Real-World Data Analysis of Adverse Manifestations Attributable to Arthroplasty Implants

Xiao Fu, Philip J. Belmont Jr., Robert Elder, Enusha Karunasena, David Saylor, Yelizaveta Torosyan

Center for Devices and Radiological Health (CDRH), Food and Drug Administration (FDA), Silver Spring, MD

Abstract

Background: Various adverse events are reported with metal implants; however, their clinical manifestations and biological underpinnings remain unclear. We initiated a research effort on implant-associated manifestations employing real-world data (RWD) from electronic health records (EHRs).

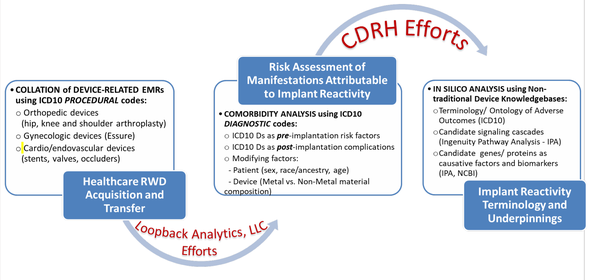
Objective: Outline the scope and risks of clinically consequential adverse manifestations attributable to arthroplasty implants.

Methods: A dataset of ~27,000 patients with large joint arthroplasties, including ~27 million diagnoses and ~9 million procedure records, was created using EHRs (Loopback Analytics, 2016 - 2019). Natural language processing (NLP) was used to link EHR-based surgical supply information to device-specific information (alloy chemistry) from regulatory submissions. Using ICD10 codes, comorbidity analysis was performed in cohorts stratified by arthroplasty types and Adverse Outcomes (AO) including Revision as well as patient demographics. Pre/post-implantation occurrence of 71 ICD10 diagnostic categories (pre-selected as immune/inflammatory conditions) was compared with respect to AO/Revision to identify potential comorbidities representing risk factors or underrecognized complications. Inter-cohort differences were assessed using chi-square test with odds ratios, relative risk ratios, time-to-event analysis, and multivariate regression. LASSO regression modelling using ~23,700 ICD10 diagnoses was used to build "unsupervised" prediction models for identifying risk factors/complications and modifying factors.

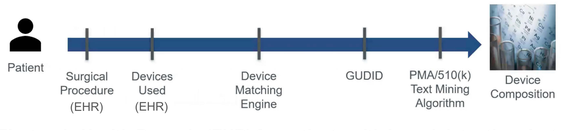
Results: Compared to Controls (recipients of large joint arthroplasty with no known arthroplasty-related complications), AO/Revision cohorts showed higher post-implantation frequencies for some immune/inflammatory conditions as arthroplasty-related complications, with likelihoods being further impacted by patient demographics and device materials.

Conclusion: Use of our transferable analytic/statistical methodology for pre-existing healthcare RWD analysis can provide insights into implant-related risk factors and complications, thus promoting the informed use and predictive evaluation of implants.

Overall Research Flow



Device-Patient Data Acquisition (Orthopedic Devices)



Electronic Health Records (EHR) for patients with large joint arthroplasty:

- ~27K subjects with hip, knee, or shoulder implants
- ICD10 codes used to characterize target arthroplasties as well as other comorbidities and procedures

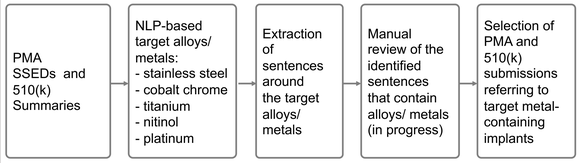
Device Matching Engine built to connect EHRs to FDA Medical Device Databases as follows:

- Standardize manufacturer names
- Apply standard formatting to part numbers
- Match EHR data to GUDID with standard manufacturer names and part #
- Match EHR data to PMA/510(k) with probabilistic matching rules using device names

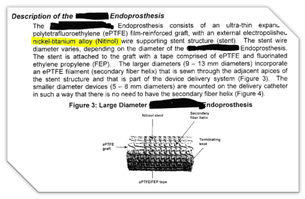
RWD Acquisition & Analysis Methodology with Respective Examples

Device Alloy Data Acquisition

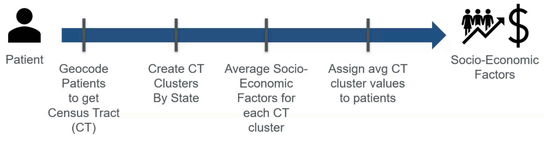
PMA/510(k) Text Mining Algorithm (Natural Language Processing - NLP) built to identify target alloys/ metals from >70K Premarket Approval (PMA) Summaries of Safety and Effectiveness (SSEDs) and 510(k) Summaries



The image below shows an example of NLP-based identification of nickel-titanium alloy Nitinol as one of device-related alloy/ metal targets (note: the acquired device composition data are not limited to arthroplasty):



Patient Socioeconomic Data Acquisition



An **algorithm** aimed to:

- Yield dataset that protects privacy but provides census tract level specificity on socio-economic factors
- Create clusters using variables such as a 3-digit zip code and a k-means model to group similar census tracts into groups of ≥20,000 inhabitants

Socio-economic factors included:

1. Median Household income
2. % receiving assisted income
3. % living below Poverty Level
4. % with at least high school education
5. % lacking health insurance
6. % houses that are vacant
7. Deprivation Index

Multivariate Logistic Regression Analysis

As a starting point, the cohort (~27K patients) with large joint arthroplasties was stratified by type of arthroplasty and by presence or absence of arthroplasty-related Revision and Adverse Outcomes such as periprosthetic osteolysis (AO+Rev and Controls, respectively). Multivariate logistic regression analysis was applied to assess the risk of certain *pre-selected* ICD-defined immune/ inflammatory conditions (n=71) with regards to AO+Rev and patient's sex and race. An example below shows the higher risk of M05 - M14 Inflammatory Polyarthropathies including Rheumatoid Arthritis in patients with Knee arthroplasty (n= 16,749), especially in Blacks and Females.

M05 - M14 Inflammatory Polyarthropathies

	Adj OR	95% CI	p-value
Race:			
Black	1.53	(1.41,1.65)	< 0.001
Other	0.69	(0.56,0.85)	< 0.001
White	Ref		
Sex:			
Female	1.0032	(0.94,1.07)	0.924
Male	Ref		
Outcome:			
AO+Rev	**1.67**	**(1.53,1.82)**	**< 0.001**
Control	Ref		

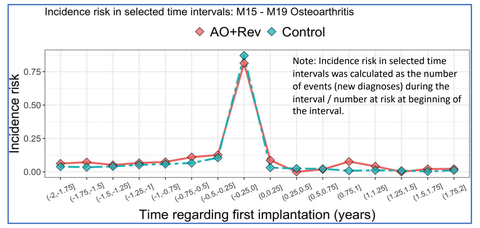
M05 - M06 Rheumatoid Arthritis

	Adj OR	95% CI	p-value
Race:			
Black	1.22	(1.06,1.39)	0.004
Other	1.21	(0.87,1.7)	0.258
White	Ref		
Sex:			
Female	1.84	(1.62, 2.1)	< 0.001
Male	Ref		
Outcome:			
AO+Rev	**1.5**	**(1.3,1.73)**	**< 0.001**
Control	Ref		

Time-To-Event Analysis using Kaplan-Meier Approach

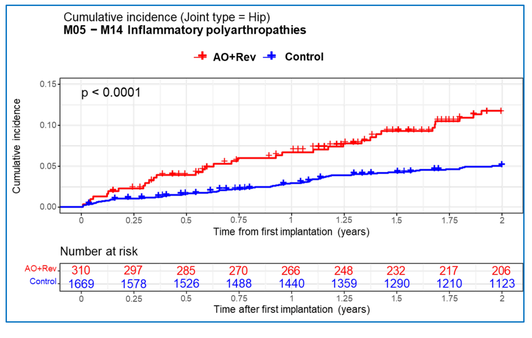
Following our hypothesis that comorbidities with higher AO+Rev vs. Control frequencies may represent potential risk factors or complications correlated with implant reactivity, we compared the incidences of pre-selected ICD10 diagnoses for immune/ inflammatory conditions in two study groups. First, ICD10-defined comorbidities in each subject were characterized based on their first appearance as: 1) *pre*-implantation diagnoses with dates *prior or on the same day* as first joint replacement procedure, and 2) *post*-implantation diagnoses with dates *after* first joint replacement procedure. In both AO/Revision and Control groups, the frequencies of most tested ICD10 codes peaked around the implantation time, likely reflecting a more thorough patient evaluation in this period.

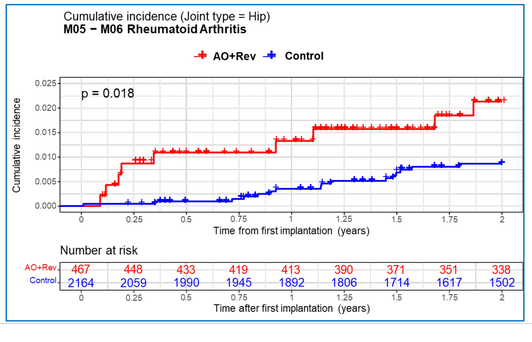
Similar incidence risks profiles for M15-M19 Osteoarthritis in both study groups, AO/Revision and Controls, were consistent with this diagnosis considered a common underlying condition and arthroplasty indication:



On the other hand, the higher incidences of Inflammatory Polyarthropathies and Rheumatoid Arthritis in the AO/Revision vs. Control subjects (not shown) suggested that these diagnoses may represent either arthroplasty-associated risk factors or adverse outcomes in the pre- and post-implantation periods, respectively.

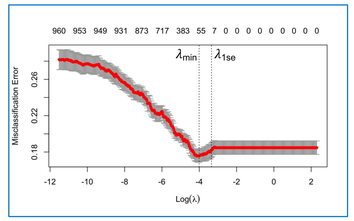
Next, we applied Kaplan-Meier based time-to-event analysis using 2-year *post*-implantation cumulative incidences, with the first appearance of selected diagnoses as failure variables and with the end of follow up (2021-01-01) or death as censored observations. As shown in the Figures below, inter-group differences between the Kaplan-Meier curves with *post*-implantation increases of cumulative Incidences in AO+Rev group suggested that these two diagnoses may represent adverse immune outcomes related to orthopedic implant reactivity:





Logistic LASSO Regression Analysis

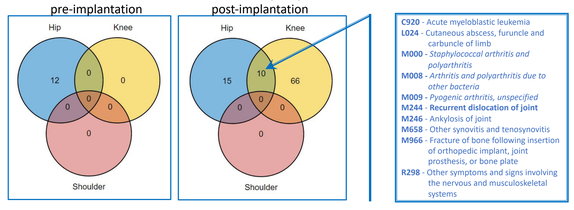
To complement our comorbidity analyses using pre-selected ICD10 diagnoses and to incorporate as candidate variables *all* diagnoses (per first 4 characters) from the main ICD10 diagnostic categories (https://www.icd10data.com/ICD10CM/Codes), we applied a logistic regression analysis with the Least Absolute Shrinkage and Selection Operator (LASSO) penalization to help reduce dimensions of ICD10 feature selection aimed at distinguishing between the AO/Revision and Control groups.



The Figure on the left presents an example of ICD10 code variable selection (regardless of pre- or post-implantation) using LASSO logistic regression in subjects with Hip arthroplasty. λ_{min} is the value of λ (lambda) that gives minimum cross-validation mean squared error and λ_{1se} is the value of λ that gives the most regularized model such that the cross-validated error is within one standard error of the minimum.

The Venn diagrams below show the numbers and overlap of LASSO-identified ICD10 features in different arthroplasty cohorts; the Table details the variables (n=10) shared by Hip and Knee Arthroplasty cohorts as *post-implantation* ICD10 features that may distinguish between AO/Revision and Control subjects in these two cohorts:

LASSO selected ICD10 features in different arthroplasty cohorts:



Per the LASSO-based coefficients and importance rankings of these ICD10 features, the *post-implantation* appearance of M244: Recurrent Dislocation of Joint (in bold), despite its relative rarity, was the top LASSO discriminator between AO/Revision and Control subjects in both Hip and Knee Arthroplasty cohorts. The AO/Revision subjects in these two cohorts also had much higher odds of post-implantation diagnoses of bacterial/ pyogenic arthritis (M000, M008, M009; italicized) and other joint/bone-related conditions such as periprosthetic fracture (M966) or ankylosis (M246).

In addition, LASSO regression analysis identified some *pre-implantation* ICD10 features (n=12; Venn diagram on the left) as potential risk factors for *post-implantation* AO and Revision in the Hip Arthroplasty cohort; however, none of these features was shared by the Knee Arthroplasty cohort. The Shoulder Arthroplasty cohort did not show any ICD10 features with statistically significant differences in their pre- or post-implantation appearance in AO/Revision subjects vs. Controls. Most importantly, none of the LASSO-identified features distinguishing between AO/Revision and Control groups in any of the Arthroplasty cohorts indicated pre/post-implantation diagnoses for (auto)immune/ inflammatory conditions as either *pre*-implantation predisposing factors or *post*-implantation manifestations of abnormal implant reactivity.

Conclusions

- **Our RWD acquisition and analysis approaches provide insights into implant-related *pre*-implantation risk factors and underlying conditions as well as *post*-implantation complications**:
  - Multivariate regression analysis reported an increased post-implantation occurrence of some infrequent immune/inflammatory diagnoses in AO/Revision subjects vs. Control, thus demonstrating a potential association between implant reactivity and conventional arthroplasty complications. The likelihood of a patient being diagnosed with a post-implantation immune/ inflammatory diagnosis, such as rheumatoid arthritis, may be further impacted by demographic risk factors.
  - On the other hand, LASSO regression analysis demonstrated the absence of systemic immune/inflammatory conditions among the generally scarce ICD10 features shared by different arthroplasties, thus underscoring the rarity of clinical manifestations that could be viewed as potential pre/post-implantation risk factors and outcome modifiers due to abnormal implant reactivity.
- **As a result, our RWD acquisition and analysis methodology can be reapplied to other healthcare RWD projects aimed to promote predictive evaluation and informed use of medical products in patient subpopulations.**