International Journal of Population Data Science

Lessons learned: Linking patient-reported outcomes data with administrative databases

Davis, L[1], Mahar, A[2], Bubis, L[3], Li, Q[4], Zhao, H[4], Moody, L[5], Sutradhar, Institute for Clinical Evaluative Sciences[4], Barbera, L[6], and Coburn, N[6]

[1]Sunnybrook Research Institute
[2]Manitoba Centre for Health Policy
[3]Sunnybrook Health Sciences
[4]Institute of Clinical Evaluative Sciences
[5]Cancer Care Ontario
[6]Odette Cancer Centre, Sunnybrook Health Sciences Centre

Introduction

Since 2007, Cancer Care Ontario (CCO) has systematically collected patient-reported outcomes (PROs) in the form of symptom data, for cancer outpatients visiting regional cancer centres or affiliate institutions. Data are used in real-time to facilitate conversation between clinicians and patients and have recently been combined with provincial administrative databases.

Objectives and Approach

CCO collects PROs using the Edmonton Symptom Assessment System (ESAS), which scores 9 symptoms on a scale of 0 (no symptoms) to 10 (worst symptom severity). Data were imported from CCO in 2015 and linked to a cancer cohort at ICES. We investigated differences between patients who completed ≥ 1 ESAS record and patients who did not, as well as the number of records, timing of data collection and missingness. We describe our experience linking and using the PRO data to administrative data, including presenting trajectories of symptoms over time and combining scores into composite indices.

Results

120,745 cancer patients had 729,861 symptom records between 2007 and 2014. Not all patients with a cancer diagnosis had ≥ 1 ESAS record and this varied by patient, disease and system level factors. Because implementation occurred from a clinical perspective, data collection was irregular within and across patients and depended on treatment and other factors; the number of records per patient varied, as well the number of contributing patients in each time period following diagnosis. Attempts were made to create meaningful composite indices by combining all symptom scores as well as combining multiple high scores for each individual symptom. As a result, selecting the best statistical analysis to use these PRO data as an exposure or outcome is still uncertain.

Conclusion/Implications

PRO data linked to provincial, administrative data holdings represent a new frontier for population-based cancer research, both in their challenging structure as well as their implications for clinical practice and health system. These lessons learned will hopefully support other researchers rigorous use of these data in the future.