perform mapped patient data into CDM, extraction, transformation, and loading (ETL) process is required. We performed the ETL process on five tables. Five core tables are involved in our data loading: person [demographic

characteristics], condition occurrence [medical history], measurement [laboratory results], procedure occurrence [procedures], and drug exposure [drug use] (Fig. 3).

To load patient data into CDM, we developed the ETL scripts as a form of Standard Query Language (SQL) according to study design, and performed them as actual ETL process (Fig. 4).

Statistical analysis

For categorical variables, chi-squared and fisher's exact test were performed using PJI as a dependent variable and each parameter as an independent variable. For blood results two variables (low vs. high) were generated based on the reference ranges used by our hospital laboratory and analyzed as categorical variables. Continuous variables were analyzed by using t-tests and categorical variables by chi-square tests. We performed logistic regression analysis to identify risk factors associated with PJI. Odd ratios (OR) and 95% confidence intervals (CI) were calculated using R package for Windows with the level of significance set at $P < 0.05$. We also calculated adjusted OR using the propensity score matching for age and sex to reduce selection bias. We used MatchIt package to conduct propensity score matching with a ratio of one to five.

Results

Patient characteristics

A total of 12,320 primary TJAs (4,758 total hip arthroplasties (THAs) and 7,562 knee total knee arthroplasties (TKA)) were identified from January 2003 and October 2017. The most common cause was osteoarthritis (THA, 1040; TKA, 7305). Others cause included inflammatory arthritis, avascular necrosis, congenital dysplasia, fracture, ankylosis, dislocation, instability. There were 120 revision surgeries (1%, 120/12,320) including 71 after THA (1.5%) and 49 TKA (0.7%). The most common causes for revision after THA were loosening (21 cases, 27.3%), followed by PJI (16 cases, 20.8%), periprosthetic fracture (16 cases, 20.8%), prosthesis failure (11 cases, 14.3%), dislocation (10 cases, 13%), and osteolysis (3 cases, 3.9%). For TKA, loosening (28 cases, 52.8%) was also the most common cause for revision, followed by PJI (18 cases, 34%), osteolysis (3 cases, 5.7%), prosthesis failure (3 cases, 5.7%), and periprosthetic fracture (1 case, 1.9%) (Table 1).

Table 1
Causes of primary and revision total hip and knee arthroplasty

Primary arthroplasty (n = 12320)				**Revision arthroplasty (n = 120)**			
Hip (n = 4758)		Knee (n = 7562)		Hip (n = 71)		Knee (n = 49)	
Diagnosis	Number (n)	Diagnosis	Number (n)	Diagnosis	Number (n)	Diagnosis	Number (n)
AVN	1648 (36.5%)	OA	7305 (92.8%)	Loosening	21 (27.3%)	Loosening	28 (52.8%)
Fracture	1063 (23.5%)	RA	366 (4.7%)	PJI	16 (20.8%)	PJI	18 (34%)
OA	1040 (23%)	Traumatic OA	76 (1%)	Periprosthetic fracture	16 (20.8%)	Osteolysis	3 (5.7%)
Inflammatory arthritis	186 (4.1%)	AVN	62 (0.8%)	Prosthesis failure	11 (14.3%)	Prosthesis failure	3 (5.7%)
RA	177 (3.9%)	Inflammatory arthritis	35 (0.5%)	Dislocation	10 (13%)	Periprosthetic fracture	1 (1.9%)
Congenital dysplasia	148 (3.3%)	Ankylosis	10 (0.1%)	Osteolysis	3 (3.9%)		
Dislocation	100 (2.2%)	Dislocation	7 (0.08%)				
Ankylosis	80 (1.8%)	Instability	6 (0.07%)				
Traumatic OA	41 (1%)	Fracture	3 (0.03%)				
Traumatic AVN	28 (0.6%)						
Instability	4 (0.1%)						
Gout	1 (0.02%)						

AVN, avascular necrosis; OA, osteoarthritis; RA, Rheumatoid arthritis; PJI, periprosthetic joint infection

*C*onversion of EMR parameters to CDM

EMR variables such as demographic/social factors (age, sex), previous medical history (comorbidities, drug history), laboratory results (albumin, cholesterol, blood cell counts, inflammatory markers, etc.) and admission days were converted into CDM. However, intraoperative factors (operative time, oxygenation, preparation, skin closure, and blood transfusion) and others (observation duration, and venous thromboembolism prophylaxis) could not be converted to CDM (Table 2).

Table 2
Conversion of EMR parameters into CDM

Category	Parameter	Conversion to CDM (Yes/No)	Obstacles to conversion
Demographic factors	Age	Yes	
	Sex	Yes	
Previous medical history	Comorbidities	Yes	
	Drug history	Yes	
Laboratory results		Yes	
Intraoperative factors	Operative time	No	No corresponding CDM code
	Oxygenation	No	No corresponding CDM code
	Preparation	No	No mapping of terminology code in EMR
	Skin closure	No	No EMR data
	Blood transfusion	No	No corresponding CDM code
Others	Admission days	Yes	
	Observation duration	No	No corresponding CDM code
	Venous thromboembolism prophylaxis	No	No mapping of terminology code in EMR

CDM, common data model; EMR, electronic medical records

Risk factors of PJI

Among the variables that could be converted to CDM, hypertension (OR, 4.6; 95% CI, 1.7–12.7; Adjusted OR, 9; 95% CI, 2.7–30.6) and urinary tract infection (OR, 14.6; 95% CI, 3.2–66.3; Adjusted OR, N/A) were found to be associated with PJI after THA.

For TKA, age bracket of 70 to 79 years (OR, 5.4; 95% CI, 1.9–15.2), male (OR, 5.7; 95% CI, 2.1–15.2), anemia (OR, 12.2; 95% CI, 2.8–54.1), steroid use (OR, 4.4; 95% CI, 1.3–15.1), and urinary tract infection (OR, 13.7; 95% CI, 3.1–

60.7) were found to be associated with PJI. We identified that anemia, steroid use and UTI were equally significant factors after adjustment (Table 3, 4, Additional file 1). However, there was no evidence of any significant associations of PJI with laboratory results (Additional file 2).

Table 3
Demographic factors as risk factors of PJI

Category	Risk factor	Hip				Knee			
		Revision (n = 16)	Non-revision (n = 4687)	P-value	Odds Ratio (95%CI)	Revision (n = 18)	Non-revision (n = 7513)	P-value	Odds Ratio (95%CI)
Demographic factors	Mean Age (years)	62.8	59.8			71.4	66.7		
	Age	2	18	0.384	1.6 (0.4 ~ 7.1)	1	1	0.532	1.4 (0.2 ~ 10.3)
	10–19	2	160	1		13	4	**0.000**	
	20–29	4	411	0.287	1.1 (0.2 ~ 4.8)	3	25	**0.002**	5.4 (1.9 ~ 15.2)
	30–39	3	585	1	1.9 (0.6 ~ 5.9)	1	310	1	
	40–49	2	762	0.752			2439		0.2 (0.1 ~ 0.6)
	50–59	3	775	0.732	0.8 (0.2 ~ 3.4)		3945		
	60–69		1011		0.6 (0.1 ~ 2.5)		780		0.5 (0.1 ~ 3.8)
	70–79		805				9		
	80–89		156		1.2 (0.3 ~ 4.3)				
	90–99		4						
	100–109								
	110–119								
	> 120								
	Male sex	9	2022	0.188	2.0 (0.7 ~ 56)	6	608	**0.002**	5.7 (2.1 ~ 15.2)

Boldface indicates statistical significance.

Table 4
Previous medical history as risk factors of PJI

Category	Risk factor	Adjusted hip group				Adjusted knee group			
		Revision (n = 16)	Non-revision (n = 80)	P-value	Odds Ratio (95%CI)	Revision (n = 18)	Non-revision (n = 90)	P-value	Odds Ratio (95%CI)
Previous medical history	HTN	8	8	**0.0006**	9 (2.7 ~ 30.6)	9	30	0.1789	2 (0.7 ~ 5.6)
	Anemia	1	0			2	0	**0.0264**	NA
	Steroid use		0			3	0	**0.0039**	NA
	UTI	2	0	**0.0263**	NA	2	0	**0.0264**	NA

HTN, hypertension; UTI, urinary tract infection

Boldface indicates statistical significance.

Detailed information on the Table 4 is given in Additional file 1.

Discussion

Rationale, Summary, Significance

We aimed to test the feasibility of applying common data model (CDM) in the orthopedic field and analyzed risk factors for periprosthetic joint infection (PJI) after total joint arthroplasty. Variables such as demographic/social factors, medical history, laboratory results and admission days could be converted into CDM, but the others such as intraoperative factors, observation duration, and venous thromboembolism prophylaxis could not be converted to CDM. When analyzed by using CDM, we found that hypertension and urinary tract infection were risk factors of PJI after THA, and age bracket of 70 to 79 years, male, anemia, steroid use and urinary tract infection were risk factors of PJI after TKA. This study demonstrates orthopedic researches using CDM is feasible although data converting to CDM was possible for limited factors.

Conversion of EMR parameters to CDM

The CDM is designed to include all observational data derived from the EMR to support the generation of reliable evidence [11, 25]. It is important to obtain what we want from the study by properly designing the algorithm with the parameters currently available in the CDM. Creating mappings the variable EMR data into the target CDM concepts is also crucial to improve patient data standardization [14, 22]. Thus, in previous studies, cohort studies have been mainly focused on the pharmacoepidemiological research as treatment of diseases and epidemiological analysis of deaths from certain diseases [11, 12, 19, 23, 32, 33]. In our case, we focused on parameters related to risk factors for PJI after TJA and constructed the algorithm directly using SQL, not through programs already created within the CDM, to achieve the desired results in our study. Of course, the code mapping

process was not easy. Four authors reviewed the code mapping, but because of incomplete concepts matching and difference between the coding systems, a little information loss was inevitable. In addition, the data in EMR are typically expressed in non-standard terms, and the textual variable values are often in free-style using different local expressions, we could not standardize these terms and the textual values into standard concepts in this study.

The main advantages of research using CDM is that such studies can be conducted on a larger scale, against lower costs, and within shorter time frames than traditional studies [5]. Also, it protects the privacy and security of patients in research because not the information of a certain patient but the information of a certain result is used in CDM tool [25]. In our study, to maintain patient confidentiality privacy and security, the original patient identifications were removed when the patient data were converted to the CDM. The CDM is also an important part of multi-organization collaborative research [19, 22]. Because each hospital has a different structure in patient information, it is necessary to cooperate with multiple hospitals to provide information for standardization of patient information through CDM tool. However, differences in data structures and coding system are still major barriers to standardize data in CDM tool [31].

Risk factors of PJI

In this study, hypertension was identified as a significant risk factor of PJI after THA, which is concurrent with some studies [1, 2, 6, 30]. The studies demonstrated that hypertension is associated with delayed wound healing following TJA.

Urinary tract infection (UTI) was a significant risk factor of PJI after both THA and TKA in this study. Usually, UTI is more common in women than in men and the reported prevalence of UTI in women undergoing primary TJA ranges from 5.1–36% [4, 6, 9, 26]. Therefore, symptomatic UTI should be treated before proceeding TJA.

We found an association between age and risk of revision, which is consistent with previous findings [5, 7, 17–21]. Although older patient age would seem to coincide with poorer nutritional status and thus elevated infection risk, some studies reported an increased risk of revision for relatively younger patients [7, 17, 18, 22].

This study found that a male sex was a significant risk factor of PJI after TKA, which coincides with some studies [3, 6, 15, 27, 29]. A study suggested that men can get a greater degree of surgical trauma and tissue necrosis than in women [27]. Also, men have a more active life-style than women after TJA. Therefore, differences in exercise volume can cause overuse differences after TJA, which may result in revision surgery.

In this study, preoperative anemia was also associated with risk factors of PJI after TKA. Anemia is usually associated with a patient’s poor nutritional status. Previous literatures have shown that primary TJA patients who have preoperative anemia are more likely to receive blood transfusions, which are associated with an increased risk of postoperative infection [2, 4, 8–10]. Therefore, patients should be preoperatively evaluated for causes of anemia, such as iron deficiency, and considered for recombinant human erythropoietin treatment in order to decrease the risk of PJI [8, 10].

We also found steroid use as a risk factor of PJI after TKA, which is consistent with previous literatures [2, 6, 18, 28]. The association between steroid use and PJI is likely to be mediated at least in part by impaired wound healing due to the anti-inflammatory and immuno-suppressive effects of steroids [20]. In addition, steroid use can

cause problems of calcium and vitamin D metabolism, zinc deficiency, and most importantly an accelerated bone mineral loss [16].

Limitations of study

There are several limitations to our study that must be noted. First, although the study objective was to utilize a CDM to identify risk factors of PJI after TJA, we couldn't analyze all of them that have been reported in the literature. We couldn't use non-matching EMR code in CDM. In our further study, we will continue improving the scalability of the converting variable data to CDM. Further data transforming technologies need to be developed to analyze more factors relevant to orthopedic area, such as intraoperative factors and imaging findings. Second, the subjects were from a single institution and our methodology has not been tested with other uses. The research of CDM designed for one use might lack credibility in terms of methodology. Therefore, the generalizability still needs to be confirmed. We will conduct subsequent research to use multi-center data for large-scale analysis and further validate our methods.

Conclusions

This study presents an approach to achieve semantic standardization among different clinical data sources by using CDM in the orthopedic field. Although data converting to CDM was possible for limited factors, we could propose reusable data transforming method. Therefore, it may differ for other uses and associated data element sets, but we consider that the methodology reported here can be applied to other researches in the orthopedic field.

Abbreviations

EMR: Electronic medical records; CDM: Common data model; TJA: Total joint arthroplasty; THA: Total hip arthroplasty; TKA: Total knee arthroplasty; PJI: Periprosthetic joint infection; ETL: Extraction-Transformation-Loading; SNOMED-CT: Systematized Nomenclature of Medicine Clinical Terms; SQL: Standard Query Language.

Declarations

Ethics approval and consent to participate

This study was approved by the Institutional Review Board of the Seoul National University Bundang Hospital. We have obtained the written informed consent for participation in the study from all participants.

Consent for publication

Not applicable.

Availability of data and materials

All relevant data are included in this manuscript. Additional data may be requested by contacting the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported by the Seoul National University Bundang Hospital Research Fund. The funding source had no role in study design, data collection, analysis or interpretation, or in writing the manuscript. The article processing fee would be expensed using the funding received.

Authors' contributions

YJC, HSG participated in the design of the study. YJC, JHS measured the data. YJC, SJJ, HSG were responsible for the statistical analysis of the study. All authors contributed to the writing of the manuscript.

Acknowledgments

We thank for the big data center in our hospital for their support, without which the present study could not have been completed.

Authors' information

Department of Orthopedic Surgery, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 300 Gumi-dong, Bundang-gu, Seongnam-si, Gyeonggi-do, 463-707, Korea

Hyun Sik Gong

Department of Orthopedic Surgery, Seoul National University Bundang Hospital, 300 Gumi-dong, Bundang-gu, Seongnam-si, Gyeonggi-do, 463-707, Korea

Young Ju Chae

Department of Big Data Center, Seoul National University Bundang Hospital, 300 Gumi-dong, Bundang-gu, Seongnam-si, Gyeonggi-do, 463-707, Korea

Jin Ho Shin, Sung Jae Jung

References

1. Ahmed AA, Mooar PA, Kleiner M, et al. Hypertensive patients show delayed wound healing following total hip arthroplasty. PLoS One. 2011;6:e23224.
2. Almustafa MA, Ewen AM, Deakin AH, et al. Risk Factors for Surgical Site Infection Following Lower Limb Arthroplasty: A Retrospective Cohort Analysis of 3932 Lower Limb Arthroplasty Procedures in a High Volume Arthroplasty Unit. J Arthroplasty. 2018;33:1861–7.
3. Bottle A, Parikh S, Aylin P, et al. Risk factors for early revision after total hip and knee arthroplasty: National observational study from a surgeon and population perspective. PLoS One. 2019;14:e0214855.
4. Bozic KJ, Lau E, Kurtz S, et al. Patient-related risk factors for periprosthetic joint infection and postoperative mortality following total hip arthroplasty in Medicare patients. J Bone Joint Surg Am. 2012;94:794–800.

5. Carey DJ, Fetterolf SN, Davis FD, et al. The Geisinger MyCode community health initiative: an electronic health record-linked biobank for precision medicine research. Genet Med. 2016;18:906–13.
6. Chen J, Cui Y, Li X, et al. Risk factors for deep infection after total knee arthroplasty: a meta-analysis. Arch Orthop Trauma Surg. 2013;133:675–87.
7. Edwards C, Counsell A, Boulton C, et al. Early infection after hip fracture surgery: risk factors, costs and outcome. J Bone Joint Surg Br. 2008;90:770–7.
8. Eka A, Chen AF. Patient-related medical risk factors for periprosthetic joint infection of the hip and knee. Ann Transl Med. 2015;3:233.
9. Friedman R, Homering M, Holberg G, et al. Allogeneic blood transfusions and postoperative infections after total hip or knee arthroplasty. J Bone Joint Surg Am. 2014;96:272–8.
10. Greenky M, Gandhi K, Pulido L, et al. Preoperative anemia in total joint arthroplasty: is it associated with periprosthetic joint infection? Clin Orthop Relat Res. 2012;470:2695–701.
11. Hong N, Zhang N, Wu H, et al. Preliminary exploration of survival analysis using the OHDSI common data model: a case study of intrahepatic cholangiocarcinoma. BMC Med Inform Decis Mak. 2018;18:116.
12. Hripcsak G, Ryan PB, Duke JD, et al. Characterizing treatment pathways at scale using the OHDSI network. Proc Natl Acad Sci U S A. 2016;113:7329–36.
13. Jasper LL, Jones CA, Mollins J, et al. Risk factors for revision of total knee arthroplasty: a scoping review. BMC Musculoskelet Disord. 2016;17:182.
14. Jiang G, Kiefer R, Prud'hommeaux E, et al. Building Interoperable FHIR-Based Vocabulary Mapping Services: A Case Study of OHDSI Vocabularies and Mappings. Stud Health Technol Inform. 2017;245:1327.
15. Johnsen SP, Sorensen HT, Lucht U, et al. Patient-related predictors of implant failure after primary total hip replacement in the initial, short- and long-terms. A nationwide Danish follow-up study including 36,984 patients. J Bone Joint Surg Br. 2006;88:1303–8.
16. Kamal R, Bansal SC, Khandelwal N, et al. Moderate zinc supplementation during prolonged steroid therapy exacerbates bone loss in rats. Biol Trace Elem Res. 2014;160:383–91.
17. Katz JNWE, Guadagnoli E, Liang MH, et al. Differencesbetween men and women undergoing major orthopedic surgery for degenerative arthritis.*Arthritis & Rheumatism.* 1994;37.
18. Kunutsor SK, Whitehouse MR, Blom AW, et al. Patient-Related Risk Factors for Periprosthetic Joint Infection after Total Joint Arthroplasty: A Systematic Review and Meta-Analysis. PLoS One. 2016;11:e0150866.
19. Lai EC, Ryan P, Zhang Y, et al. Applying a common data model to Asian databases for multinational pharmacoepidemiologic studies: opportunities and challenges. Clin Epidemiol. 2018;10:875–85.
20. Lee DM, Weinblatt ME. Rheumatoid arthritis. Lancet. 2001;358:903–11.
21. Martin-Sanchez FJ, Aguiar-Pulido V, Lopez-Campos GH, et al. Secondary Use and Analysis of Big Data Collected for Patient Care. Yearb Med Inform. 2017;26:28–37.
22. Overhage JM, Ryan PB, Reich CG, et al. Validation of a common data model for active safety surveillance research. J Am Med Inform Assoc. 2012;19:54–60.
23. Panaccio MP, Cummins G, Wentworth C, et al. A common data model to assess cardiovascular hospitalization and mortality in atrial fibrillation patients using administrative claims and medical records. Clin Epidemiol. 2015;7:77–90.

24. Park CH, Lee YK, Koo KH. Lower Urinary Tract Infection and Periprosthetic Joint Infection after Elective Primary Total Hip Arthroplasty. Hip Pelvis. 2017;29:30–4.
25. Park RW. A clinical research strategy using longitudinal observational data in the post-electronic health records era. *Journal of the Korean Medical Association.* 2012;55.
26. Parvizi J, Koo KH. Should a Urinary Tract Infection Be Treated before a Total Joint Arthroplasty? Hip Pelvis. 2019;31:1–3.
27. Pedersen AB, Svendsson JE, Johnsen SP, et al. Risk factors for revision due to infection after primary total hip arthroplasty. A population-based study of 80,756 primary procedures in the Danish Hip Arthroplasty Registry. Acta Orthop. 2010;81:542–7.
28. Peel TN, Dowsey MM, Daffy JR, et al. Risk factors for prosthetic hip and knee infections according to arthroplasty site. J Hosp Infect. 2011;79:129–33.
29. Prokopetz JJ, Losina E, Bliss RL, et al. Risk factors for revision of primary total hip arthroplasty: a systematic review. BMC Musculoskelet Disord. 2012;13:251.
30. Saleh K, Olson M, Resig S, et al. Predictors of wound infection in hip and knee joint replacement: results from a 20 year surveillance program. J Orthop Res. 2002;20:506–15.
31. Conversion and Data Quality Assessment of Electronic Health Record Data at a Korean Tertiary Teaching Hospital to a Common Data Model for
Yoon D, Ahn EK, Park MY, et al. Conversion and Data Quality Assessment of Electronic Health Record Data at a Korean Tertiary Teaching Hospital to a Common Data Model for Distributed Network Research. Healthc Inform Res 2016;22:54–8.
32. Zhang X, Wang L, Miao S, et al. Analysis of treatment pathways for three chronic diseases using OMOP CDM. J Med Syst. 2018;42:260.
33. Zhou X, Murugesan S, Bhullar H, et al. An evaluation of the THIN database in the OMOP Common Data Model for active drug safety surveillance. Drug Saf. 2013;36:119–34.
34. Zhu Y, Zhang F, Chen W, et al. Risk factors for periprosthetic joint infection after total joint arthroplasty: a systematic review and meta-analysis. J Hosp Infect. 2015;89:82–9.

Additional Files

The additional files for this article can be found as follows:

Additional file 1. Previous medical history and others related risk factors for revision due to PJI after primary TJAs.

Additional file 2. Laboratory results related risk factors for revision due to PJI after primary TJAs.

Figures

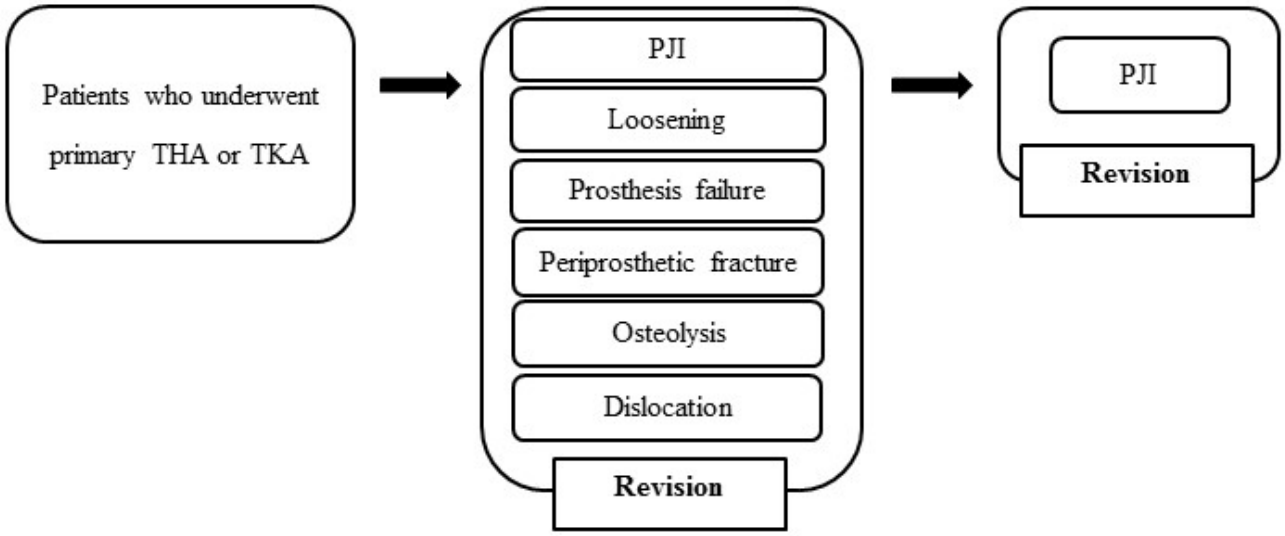

Figure 1

The patient selection flow chart.

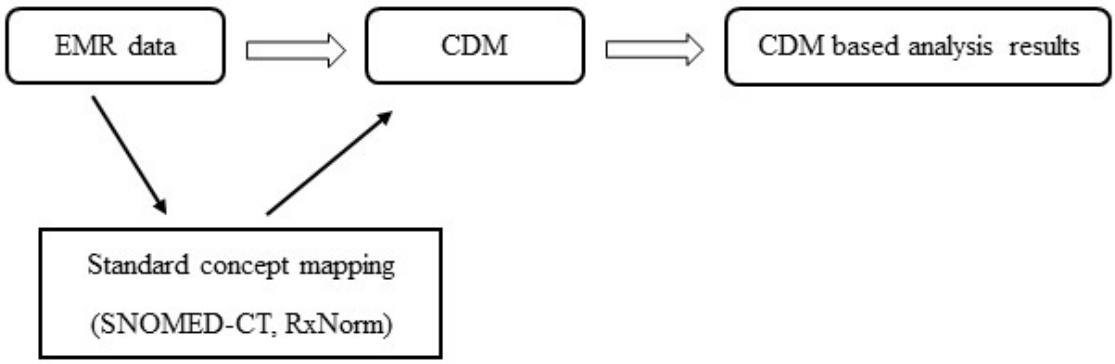

Figure 2

The algorithm that electronic medical records (EMR) data converted to common data model (CDM).

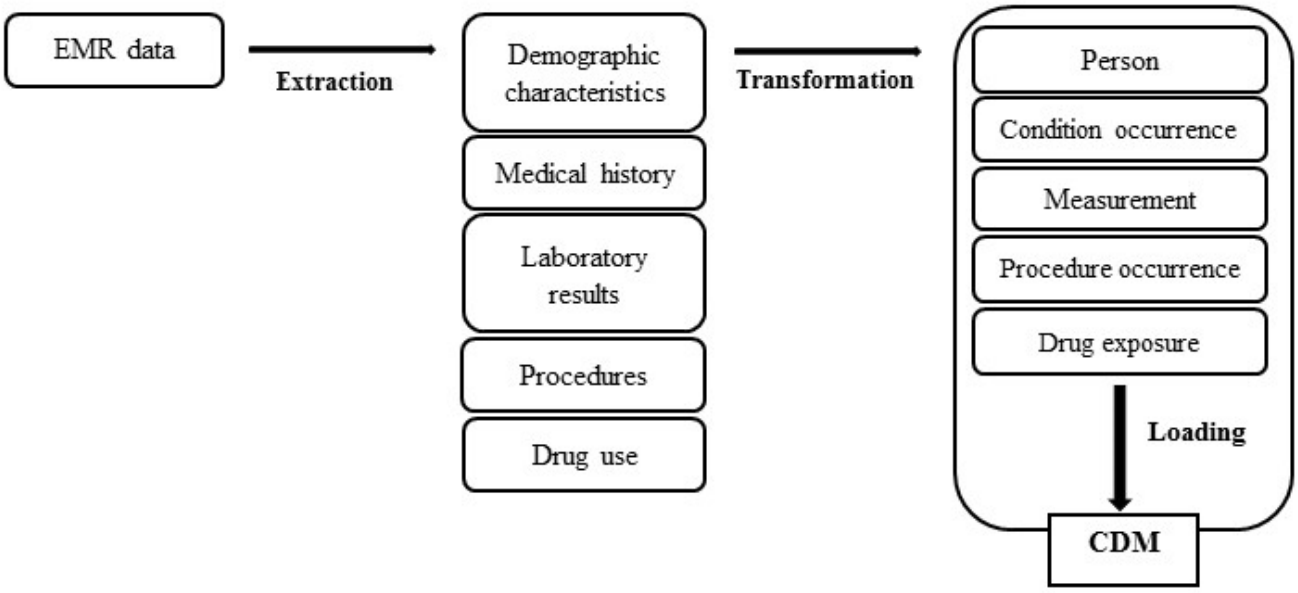

Figure 3

Extraction, Transformation, and Loading (ETL) process.

```
SELECT PROCEDURE_OCCURRENCE_ID, PERSON_ID,
      PROCEDURE_CONCEPT_ID, PROCEDURE_DATE,
      PROCEDURE_DATETIME, PROCEDURE_SOURCE_VALUE,
      MIN(REVISED_YN) AS REVISED_YN,
      MIN(HTN) AS HTN, MIN(RA) AS RA, MIN(Liver) AS Liver, MIN(Anemia) AS Anemia,
      MIN(Hyperchol) AS Hyperchol, MIN(Depression) AS Depression, MIN(Coagulation) AS Coagulation,
      MIN(Heart) AS Heart, MIN(Previous) AS Previous, MIN(DM) AS DM,
      MIN(Renal) AS Renal, MIN(Pulmonary) AS Pulmonary, MIN(FST_PROC_DATE) FST_PROC_DATE
 FROM ( SELECT  A.*
      ,LEAD(REV_GB,1) OVER (PARTITION BY A.PERSON_ID,NM ORDER BY A.PROCEDURE_DATETIME ASC) LEAD_REV_GB
      ,CASE WHEN REV_GB='PRIMARY' AND LEAD(REV_GB,1) OVER (PARTITION BY A.PERSON_ID,NM ORDER BY A.PROCEDURE_DATETIME ASC)='REVISION' THEN 'REVISED_Y' ELSE 'REVISED_N' END REVISED_YN
      ,CASE WHEN DIS_CLS_CD='HTN'          AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS HTN
      ,CASE WHEN DIS_CLS_CD='RA'           AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS RA
      ,CASE WHEN DIS_CLS_CD='Liver'        AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Liver
      ,CASE WHEN DIS_CLS_CD='Anemia'       AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Anemia
      ,CASE WHEN DIS_CLS_CD='Hyperchol'    AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Hyperchol
      ,CASE WHEN DIS_CLS_CD='Depression'   AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Depression
      ,CASE WHEN DIS_CLS_CD='Coagulation' AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Coagulation
      ,CASE WHEN DIS_CLS_CD='Heart'        AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Heart
      ,CASE WHEN DIS_CLS_CD='Previous'     AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Previous
      ,CASE WHEN DIS_CLS_CD='DM'           AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS DM
      ,CASE WHEN DIS_CLS_CD='Renal'        AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Renal
      ,CASE WHEN DIS_CLS_CD='Pulmonary'    AND CONDITION_START_DATE>=FST_PROC_DATE THEN CONDITION_START_DATE END AS Pulmonary
      ,FST_PROC_DATE
 FROM (select a.*,'REVISION' REV_GB from PROCEDURE_OCCURRENCE_REV a
        UNION ALL
       select a.*,'PRIMARY' REV_GB  from PROCEDURE_OCCURRENCE_PRI a) A
 LEFT OUTER JOIN (select a.* ,
                         ROW_NUMBER() OVER (PARTITION BY PERSON_ID, DIS_CLS_CD ORDER BY CONDITION_START_DATE ASC) RN
                  from  CONDITION_OCCURRENCE_OSSTUDY a) B ON A.PERSON_ID=B.PERSON_ID AND B.RN=1
 LEFT OUTER JOIN (select PERSON_ID,MIN(PROCEDURE_DATE) FST_PROC_DATE
                  from PROCEDURE_OCCURRENCE_PRI a
                 group by PERSON_ID) c ON A.PERSON_ID=c.PERSON_ID
WHERE REV_GB='PRIMARY' )
GROUP BY PROCEDURE_OCCURRENCE_ID
        ,PERSON_ID
        ,PROCEDURE_CONCEPT_ID
        ,PROCEDURE_DATE
        ,PROCEDURE_DATETIME
        ,PROCEDURE_SOURCE_VALUE;
```

Figure 4

ETL scripts as a form of Standard Query Language (SQL).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplement8.docx
- supplement9.docx