Comparability Assessment of Cohorts with and without Laboratory Values

Rupa Makadia,MS[1,2], Jill Hardin,MS,PhD[1,2], Laura Hester, PhD[1,2], Chris Knoll, BCS[1,2], Ajit Londhe, MPH[1,2], Joel Swerdel, MPH[1,2]

[1]Janssen Research & Development, LLC, Titusville, NJ [2]OHDSI collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY

BACKGROUND

- Laboratory values provide valuable information for characterizing patients with a new condition diagnosis or drug use. However, studies using US administrative claims databases have found that organization- and patient-level factors can influence whether a laboratory test is ordered, received, and recorded in claims data.
- Organizational-level factors influencing availability of lab test values include linkage of claims to specific national lab test providers, which excludes lab values obtained from other lab providers.(1)
- Patients may be more likely to receive lab tests if they need diagnostic or monitoring work for disease risk factors, drug use, or other comorbidities(2-4) or if they exhibit health-seeking behaviors.(5) As a result, meaningful lab values may only be available for certain cohort subpopulations and may not accurately reflect true lab values of all patients in the clinical cohort.
- This study provides a method for comparing and visualizing covariates of patients with and without lab tests ordered within claims-based clinical cohorts as well patients with and without lab test values within clinical cohorts receiving lab tests.

METHODS

- This study used deidentified claims from the Optum Clinformatics® Extended Data Mart, Socio-Economic Status version converted to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), version 5.0.1. Cohorts were developed using the OHDSI Atlas tool.
- Eligible subjects had to: be 18 years and older, have an initial diagnosis and at least one prior confirmatory diagnosis 365 to 180 days before index, and continuous observation of at least 365 days before the index event
- We identified four clinical cohorts:

Clinical cohort (index date)	Lab Measurement	Lab Measurement Evaluation Period
Hepatitis B (first diagnosis)	Alanine aminotransferase (ALT)	365 days before/after index date
Crohn's disease (first diagnosis)	c reactive protein (CRP)	366 days before/after index date
Hyperlipidemia on statins (any use)	Low-density lipoprotein (LDL-c)	180 days before index date
Age >50 years (date met)	Total cholesterol	730 days after index date

- For each cohort we identified: subjects with (target) and without (comparison) lab measurement and subjects with a lab measurement and value (target) and without (comparison) value.
- The unit of analyses are covariates within each domain which are represented as concepts.
- Cohort comparisons were made by calculating the absolute standardized difference in means (aSMD) for all covariates in units of the pooled standard deviation. The expression for calculating the pooled standard deviation and is suitable for dichotomous variables and is not influenced by large sample sizes (6).

Table 1. Condition cohort descriptions and statistics

Measurement	Target Cohort (Absence) (n subjects)	Comparator Cohort (Presence) (n subjects)	Total n covariate compare (all domains)	Condition (n)	Drug (n)	Measure-ment (n)	Procedure (n)	(N) covariates w/ abs std difference > 0.1 (All domains)	Covariate differences by domain w/ abs std difference > 0.1 (N, %)**
With or without measurement									
ALT measure	656	2,076	18,326	10,523	2,622	1,916	3,023	1,576	
c reactive protein	4,813	3,846	29,545	17,430	3,297	3,861	4,688	1,279	
LDL-c measure	**217,867**	**80,752**	**52,085**	**30,440**	**4,034**	**8,451**	**8,864**	**580**	
total cholesterol	**67,565**	**192,689**	**48,520**	**28,982**	**3,981**	**7,612**	**7,663**	**1,594**	
With or without measurement value									
ALT measure	**391**	**1,685**	**14,554**	**7,407**	**2,383**	**2,433**	**2,139**	**2,657**	
c reactive protein	**3,082**	**764**	**20,039**	**11,167**	**2,855**	**2,755**	**3,009**	**1,805**	
LDL-c measure	5,320	75,432	31,046	18,056	3,263	4,402	5,035	590	
total cholesterol	14,368	178,321	34,977	19,600	3,499	6,699	4,901	1,948	

**Domains with N less than 10 were not displayed (Gender, Race, Ethnicity, Metadata)

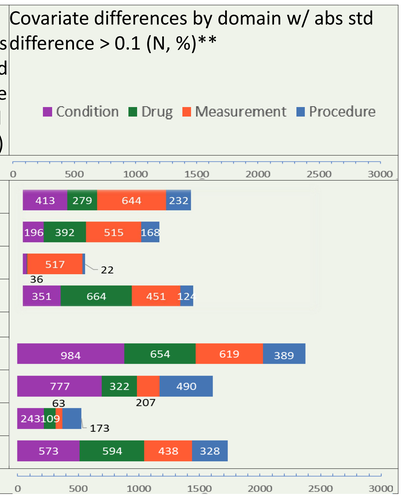
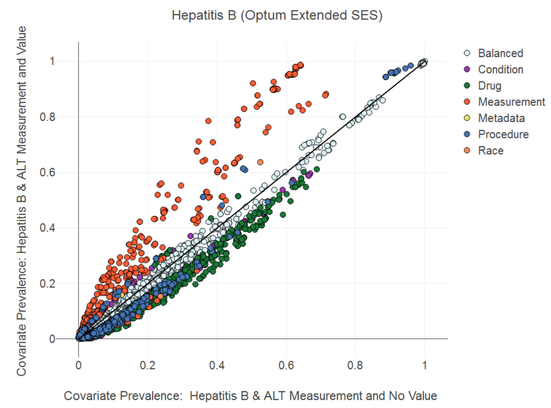
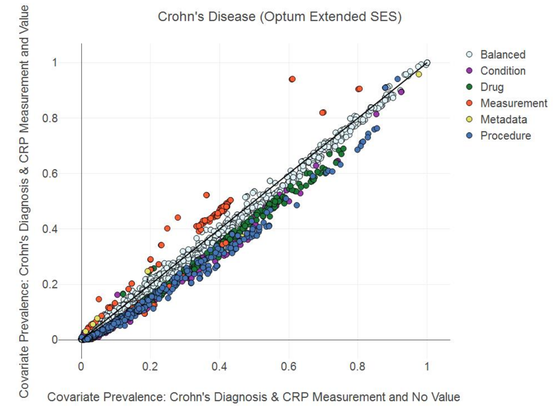
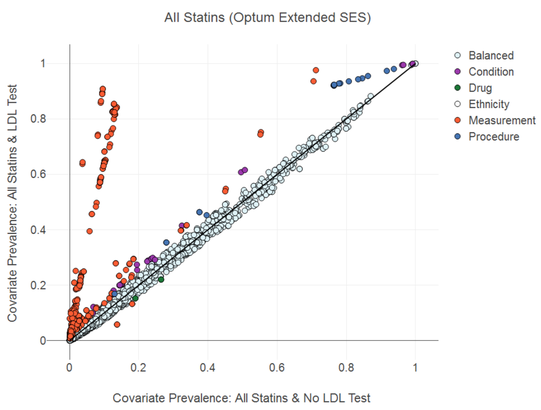
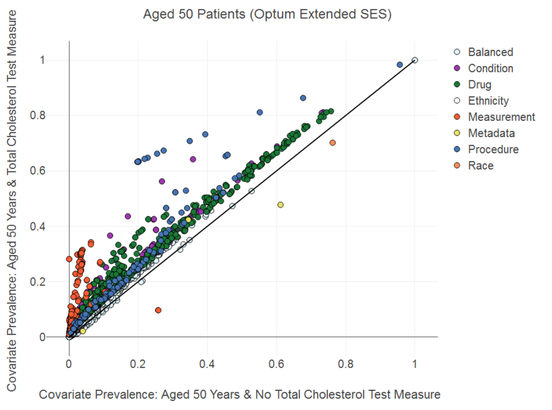


Figure 1. Plots of covariate prevalence for 4 cohorts









Cohort covariate prevalence plots with absolute standardized difference of mean (aSMD) comparisons. The Y-Axis represents the prevalence of covariates in the target cohort and the X-Axis represents the prevalence of covariates in the comparator cohort. The covariates are colored based on vocabulary domain. The 45 degree line represents full covariate balance; that is, an aSMD of 0. Plot points that are light blue have an aSMD < 0.1. Covariates further away from the 45 degree line have an aSMD > 0.1, and thus should be reviewed before utilizing the target cohort as a representative subset of the broader diseased population.

METHODS cont

- Plots were generated comparing cohorts with and without laboratory measurements and values prior to the index date.

RESULTS

- The comparisons of cohorts with a measurement and value vs. cohorts with a measurement and no value have more unbalanced covariates than comparisons between cohorts with laboratory measurements vs. cohorts without laboratory measurements.
- Cohorts of subjects with hepatitis B and Crohn's disease had a higher proportion of covariates with (aSMD) in means > 0.1 compared to hyperlipidemia and patients aged 50 cohorts.
- There are differences among the number and domain of unbalanced covariates between the cohorts with and without laboratory measurements and the cohorts with laboratory measurements with and without values across each outcome.

LIMITATIONS

- Lab results are only available within certain datasets, therefore studies requiring lab data may not be generalizable to broader populations.
- Claims data only capture laboratory tests that are reimbursed and therefore some measurements (e.g. body weight) are more difficult to assess.

CONCLUSIONS

- We developed a systematic framework to assess if use of laboratory measurement data is appropriate to represent subjects without measurements.
- When using measurements to define cohorts, we suggest thoroughly examining the underlying cohorts for comparability.

NEXT STEPS

- This study serves as the basis for developing a set of criteria to illustrate similarity between cohorts of subjects with and without measurements. Further research will include metrics and tests to evaluate the use of cohorts with and without measurements in analyses.

REFERENCES

1. McCullough E, Sullivan C, Banning P, Goldfield N, Hughes J: Challenges and benefits of adding laboratory data to a mortality risk adjustment method. Qual Manag Health Care 2011, 20:253–262.
2. Maciejewski ML, Mi X, Curtis LH, Ng J, Haffer SC, Hammill BG. Frequency of disparities in laboratory testing after statin initiation in subjects ≥65 Years. Am J Cardiol. 2016;118(3):376-82.
3. Maciejewski ML, Mi X, Curtis LH, Ng J, Haffer SC, Hammill BG. Few disparities in baseline laboratory testing after the diuretic or digoxin initiation by Medicare Fee-For-Service beneficiaries. Circ Cardiovasc Qual Outcomes. 2016;9(6):714-22.
4. Schneeweiss S, Rassen JA, Glynn RJ, Myers J, Daniel GW, Singer J, et al. Supplementing claims data with outpatient laboratory test results to improve confounding adjustment in effectiveness studies of lipid-lowering treatments. BMC Med Res Methodol. 2012;12:180-95.
5. Dormuth CR, Patrick AR, Shrank WH, Wright JM, Glynn RJ, Sutherland J, et al. Statin adherence and risk of accidents: a cautionary tale. Circulation. 2009;119(15):2051-7.
6. Austin, P.C., An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res, 2011. 46(3): p. 399-424.

CONFLICT OF INTEREST STATEMENT

The authors are full time employees of Janssen Research and Development, a unit of Johnson and Johnson. The work on this study was part of their employment. They also hold pension rights from the company and own stock and stock options.