The Impact of Variable Days of Coverage on Estimating Diagnosis and Treatment Prevalence from Administrative Claims Data

MSR78

Pat Hilden DrPH, Kathryn Haglich MS, Ken Gross PhD, Sandy Balkin PhD
Strategy and Analytics
ROYALTY PHARMA

Objectives

Closed administrative claims data are a mainstay in HEOR and commonly used to estimate the annual prevalence of a diagnosis or treatment within the US population.

Prevalence estimation is often complicated by incomplete pharmacy and/or medical coverage for a subset of persons within a given year, which can be attributed to two possible reasons 1) The individual has changed health plans and their claims are no longer captured by the data provider 2) The individual died during the year. This missingness is inherently non-random and can occur with relative frequency in closed claims data.

When it is known that the incomplete enrollment is due to a change in health plan, then excluding these persons from prevalence estimation is a reasonable approach as projection methodologies may account for this missingness by increasing the associated weights. In contrast, when individuals die, excluding them from the prevalence estimation would result in over-estimation of true population prevalence.

In practice, the goal of prevalence estimation is to determine the population prevalence under the scenario where the only missing enrollment information is due to death, however closed claims data typically only note that a person was not enrolled at a given date and do not provide details as to why this was the case.

Given the above, it remains an open research question as to what extent persons with incomplete coverage can or should be included when estimating annual prevalence from closed claims data.

Using a simulation study, this analysis looks to assess the impact of including persons with variable amounts of annual enrollment on the accuracy of prevalence estimation.

Additionally, we highlight a real-world case study related to obstructive hypertrophic cardiomyopathy (HCM), a space of recent interest to our organization, and highlighted previously[1,2].

Simulation Methods

Figure 1 provides an outline of the simulation structure.

Closed payor claims, sourced from Inovalon and provided by HealthVerity, from 2022 are used for this analysis (data cohort, ~89M persons).

Using the data cohort, we identify the subset of persons with complete pharmacy and medical enrolment over the course of the year (analysis population, ~53M persons), and randomly assigned dates of death to a subset of persons using gender and age group specific rates representative of the US population.

Our goal is to estimate the prevalence in this analysis population as it reflects the situation where missing enrollment data is solely attributable to mortality.

Each simulation consists of sampling 1/3 of the analysis population (sample data, ~18M persons) and matching these persons 1:1 with randomly sampled persons from the larger data cohort. A sample size of 1/3 was selected as it is reflective of the fraction of the US population the data cohort represents for a given year.

For each 1:1 match, the observed (often incomplete) coverage patterns from the matched person are imposed on the sample data person, resulting in enrollment patterns which are representative of the larger data cohort, but for which true enrollment and claims are known during all periods (analysis data).

For each simulation, we evaluate scenarios with pharmacy and medical coverage minimums of 1+, 90+, 180+, 270+ and 365 days, and for each coverage minimum we weight the analysis data by gender and age group to estimate the population prevalence.

200 total simulations are run, and the percent error between the true prevalence in the analysis population and the estimated prevalence in the analysis data is used to compare between different coverage minimums.

For presentation purposes, we focus on estimating the prevalence of various diagnosis categories for diseases of the circulatory system (**Table 1**).

Figure 1

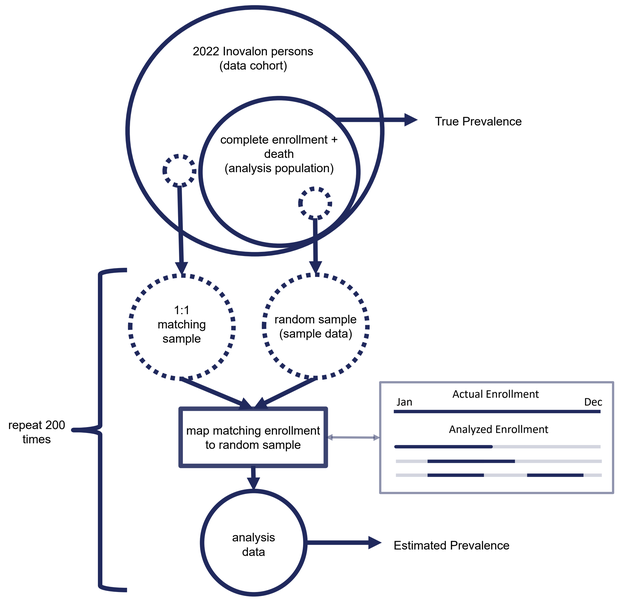


Table 1

Diagnosis Group	Description	Prevalence per 100,000
I00-I02	Acute rheumatic fever	8
I05-I09	Chronic rheumatic heart diseases	514
I10-I1A	Hypertensive diseases	16,780
I20-I25	Ischemic heart diseases	3,266
I26-I28	Pulmonary heart disease and diseases of pulmonary circulation	693
I30-I5A	Other forms of heart disease	5,653
I60-I69	Cerebrovascular diseases	1,765
I70-I79	Diseases of arteries, arterioles and capillaries	3,291
I80-I89	Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified	2,254
I95-I99	Other and unspecified disorders of the circulatory system	992

Simulation Results

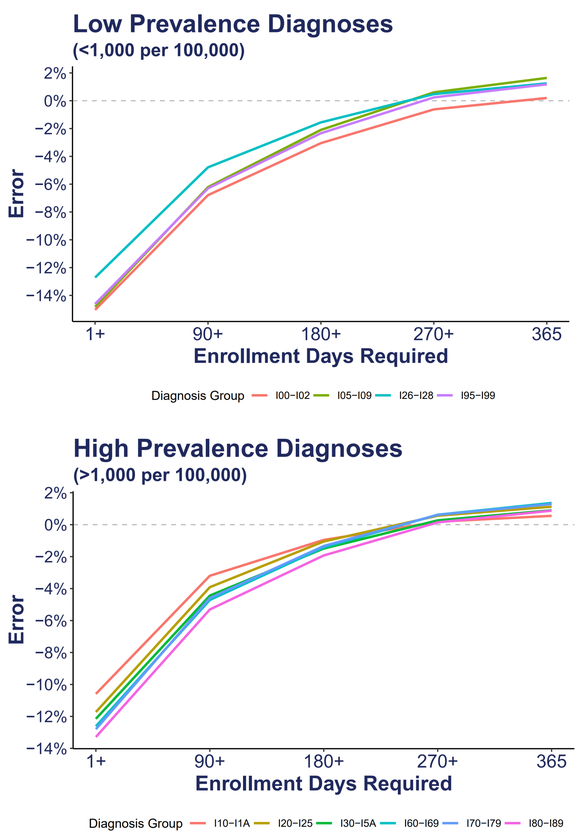
The true population prevalence in the analysis population varied from 8 (acute rheumatic fever) to 16,780 (hypertensive diseases) per 100,000 persons across the diagnosis groups of interest (**Table 1**).

Figure 2 highlights the simulation results with respect to percent error for each diagnosis category across the varying coverage minimums considered. Across both low prevalence diagnoses (top, <1,000 per 100,000) and high prevalence diagnoses (bottom, >1,000 per 100,000) scenarios which included persons with 1+, 90+, and 180+ enrollment days tended to under-estimate the true population prevalence, while including only persons with 365 days resulted in over estimation, across all diagnosis categories.

Using 270+ days as a threshold for inclusion tended to perform well across both low and high prevalence diagnoses as well as differing diagnosis groups.

As expected, diagnoses with high prevalence resulted in lower error rates as compared to diagnoses with low prevalence.

Figure 2



Conclusions & Limitations

Based on the simulation study results, including persons with 270+ days of annual enrollment in prevalence estimation tends to result in accurate estimation of the unknown population prevalence in the scenario where missing enrollment is only due to mortality.

Using a lower threshold for enrollment tends to lead to under-estimation of the true prevalence and including only patients with complete enrollment over-estimation.

While this analysis was able to simulate deaths and account for them in the prevalence estimation, the likelihood of death was independent of their underlying diagnoses, and future work should look to build upon this by accounting for the correlation between diagnosis and likelihood of death.

HCM Case Study

Obstructive hypertrophic cardiomyopathy (HCM) is a disease in which the muscles of the heart become abnormally thick or enlarged ultimately impacting the heart's ability to function properly.

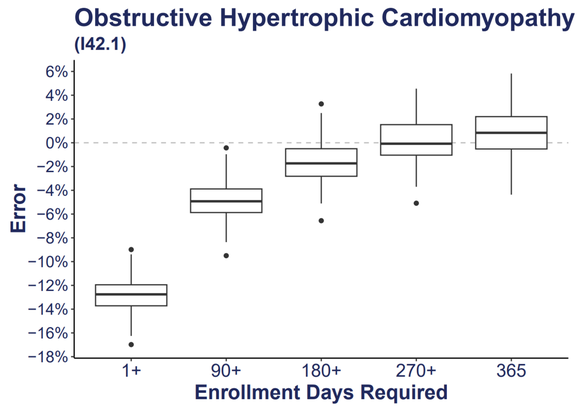
The current treatment paradigm for HCM includes the use of beta and calcium channel blockers, but significant unmet need still exists in the treatment of HCM.

Aficamten is an investigational therapy currently in development for HCM, and as part of our due diligence process noted previously[1,2], we were interested in estimating the prevalence of HCM in the US.

Using our simulation study framework, we evaluated the prevalence estimation of HCM (I42.1), and the results across the 200 simulations can be found in **Figure 3**. Similarly as seen in our larger study, using a criteria of 270+ days of enrollment for inclusion in the prevalence estimation resulted in relatively unbiased estimates of the true prevalence in the analysis population. Additionally, variability in estimation remained relatively similar across the different coverage minimums (± 5%).

Using the entire data cohort from 2022 as well as 270+ days of pharmacy and medical enrollment as our minimum threshold for inclusion, we estimated that ~104,000 persons had a claim with an HCM diagnosis in 2022 representing a prevalence of ~31.2 per 100,000 persons, results which seem reasonable compared to those reported elsewhere[3,4].

Figure 3



References

1. OP10,2022-11, ISPOR Europe 2022, Vienna, Austria
2. Royalty Pharma. (2022, January 7). Royalty Pharma and Cytokinetics Announce Funding Agreements Totaling up to $450 Million [Press release]. https://www.royaltypharma.com/news/royalty-pharma-and-cytokinetics-announce-funding-agreements-totaling-up-to-450-million/
3. Butzner, Michael, et al. "Stable rates of obstructive hypertrophic cardiomyopathy in a contemporary era." Frontiers in Cardiovascular Medicine 8 (2022): 765876.
4. Maron, Martin S., et al. "Occurrence of clinically diagnosed hypertrophic cardiomyopathy in the United States." The American journal of cardiology 117.10 (2016): 1651-1654.