Drug Saf (2015) 38:683–686
DOI 10.1007/s40264-015-0313-9

EDITORIAL

Common Models, Different Approaches

Joshua J. Gagne[1]

Published online: 19 June 2015

In recent years, a number of initiatives have established database networks for studying drug safety, including the Mini-Sentinel [1] and Observational Medical Outcomes Partnership (OMOP) [2] programs in the US, the Canadian Network for Observational Drug Effect Studies (CNODES) [3], the Asian Pharmacoepidemiology Network (AsPEN) [4], and the Exploring and Understanding Adverse Drug Reactions (EU-ADR) in Europe [5]. These networks, each comprising data for up to hundreds of millions of individuals, facilitate analyses on unprecedented numbers of patients, which can be particularly useful for evaluating very rare adverse outcomes, investigating heterogeneity across patient subgroups, or assessing outcomes shortly after drug launch, when the number of exposed individuals in any one database may be limited.

Some, but not all, initiatives have adopted common data models (CDMs) to standardize the data structure across the often diverse databases. In particular, the US-based programs, Mini-Sentinel and OMOP, developed separate CDMs and have created tools compatible with the respective CDMs to quickly perform standardized analyses across the database networks [6, 7]. The US Food and Drug Administration (FDA) is now using results from analyses conducted in the Mini-Sentinel CDM to inform regulatory decision making, and data transformed into the OMOP CDM are available from the Reagan-Udall Foundation for the FDA's Innovation in Medical Evidence Development and Surveillance program, which aims to facilitate methods research for medical product safety monitoring, among other objectives. Given the potential regulatory and public health importance of results arising from these programs, it is critical to understand the impact of a CDM on the ability to conduct robust medical product safety surveillance.

The ambitious study by Xu and colleagues in this issue of *Drug Safety* is an important step in this direction [8]. Using Humana's claims database, which they transformed into both the Mini-Sentinel and the OMOP CDMs, the authors conducted what they call an 'ecosystem' comparison by evaluating the results of analyses in the two CDMs using tools that were developed for use in these environments, holding the underlying data constant. Using six drug–outcome pairs for which positive associations are expected, the authors compared what they call the 'high-dimensional propensity score- (hdPS-) based analysis procedures' developed for each CDM and the 'self-controlled case series (SCCS) analysis procedure' for each CDM.

The authors also compared the CDMs on a conceptual level, elucidating a number of important differences between the two. A particularly salient difference is evident in the sometimes large differences in numbers of patients each approach identifies as being exposed to particular medical products. For example, the application of the 'hdPS-based analysis procedure' in the OMOP environment identified 356,078 new users of ketorolac, whereas the application in the Mini-Sentinel environment identified 30,322 new users (<9 % of OMOP total). As the authors explain, they used only national drug codes (NDCs) to identify ketorolac exposure in the Mini-Sentinel analysis, but the concept-based identification process used in the OMOP CDM also captured drug exposure using procedure

✉ Joshua J. Gagne
jgagne1@partners.org

[1] Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 1620 Tremont Street, Suite 3030, Boston, MA 02120, USA

codes (e.g., J-codes). It is important to note that, while the authors used only NDCs to capture drug exposure, the Mini-Sentinel CDM permits drug exposure definitions using any combination of NDC and procedure codes, if that is of relevance for a particular assessment.

It is important to note also that there are differences between the tools that the authors implemented and the tools made available by Mini-Sentinel. The self-controlled analysis program developed by Mini-Sentinel is a self-controlled risk interval method [9], which is a variant of the SCCS, but, because the tool was not available at the time of the analysis, the authors wrote their own code to implement the SCCS in the Mini-Sentinel CDM. In implementing the 'hdPS-based analysis procedure,' the authors included only empirically identified covariates, whereas use of this approach typically involves enriching a set of pre-defined, investigator-specified covariates with those identified by the hdPS algorithm.

Despite differences in the CDMs and in configurations of the tools, Xu and colleagues found that, when properly applied, these methods yielded estimates of association consistent with expectation in both CDMs. When fixed risk windows were used, both methods produced estimates in the expected direction across all six drug–outcome pairs in both Mini-Sentinel and OMOP CDMs. While not all of the associations reached statistical significance, it is important to emphasize that the analyses were conducted in a single database, whereas the purpose of CDMs is to facilitate analyses across multiple databases, which would provide more statistical power. Indeed, only one exposed event contributed to the analysis using the 'hdPS-based analysis procedure' in the OMOP CDM, demonstrating the lack of power to detect statistically significant associations for some pairs in this single database study.

Another way to examine the potential impact of the CDM on variability in results is to compare differences in results between the Mini-Sentinel and OMOP ecosystems to differences arising from other sources. For example, differences in results due to use of the cohort-type 'hdPS-based analysis procedure' vs the self-controlled 'SCCS analysis procedure' and from adjusting vs not adjusting for confounding in the cohort-type analyses can be evaluated holding constant both the underlying data and the CDM. The comparison of the effect of confounding can further be made holding constant other design features.

To illustrate the relative impact of the ecosystem in this context, Table 1 compares differences in estimates observed in the ecosystem comparisons to other factors. The numbers in the table are absolute differences in the natural logs of the estimates reported by Xu and colleagues comparing each strategy. Larger values indicate that the compared strategies produced more divergent estimates; identical estimates would result in values of zero. In four out of six of the drug–outcome pairs, the largest discrepancies were observed comparing the variable vs fixed follow-up implementations of the OMOP 'hdPS-based analysis procedure'. OMOP's implementation of the 'hdPS-based analysis procedure' with a variable risk window allows patients to contribute variable amounts of person-time to the analysis for as long as they are exposed to the drug of interest, plus an extension window. However, the subsequent analysis uses a logistic regression model to estimate an odds ratio, which does not account for the variable length of follow-up. Obviously, patients with longer follow-up have greater opportunity to experience the outcome of interest independent of any action of the drug. This implementation yielded odds ratios that were less than 1.0, and therefore in the opposite direction of expectation, for four out of six drug–outcome pairs. Indeed, the average follow-up time for patients in the comparator drug group was substantially larger than the average follow-up time for patients in the drug of interest group [e.g., mean follow-up for amoxicillin (drug of interest), 12 days vs mean follow-up for comparator, 471 days], explaining the erroneous findings.

In contrast, adjusting vs not adjusting for confounding yielded the largest difference in log estimates for one drug–outcome pair (i.e., valproic acid and acute liver injury). Similarly, differences between the ecosystems yielded the largest difference in estimates for only one drug–outcome pair (i.e., carbamazepine and acute liver injury), when comparing the Mini-Sentinel and OMOP implementations of the 'hdPS-based analysis procedure' using both confounding adjustment and the fixed risk window. However, as mentioned above, the estimate from the application of the 'hdPS-based analysis procedure' in the OMOP CDM was based on a single exposed event and the estimate in the Mini-Sentinel CDM was based only on two exposed events, so large differences in results are plausibly attributable to chance.

Placing the impact of differences in ecosystems in the context of differences in results arising from other decisions required of investigators reveals that use of appropriate analysis techniques is at least as important as the choice of well designed CDM. While this finding supports the use of CDMs in multi-database evaluations, it also highlights that CDMs and standardized analytic tools developed to interface with them must enable investigators to implement the most appropriate design and analysis plans for given drug–outcome pairs. To the extent that

Table 1 Relative impact of ecosystem, study design, confounding adjustment, and analysis strategy on differences in results of six drug safety assessments[a]

	OMOP vs MS ecosystem		SCCS vs hdPS		Confounding adjustment vs no adjustment		Appropriate vs inappropriate analysis[b]
Drug–outcome pair	OMOP hdPS, adjusted, fixed vs MS, hdPS, adjusted, fixed	OMOP SCCS, fixed vs MS SCCS, fixed	OMOP SCCS, fixed vs OMOP hdPS, adjusted, fixed	MS SCCS, fixed vs MS hdPS, adjusted, fixed	OMOP hdPS, adjusted, fixed vs OMOP hdPS, unadjusted, fixed	MS hdPS, adjusted, fixed vs MS hdPS, unadjusted, fixed	OMOP, hdPS, adjusted fixed vs OMOP, hdPS, adjusted variable
Indomethacin and acute myocardial infarction	0.09	0.00	0.17	0.08	0.10	0.07	**0.71**
Ketorolac and gastrointestinal bleeding	0.08	0.03	0.01	0.04	0.31	0.48	**0.98**
Benzodiazepines and hip fracture	0.15	0.15	0.03	0.32	0.03	0.22	**0.66**
Valproic acid and acute liver injury	0.28	0.09	0.07	0.12	0.63	**0.97**	0.09
Carbamazepine and acute liver injury	**1.65**	0.44	0.23	1.44	0.04	0.58	0.14
Anaphylactic shock	0.63	0.14	0.08	0.41	0.11	0.02	**1.87**

Bold values indicate the comparison with the largest difference in natural logs of the estimates for each of the six drug-outcome pairs

Adjusted and unadjusted indicate whether analyses adjusted for potential confounders, respectively; fixed and variable indicate whether analyses used fixed follow-up window or allowed patients to contribute variable follow-up time, respectively; numbers are absolute differences in the natural logs of the estimates reported by Xu and colleagues between the compared strategies

hdPS High-dimensional propensity score-based analysis procedure, *MS* Mini-Sentinel, *OMOP* Observational Medical Outcomes Partnership, *SCCS* self-controlled case series analysis

[a] Numbers are absolute differences in the natural logs of the estimates reported by Xu and colleagues [8] between the compared strategies

[b] Inappropriate analysis indicates the use of variable risk windows in the OMOP implementation of the 'high-dimensional propensity score-based analysis procedure' and estimation of an odds ratio without accounting for the variable lengths of follow-up; note that the 'high-dimensional propensity score-based analysis procedure' developed by Mini-Sentinel does not permit this inappropriate analysis

CDMs facilitate scaling of the most rigorous design and analysis plans, bigger will be better. However, scaling of inappropriate design and analysis methods will lead to more results that are precisely wrong.

Compliance with Ethical Standards

Funding No funding was used in the preparation of this editorial.

Conflicts of interest Joshua J. Gagne was previously Principal Investigator of an investigator-initiated grant from Novartis Pharmaceuticals Corporation to the Brigham and Women's Hospital for work unrelated to this editorial. He is the Principal Investigator of a grant from the Reagan-Udall Foundation for the FDA's Innovation in Medical Evidence Development and Surveillance program to the Brigham and Women's Hospital, an investigator on subcontracts from the Harvard Pilgrim Healthcare Institute to the Brigham and Women's Hospital for work funded by the FDA Sentinel program, the Principal Investigator on a cooperative agreement from the US FDA to the Brigham and Women's Hospital, a consultant to Aetion, Inc., and is co-lead of the Sentinel Methods Core.

References

1. Platt R, Carnahan RM, Brown JS, et al. The US Food and Drug Administration's Mini-Sentinel program: status and direction. Pharmacoepidemiol Drug Saf. 2012;21(Suppl 1):1–8.
2. Stang PE, Ryan PB, Racoosin JA, et al. Advancing the science for active surveillance: rationale and design of the Observational Medical Outcomes Partnership. Ann Intern Med. 2010;153:600–6.
3. Suissa S, Henry D, Caetano P, et al. CNODES: the Canadian Network for Observational Drug Effect Studies. Open Med. 2012;6:134–40.
4. Andersen M, Bergman U, Choi NK, et al. The Asian Pharmacoepidemiology Network (AsPEN): promoting multi-national collaboration for pharmacoepidemiologic research in Asia. Pharmacoepidemiol Drug Saf. 2013;22:700–4.
5. Coloma PM, Schuemie MJ, Trifiro G, et al. Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. Pharmacoepidemiol Drug Saf. 2011;20:1–11.
6. Curtis LH, Weiner MG, Boudreau DM, et al. Design considerations, architecture, and use of the Mini-Sentinel distributed data system. Pharmacoepidemiol Drug Saf. 2012;21(Suppl 1):23–31.
7. Overhage JM, Ryan PB, Reich CG, Hartema AG, Stang PE. Validation of a common data model for active safety surveillance research. J Am Med Inform Assoc. 2012;19:54–60.
8. Xu Y, Zhou X, Suehs BT, et al. A comparative assessment of Observational Medical Outcomes Partnership and Mini-Sentinel common data models and analytics: implications for active drug safety surveillance. Drug Saf 2015 (this issue). doi:10.1007/s40264-015-0297-5.
9. Coronel-Moreno C, Vilk Y, Silva I, Yih K, Kulldorff M. Mini-Sentinel Prospective Routine Observational Monitoring Program Tool: continuous and group sequential analysis with self-control design. Technical Users' Guide version: 1.0. Available at: http://www.mini-sentinel.org/work_products/Statistical_Methods/Mini-Sentinel_PROMPT_Self-Control-Design-Tool_Technical-Users-Guide.pdf. Accessed 8 June 2015.