
Imperfect Administrative Health Data: Can we do better?

Organizer and Chair: Lisa Lix (University of Saskatchewan)

YULEI HE, Harvard University

Combining Information on Health Services Variables from Multiple Sources

Measures of certain services variables (e.g., hospice use) are important quality indicators of cancer care. The Cancer Care Outcomes Research and Surveillance (CanCORS) study collected these variables from several sources including patient survey, medical records abstraction, cancer registry databases, and Medicare claims data. The sample coverages of these sources are different, and the reporting might be subject to error. We first estimate the patterns of misreporting assuming one source (e.g., claims) is a gold standard. The misreported outcome in other sources can therefore be corrected/imputed using a statistical model. Valid analysis can then be based on multiply imputed data.

MARTIN LADOUCEUR, McGill University

Investigating Robustness of Prevalence Estimates from Misclassified Administrative Data Using Bayesian Latent Class Models

Information collected in medical administrative databases is typically collected for reasons other than research, and many such databases have been shown to contain substantial proportions of misclassification errors. For example, errors in ascertaining which individuals have a given disease may bias prevalence estimates. We attempt to estimate the prevalence of osteoarthritis (OA) among elderly Quebecers using a government administrative database. We compare a naive estimate relying solely on the physician diagnoses of OA listed in the database to estimates from several different Bayesian latent class models which adjust for misclassified physician diagnostic codes via use of other available diagnostic clues.

KHOKAN SIKDAR, Newfoundland and Labrador Centre for Health Information

Challenges and Opportunities for Using Administrative Health Data to Improve Chronic Disease Case Ascertainment

Administrative data plays an important role in population health research and disease surveillance, especially in case ascertainment. These data are often readily available, cost-effective and typically encompass large populations. However, gaps in clinical information including physician billing practices and coding can compromise the ability to ascertain valid chronic disease cases; this can affect accuracy and completeness of data. This presentation uses various methods including ratio estimation and capture-recapture to adjust for data incompleteness in estimating diabetes incidence and prevalence in Newfoundland and Labrador. This research will show how these methods can aid in improving case ascertainment from incomplete data.