Validation of canadian health administrative data algorithms for estimating trends in the incidence and prevalence of osteoarthritis

Jessica Widdifield a, b, c, *, R. Liisa Jaakkimainen a, b, c, d, Jodi M. Gatley b, Gillian A. Hawker b, e, f, Lisa M. Lix g, Sasha Bernatsky h, Bheeshma Ravi j, David Wasserstein g, Bing Yu b, Karen Tu c, i, j

a Sunnybrook Research Institute, Toronto, Canada
b ICES, Toronto, Canada
c Institute of Health Policy, Management & Evaluation, University of Toronto, Toronto, Canada
d Department of Family and Community Medicine, University of Toronto, Toronto, Canada
e Department of Medicine, University of Toronto, Toronto, Canada
f Women’s College Hospital, Toronto, Canada
g Department of Community Health Sciences, University of Manitoba, Winnipeg, Canada
h Research Institute of McGill University Health Centre, Montreal, Canada
i Toronto Western Hospital Family Health Team-University Health Network, Toronto, Canada
j North York General Hospital, Toronto, Canada

A R T I C L E I N F O

Keywords:
Health administrative data
Epidemiology
Osteoarthritis
Electronic medical records
Incidence and prevalence
Validation

S U M M A R Y

Objective: To estimate the 1) accuracy of algorithms for identifying osteoarthritis (OA) using health administrative data; and 2) population-level OA prevalence and incidence over time in Ontario, Canada.

Method: We performed a retrospective chart abstraction study to identify OA patients in a random sample of 7500 primary care patients from electronic medical records. The validation sample was linked with several administrative data sources. Accuracy of administrative data algorithms for identifying OA was tested against two reference standard definitions by estimating the sensitivity, specificity and predictive values. The validated algorithms were then applied to the Ontario population to estimate and compare population-level prevalence and incidence from 2000 to 2017.

Results: OA prevalence within the validation sample ranged from 10% to 23% across the two reference standards. Algorithms varied in accuracy depending on the reference standard, with the sensitivity highest (77%) for patients with OA documented in medical problem lists. Using the top performing administrative data algorithms, the crude population-level OA prevalence ranged from 11% to 25% and standardized prevalence ranged from 9 to 22% in 2017. Over time, prevalence increased whereas incidence remained stable (~1% annually).

Conclusion: Health administrative data have limited sensitivity in adequately identifying all OA patients and appear to be more sensitive at detecting OA patients for whom their physician formally documented their diagnosis in medical problem lists than individuals who have their diagnosis documented outside of problem lists. Irrespective of the algorithm used, OA prevalence has increased over the past decade while annual incidence has been stable.

1. Introduction

Osteoarthritis (OA) is the most common form of arthritis (1), but incidence and prevalence estimates vary across studies. Various case definitions of OA (symptomatic, radiographic, self-reported, or doctor-diagnosed) [2–6] have been used across different clinical studies. Similarly, epidemiological studies using health administrative and electronic medical record (EMR) data have used various OA algorithms [7–15].

https://doi.org/10.1016/j.ocarto.2020.100115
Received 15 July 2020; Accepted 26 October 2020
2665-9131/© 2020 The Authors. Published by Elsevier Ltd on behalf of Osteoarthritis Research Society International (OARSI). This is an open access article under the

* Corresponding author. Holland Bone & Joint Research Program, Sunnybrook Research Institute, MG 352 - 2075 Bayview Avenue, Toronto, ON, M4N 3M5, Canada.
E-mail addresses: jessica.widdifield@utoronto.ca (J. Widdifield), liisa.jaakkimainen@ices.on.ca (R.L. Jaakkimainen), Jodi.Gatley@ices.on.ca (J.M. Gatley), g.hawker@utoronto.ca (G.A. Hawker), Lisa.Lix@umanitoba.ca (J.M. Lix), sasha.bernatsky@mcgill.ca (S. Bernatsky), bheeshma.ravi@sunnybrook.ca (B. Ravi), david.wasserstein@utoronto.ca (D. Wasserstein), Bing.Yu@ices.on.ca (B. Yu), k.tu@utoronto.ca (K. Tu).
which may contribute to varying disease estimates across studies.

Health administrative data are generated at every patient encounter within a health care system, whether through a physician visit, a diagnostic or therapeutic procedure, hospital admission, or receiving/filling prescription medication [16]. There is uncertainty surrounding the optimal approach to identifying OA in administrative data as algorithms have been shown to have poor sensitivity and low positive predictive values (PPV) [17–20]. Potential reasons for administrative data having suboptimal performance may be a consequence of OA care not being prioritized during physician encounters to warrant assignment of a diagnosis or a diagnosis code (6), or perhaps imperfect reference standards being used during the validation of algorithms.

The key aspect that differentiates health administrative data from EMRs is that the latter contains more detailed clinical data comprised of a mix of coded (structured) and free text data. As the coded data in EMRs (such as diagnoses in problem lists) is much more readily accessible to researchers, clinical diagnoses embedded in free text data are often missed. Therefore, ascertaining patient populations using coded data like problem lists may be imperfect, as data completeness varies substantially across providers [21–23]. Thus, when the case identification process is confined to coded EMR data alone, classification errors in disease status can be significant [24–26] and can considerably bias study findings [27].

In an attempt to resolve these issues, we conducted a detailed review of EMRs to identify OA patients in order to test the accuracy of administrative data OA algorithms against two different definitions for classifying OA patients (based on the location of the diagnosis within the EMR). We then applied the top-performing algorithms to the entire Ontario population to compare the annual estimates of OA incidence and prevalence from 2000 to 2017.

2. Method

Setting and design. Ontario residents are covered by single-payer, public health insurance that covers hospital care and physicians’ services. We first conducted a retrospective chart abstraction study on a random sample of patients from primary care EMRs to identify patients with a physician-documented OA diagnosis. This validation sample was used to construct two different reference standard definitions, which varied based on the location of where an OA diagnosis was documented in the EMR. We tested multiple administrative data algorithms across both reference standards. Subsequently, we conducted a population-based study using the top performing validated algorithms to assess annual OA incidence and prevalence.

Validation cohort selection. The validation sample was derived from the Electronic Medical Record Primary Care Database. The patient profile of this primary care population closely approximates that of the general population in terms of demographics and disease burden [28], making these data a preferred source for validation studies [29–32].

The data contains a full-chart extract of all clinical data from a patient’s EMR, including cumulative patient profiles with medical problem lists, progress notes, laboratory and diagnostic test results, prescriptions, specialist consultation letters and hospital discharge summaries. The comprehensiveness of the data has been evaluated and all data go through data quality assessments after collection and before research use [28,33]. These data quality assessments include verifying individual patient-level and practice-level data entry fields meet data quality thresholds. At the time of study, the database was comprised of 400 primary care physicians distributed across Ontario who use the PS (Practice Solutions) Suite EMR. The database consisted of 443,253 unique patients, of which 350,863 met potential eligibility criteria for this study.

Patient inclusion criteria required a valid health insurance number and date of birth; one or more physician visits in the past year; and enrolled in the physician’s practice (to allow for sufficient clinical information to verify disease status). Physician inclusion criteria required at least two years of data for their individual practice population to ensure the EMR was adequately populated. Among 73,014 eligible patients from 83 primary care physicians meeting both patient and physician eligibility criteria, we randomly sampled 7500 patients aged 20 years and older.

Data abstraction. For each of the 7500 patients, the complete EMR was reviewed by one of five trained chart abstractors to review every entry in the EMR and identify documented diagnosis of OA (at any joint) using a standardized data abstraction tool. Trained abstractors coded each patient encounter (including progress notes, consultation letters and imaging reports) as to whether or not a diagnosis of OA was documented. A diagnosis was not inferred if none was explicitly stated. Abstractors also coded whether the OA diagnosis was documented in medical problem lists, and/or in any location within the EMR, including the free text portions of the EMR (i.e. progress notes, consultation letters).

To assess the intra- and inter-rater reliability of abstractors, an initial 10% sample was abstracted a second time by the same abstractor, and once by a different abstractor. Kappa scores for inter- and intra-rater reliability all exceeded 0.85 indicating good agreement for all five chart abstractors.

Reference Standard Definitions. We a priori defined two alternative reference standard definitions of OA within the validation sample. Reference standard 1 classified all individuals with OA if the diagnosis was present in any location of their EMR. Reference standard 2 classified individuals with OA if the diagnosis was on their medical problem list only. Those who did not meet the stated OA case criteria were defined as non-cases. Thus, for reference standard 2, patients with an OA diagnoses only documented outside their medical problem list (i.e. in unstructured free text data within clinical notes) were classified as a non-case.

Health administrative data sources. Several different administrative data sources were obtained and records were linked between databases using unique, encoded identifiers [16]. We used the Ontario Health Insurance Plan (OHIP) Claims Database to identify physician billing diagnosis codes for OA [34] using International Classification of Diseases (ICD) code 715. Within OHIP, physicians provide only one diagnosis code, representing the main ‘reason for the visit’, with each billing claim. Physician specialty (associated with OHIP billing claims for OA) was identified by linking with the ICES Physician Database. From the Canadian Institute for Health Information Discharge Abstract Database, we identified OA diagnosis codes (ICD9 715, ICD10 M15-M19) within all diagnoses listed in the hospital discharge abstract. The National Ambulatory Care Reporting System identified OA diagnoses recorded during emergency department visits [35]. All hospital data prior to April 1, 2002 have diagnoses coded in ICD-9 [36] and can contain up to 16 recorded diagnoses per hospital encounter. Hospitalizations and emergency department visits after April 1, 2002 are coded using ICD-10-CA [36], and each record contains up to 25 diagnoses per admission. Patient demographic information and vital status were obtained from the Registered Persons Database (RPDB). Annual population denominators were ascertained for all OHIP beneficiaries that were 20 years and older, alive, and who have accessed the healthcare system (at least once in a seven-year period of the reporting year).

Derivation of health administrative data algorithms. Administrative data algorithms are simple rule-based case definitions which are derived from coded data (such as diagnosis codes and pharmacy dispensations) from one or more data sources. Administrative data OA algorithms were developed using combinations of physician billing diagnosis for OA (ICD 715), primary and secondary hospital discharge diagnoses for OA (ICD9 715; ICD10 M15-M19), prescription drug claims (for NSAIDs, COX-2 inhibitors), by varying time windows between diagnosis codes or the period in which diagnosis codes appeared (e.g. within 2 years), and whether the diagnosis codes were rendered by a musculoskeletal specialist (orthopedic surgeon, rheumatologist, or internist). We a priori were interested in identifying administrative data OA algorithms which maximized sensitivity and PPV (to identify all true OA patients and minimize false positives) [37].

Statistical analysis. We descriptively characterized the validation
sample and assessed the prevalence of OA according to the two reference standard definitions. We estimated the accuracy of administrative data algorithms against the two reference standards by computing the sensitivity, specificity, PPV, and negative predictive value (NPV), accompanied by their 95% confidence intervals (CIs) to identify those with the combined optimal sensitivity (to capture all cases) and PPVs (to reduce false positives). We then applied the selected administrative data algorithms to the entire Ontario adult population to provide comparative estimates of incidence and prevalence of OA in the population. For each selected algorithm, annual crude and age- and sex-standardized prevalence and incidence over time were estimated. Annual prevalence were calculated as the total number of patients aged 20 years or older classified as having OA divided by the number of Ontario residents aged 20 years or older. The annual denominators consisted of all Ontario residents for each calendar year, excluding those <20 years of age or those who died in the year prior. Once a case was included, it remained a case in the next year unless they died. The annual incidence was calculated by dividing the number of new OA patients aged 20 or older by the population at risk among Ontario residents aged 20 or older. Incidence was defined as the first occurrence of OA per 100 population. As administrative data are available from 1990 onwards, to reduce the potential of observational period effects (misclassification of incident and prevalent cases after the establishment of administrative data), cases were ascertained over all available years of data but annual estimates are reported from 2000 onwards. Therefore a 10-year washout period was used to report incidence. To adjust for differences in population distribution over time, direct age and sex standardization was done using 2001 Ontario census population estimates.

All analyses were performed at ICES (www.ices.on.ca) using SAS version 9.2 (SAS Institute, Cary, NC, USA). This study was approved by the Research Ethics Board at Sunnybrook Health Sciences Centre.

3. Results

3.1. Description of the validation sample: Table 1

When the reference standard was considered to be a physician-documented OA diagnosis within any part of the EMR (reference standard 1), we identified 1736 cases and 5764 non-cases, corresponding to a 23.1% OA prevalence. With the second reference standard (physician-documented OA diagnosis only in the medical problem list), 738 cases and 6762 non-cases were identified, corresponding to a 9.8% OA prevalence. Patient characteristics by reference standard definition are presented in Table 1.

3.2. Validation results: Table 2

The diagnostic test characteristics for different administrative data algorithms to detect each reference standard are presented in Table 2. The test characteristics of administrative data algorithms identifying reference standards 2 (OA diagnosis documented in EMR problem list) differed significantly (with non-overlapping confidence intervals) from administrative data algorithms identifying reference standard 1, which also included individuals labelled as having OA elsewhere in their EMR. Across all administrative data algorithms tested, sensitivity was significantly and consistently higher for detecting OA cases for reference standards 2 (EMR problem list only), as compared to reference standard 1. Overall, specificity estimates across algorithms were modest to excellent, and tended to be highest within reference standard 1 (likely due to fewer cases misclassified in the reference standard). PPV estimates were highest for reference standard 1 (also likely due to fewer cases misclassified in the reference standard). In all instances, PPVs increased as the number of physician diagnosis codes in administrative data increased. Administrative data algorithms that incorporated additional features, such as being seen by a musculoskeletal specialist or a NSAID dispensation, did not improve accuracy. Further, only 28.3%, and 36.9% of OA patients in reference standards 1 and 2, respectively, had OA billing claims by a musculoskeletal specialist, thus dramatically decreasing the sensitivity of algorithms requiring specialty claims.

Algorithms that required “at least one hospital admission with an OA diagnosis or one (or two) physician OA diagnosis code(s) ever” were associated with maximal sensitivity, while algorithms that required “at least one hospital admission associated with an OA diagnosis or at least two (or three) physician diagnosis codes within five years” had the highest PPVs. These four algorithms were selected to compare estimates of population-level prevalence.

3.3. Estimated population incidence and prevalence of OA: Fig. 1

From 2000 to 2017, the number of Ontario adults aged 20 years of age and older (the population denominator) increased from 8,643,521 to 10,969,729 (Supplementary Table 1). Based on the best algorithms, the crude OA prevalence more than doubled over this time period, from 5.4–11.2% in 2000 to 11.4–25.3% in 2017 (Fig. 1). Age- and sex-standardized prevalence estimates nearly doubled over the study period, with an estimated 9.3–22.0 cases per 100 population in 2017 (Fig. 2, Supplementary Table 2).

The crude incidence ranged from 1.0 to 2.1 cases per 100 population in 2000 to 0.7–1.5 cases per 100 population in 2017 (Fig. 3, Supplementary Table 3). Standardized incidence across definitions estimated about 1 newly identified OA case per 100 persons in each year, Fig. 4. Annual incidence estimates were slightly elevated in years 2000–2007, which may have been influenced by observational period effects rather than a true decline in incidence over time, Supplementary Table 4.

The age-standardized estimates in males and females showed similar increasing trends to the overall estimates, with higher number for females versus males, Supplementary Fig. 1.

In 2017, among the 10,969,729 adult Ontario population, a large variation in the counts of the number of new and total OA cases across different algorithms were noticeable, Fig. 4. Twice as many new annual cases were ascertained from the population by the most sensitive case definition (n = 121,068) versus the algorithm with the highest PPV (but lower sensitivity), which only identified 67,187 new annual cases. The total number of OA cases to be identified over the study period was...
similarly significantly higher for the most sensitive algorithm (n = 2,775,888, 25.3%) than the lower sensitivity algorithm (n = 1,250,161, 11.4%). The most sensitive algorithm achieved a population-level prevalence that most closely approximates the pre-test prevalence of OA ascertained from the primary care random sample.

4. Discussion

We present a comprehensive assessment on the accuracy of algorithms for detecting OA in health administrative data. Our findings report that administrative data have limited sensitivity in adequately identifying all OA patients and appear to be most sensitive at detecting OA patients for whom their physician formally documented their diagnosis in medical problem lists. These findings underscore the need for careful scrutiny of the use of administrative data for OA surveillance as they may substantially underestimate the number of individuals with OA. By contrasting different reference standard definitions used for administrative data validation studies, we observed the effects of alternative reference standards on algorithm performance. These findings illustrate under-documentation of OA diagnoses in medical problem lists, and defining OA cases by medical problem lists alone estimated lower PPVs for administrative data algorithms because the reference standard definition completely misses some true OA cases who have their diagnosis documented outside of problem lists. Our population-based findings estimated an increasing OA prevalence over time irrespective of the algorithm used to identify OA, whereas annual incidence remained steady.

Table 2
Accuracy of administrative data algorithms to identify OA against 2 different reference standard definitions.

| Algorithm | Reference Standard | Sensitivity (95% CI) | Specificity (95% CI) | PPV (95% CI) | NPV (95% CI) | Post-Test Prevalence |
|-----------|-------------------|----------------------|----------------------|--------------|--------------|---------------------|
| ≥ 1 H ever or ≥ 1 P ever | 1 | 68.5 (66.3-70.7) | 84.3 (83.4-85.3) | 56.8 (54.7-59.0) | 89.9 (89.1-90.7) | 27.9% |
| | 2 | 77.2 (74.2-80.3) | 77.5 (76.5-78.5) | 27.2 (25.3-29.2) | 96.5 (96.4-97.4) | |
| ≥ 2 P in 2 years | 1 | 45.3 (42.9-47.6) | 95.0 (94.5-95.6) | 73.3 (70.7-76.0) | 85.2 (84.4-86.1) | 14.3% |
| | 2 | 56.4 (52.8-59.9) | 90.3 (89.6-91.0) | 38.8 (35.9-41.7) | 95.0 (94.5-95.5) | |
| ≥ 2 P in 3 years | 1 | 47.1 (44.8-49.5) | 94.6 (94.0-95.2) | 72.3 (69.9-75.1) | 85.6 (84.7-86.3) | 15.0% |
| | 2 | 58.8 (55.3-62.4) | 89.7 (89.0-90.5) | 38.9 (35.6-41.3) | 95.2 (94.7-95.8) | |
| ≥ 3 P in 3 years | 1 | 33.8 (31.5-36.0) | 97.3 (96.8-97.7) | 78.8 (75.8-81.7) | 83.0 (82.1-83.9) | 9.9% |
| | 2 | 43.5 (39.9-47.1) | 93.7 (93.2-94.3) | 43.1 (39.6-46.7) | 93.8 (93.3-94.4) | |
| ≥ 1H ever | 1 | 13.0 (11.4-14.6) | 99.4 (99.3-99.6) | 87.6 (83.6-91.6) | 79.1 (78.2-80.1) | 3.4% |
| | 2 | 17.6 (14.9-20.4) | 98.1 (97.8-98.4) | 50.4 (44.3-56.5) | 91.6 (91.0-92.2) | |

Top Performing Definitions

| ≥ 1 H ever or ≥ 1 P ever | 1 | 68.6 (66.4-70.8) | 84.2 (83.3-85.2) | 56.7 (54.6-58.8) | 89.9 (89.1-90.7) | 28.0% |
| | 2 | 77.5 (74.5-80.5) | 77.4 (76.4-78.4) | 27.2 (25.3-29.1) | 96.9 (96.5-97.4) | |
| ≥ 1 H ever or ≥ 2 P ever | 1 | 50.9 (48.6-53.3) | 93.4 (92.7-94.0) | 69.9 (67.4-72.4) | 86.3 (85.5-87.2) | 16.9% |
| | 2 | 62.7 (59.2-66.2) | 88.1 (87.4-88.9) | 36.6 (33.9-39.2) | 95.6 (95.1-96.1) | |
| ≥ 1H or (≥2 P in 5 years) | 1 | 48.4 (46.0-50.7) | 94.1 (93.5-94.7) | 71.3 (68.7-73.9) | 85.8 (85.0-86.7) | 15.7% |
| | 2 | 60.0 (56.5-63.6) | 89.1 (88.4-89.9) | 37.6 (34.8-40.4) | 95.3 (94.8-95.9) | |
| ≥ 1H ever or (≥3 P in 5 years) | 1 | 48.4 (46.0-50.7) | 94.1 (93.5-94.7) | 71.3 (68.7-73.9) | 85.8 (85.0-86.7) | 15.7% |
| | 2 | 60.0 (56.5-63.6) | 89.1 (88.4-89.9) | 37.6 (34.8-40.4) | 95.3 (94.8-95.9) | |

Abbreviations: H = hospitalization; P = physician billing claim.

a Reference Standard 1: 1736 cases with a physician-confirmed OA diagnosis and 5764 non-cases based on a complete review of the entire medical record, corresponding to a 23.1% OA prevalence. Reference Standard 2: 738 cases with physician-documented OA diagnosis in provided in their medical problem list only and 6762 non-cases, corresponding to a 9.8% OA prevalence.

b Post-Test Prevalence: Prevalence of OA within health administrative data among the 7500 patients in the validation cohort (true positive + false positive cases)/7500.

Fig. 1. Annual crude prevalence of OA per 100 population among adults aged 20 years and older in Ontario for the years 2000–2017. Abbreviations: H = hospitalization; P = physician billing claim.
Fig. 2. Annual age- and sex-standardized prevalence per 100 population among adults aged 20 years and older in Ontario for the years 2000–2017. Abbreviations: H = hospitalization; P = physician billing claim.

Fig. 3. Annual age- and sex-standardized incidence of OA per 100 population among adults aged 20 years and older in Ontario for the years 2000–2017. Abbreviations: H = hospitalization; P = physician billing claim.

Fig. 4. Counts of New OA cases and Total OA cases across algorithms for 2017 identified among the 10, 969, 729 adult Ontario population. Abbreviations: H = hospitalization; P = physician billing claim.
over the past decade.

In our comprehensive review of primary care EMRs of a large random sample, only 58% of OA patients had their diagnosis documented in their medical problem list. These findings are consistent with prior studies [38–40] that have found OA to be under-recognized in primary care. As many as 60–90% of people with OA have at least one other condition, most often diabetes, hypertension, dyslipidemia and heart disease [41, 42]. In the setting of other ‘more important’ health problems, people with OA and their physicians may place OA at a lower priority for attention. However, this under-documentation of OA diagnoses may suggest that medical problem lists may be inadequate to identify OA patients for secondary research purposes. Although there are limited reports on identifying OA patients in Canadian primary care EMRs, Birtwistle et al. identified OA patients using a combination of both ICD-9 codes and diagnoses documented in problem lists and identified an OA prevalence of 14% [43]. In our large validation sample, our OA prevalence based on diagnoses in medical problem lists was only 10%. However, we did not also ascertain patients using ICD codes contained within the EMR (as it would bias our algorithm performance). More recently, reports of lower overall completeness of problem lists [22] further suggest caution when relying only on problem lists for disease surveillance. Outside of Canada, the under-recording of arthritis diagnoses in primary care has been similarly reported [25,26]. It may be that only more severe OA is associated with formal documentation; OA may not be prioritized as a co-existing medical condition warranting a documentation, or other reasons.

Among previous studies examining the accuracy of administrative data OA algorithms [17–20,44–49], only a few were conducted in Canada [17–20]. Past OA reference standards used in Canadian studies have varied (from patient-self reported diagnosis to diagnosis confirmation via imaging reports), and have uniformly been associated with suboptimal sensitivities and predictive values [17–20]. A prior study also used a range of reference standards that variable combined OA on imaging, self-report, and OA diagnostic criteria in 171 Canadians with knee pain invited for OA screening, and reported sensitivities of 21–57% using administrative data but high PPVs (82–100%) [19]. Their low sensitivity of identifying OA patients meeting diagnostic criteria may be explained by their screening process (potentially identifying previously undiagnosed OA) and their high PPVs may be explained by the high prevalence of OA within their reference standard and exclusion of individuals with other musculoskeletal issues. In our study, we chose a priori to evaluate the performance of administrative data algorithms against two OA reference standards. We opted not to include OA diagnostic criteria as a standard, or confine OA to those with confirmed OA on imaging as primary care physicians do not consistently use these tools and they are not required for diagnosis [50]. While there was less classification of OA when classifying diagnoses during all medical encounters, we suspected reference standard 1 would most likely include individuals across all spectrums of the disease, and thus administrative data algorithms would have lower sensitivity, but higher PPVs due to fewer false positives in administrative data. We also suspect that individuals labelled with OA in problem lists (reference standard 2) may have more health contacts related to their OA (to warrant population the problem list) and therefore administrative data algorithms would have higher sensitivity at detecting these OA cases. However, as reference standard 2 completely misses some true OA, the PPVs estimated against reference standards 2 should be interpreted with caution.

Our findings also have important implications for other settings, particularly US insurance claims databases, which may have less extensive data histories on individual patients than here in Canada. The sensitivity of administrative data algorithms to identify OA in these settings may be even lower, leading to even greater underestimates of the burden of OA. Additionally, the doubling of the OA prevalence in the population over our study’s timeframe has important implications for policy making, health services planning, raising awareness, and informing funding allocations for OA care and research. Over time, the increasing pain, loss of function, diminished quality of life and increased disability [51–54] associated with OA will contribute to increasing financial and physical burden on individuals and society [55–58]. OA care is provided mainly in ambulatory or outpatient settings, and already comprises a substantial burden of primary care office visits [59–63]. Moreover, Canadian wait times to see musculoskeletal specialists – including rheumatologists and orthopedic surgeons – are the longest among subspecialists indicating capacity of providing quality OA care needs to be increased [64,65]. With our findings indicating that OA is now affecting 1 in 5 adults, mobilizing planning for appropriate health service needs to manage these individuals is urgently needed.

There have been some reports using computerized health data to describe the incidence and prevalence of OA [1,7,66–72] but limited reports evaluating trends over time. A UK study on temporal trends between 1992 and 2013 reported that the annual age–sex standardized incidence of diagnosed OA showed a small increase over the period 1992–2004 but remained stable or even declined thereafter and that similar trends were observed for both sexes [70]. In contrast, an earlier study from British Columbia, Canada reported that the age-standardized incidence did not change among males but increased among females over 1996–2004 [66]. In our study which spans more recent years, we observed higher incidence during the initial observational period than in more recent years; however, we did not observe any differences in incidence trends between sexes.

In our validation sample, we detected a high prevalence of doctor-diagnosed OA in primary care (23%), with 36% having OA among individuals 45 years and older. The overarching purpose of public health surveillance for OA is to facilitate the primary, secondary and tertiary prevention of OA. This includes being able to accurately identify the total number of individuals with OA and the number of patients diagnosed annually. Across the top performing algorithms, the crude population-level OA prevalence ranged from 11% to 25% in 2017. Algorithms with lower sensitivity identified lower prevalence and may not be as useful for public health surveillance. Using our most sensitive algorithm, we identified 2,775,888 Ontario patients with at least one hospital or physician diagnosis code over the study period, and over a million OA patients are undetected if requiring one additional diagnosis code. Thus, careful consideration of the algorithm is required in any surveillance system aimed at detecting OA individuals.

There are potential limitations of our study. First, we assumed that a physician diagnosis of OA in our validation sample was correct and complete. We did not require patients to satisfy classification criteria as these are not routinely employed or documented during the course of clinical care. OA diagnoses were taken verbatim and no inferences were made from signs or symptoms. However, the validity of the diagnosis made by a physician was not the study’s objective. Second, our findings are based on patients who are rostered to a primary care physician and thus may not be generalizable to patients without access to primary care. Third, changes to physician coding practices in Ontario, including the introduction of different payment models, and higher remuneration for the care of other diseases, such as diabetes, may have influenced the sensitivity of administrative data to detect OA among individuals with multi-morbidity. This is not a limitation per se, but a caveat of the potential for diagnostic overshadowing in billing claims where only one diagnosis code is permitted per billing claim. Finally, we may have misclassified prevalent OA as incident OA in the population analyses, especially during the early years. The early years of data coverage are most prone to misclassification of prevalent OA as incident OA whereby incidence appears to be highest in the earlier years as there are more prevalent OA cases being misclassified as incident. Previous studies have shown that a 5-year washout period may be sufficient but a 10-year washout period (which we employed) further reduces misclassification and is preferred [73,74]. Moreover, it is difficult to state when OA actually becomes ‘incident’, since mild OA can be present quite early in life, and worsen slowly over decades.
In conclusion, improving efforts to encourage the documentation of OA diagnoses in both EMR and administrative data will enhance the surveillance and evaluation efforts of OA. Health administrative data appear to be more sensitive at detecting OA patients for whom their physician formally documented their diagnosis in medical problem lists. Potential reasons for this require ongoing investigation to determine whether these individuals have more severe OA, are actively receiving care for their OA, diagnostic overshadowing in the presence of multimorbidity, or other reasons. Additionally, we observed the prevalence of OA to be increasing over time, irrespective of the algorithm used to identify OA, whereas annual incidence appears to be stable.

Author contributions

Jessica Widdifield was involved in conception and design, obtaining funding, acquisition of data analysis and interpretation of the data, manuscript preparation, and takes responsibility for the integrity of the work.

R. Liisa Jaakkimainen was involved in conception and design, obtaining funding, acquisition of data analysis and interpretation of the data, and manuscript preparation.

Jodi Gatley was involved in design, analysis and interpretation of the data, administrative, technical, and logistic support, and manuscript preparation.

Lisa Lix was involved in design, analysis and interpretation of the data, and manuscript preparation.

Gillian Hawker was involved in design, obtaining funding, analysis and interpretation of the data, and manuscript preparation.

Sasha Bernatsky was involved in design, analysis and interpretation of the data, and manuscript preparation.

Bheeshma Ravi was involved in design, analysis and interpretation of the data, and manuscript preparation.

David Wasserstein was involved in design, analysis and interpretation of the data, and manuscript preparation.

Bing Yu was involved in design, acquisition of data, analysis and interpretation of the data, administrative, technical, and logistic support, and manuscript preparation.

Karen Tu was involved in conception and design, obtaining funding, acquisition of data, analysis and interpretation of the data, and manuscript preparation.

Role of the funding source

This work was supported by the Canadian Institutes of Health Research and the Public Health Agency of Canada, who played no role in the design or conduct of the study.

Declaration of competing interest

None of the authors have conflicts of interests related to this study.

Acknowledgements

First and foremost, we wish to thank all the family physicians who provide data to the Electronic Medical Record Primary Care Database (EMRPC), also known as the Electronic Medical Record Administrative data Linked Database (EMRALD). We would like to thank our five chart abstractors for their work (Nancy Cooper, Abayomi Fowora, Diane Kerbel, Anne Marie Mior and Barbara Thompson), health information analysts Myra Wang and Jacqueline Young for assistance in preparing the reference standard sample, and project management support from Raquel Duchen (ICES).

This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). This study also received funding from the Public Health Agency of Canada and the Canadian Institutes of Health Research.

Parts of this material are based on data and information compiled and provided by the MOHLTC, and the Canadian Institutes of Health Information (CIHI). The analyses, conclusions, opinions and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

Karen Tu is supported by a research scholar award by the Department of Family and Community Medicine at the University of Toronto. Jessica Widdifield receives support from the Arthritis Society Stars Career Development Award (STAR-19-0610). Lisa Lix is supported by a Tier 1 Canada Research Chair.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ocearto.2020.100115.

References

[1] Global Burden of Disease Study Collaborators, Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the global burden of disease study 2013, Lancet 86 (2015) 743-800.

[2] S.G.S. Dagenais, E.K. Wai, Systematic review of the prevalence of radiographic primary hip osteoarthritis, Clin. Orthop. Relat. Res. 467 (2009) 623-637.

[3] Public Health Agency of Canada, National Surveillance of Osteoarthritis and Rheumatoid Arthritis in Canada: Results from the Canadian Chronic Disease Surveillance System, PHAC, Ottawa, 2017.

[4] R.K.N. Plotnikoff, E. Lytvyak, C. Penfold, D. Schofplocher, I. Imayama, S.T. Johnson, K. Raine, Osteoarthritis prevalence and modifiable factors: a population study, BMC Public Health 15 (2015) 1195.

[5] K.V.S.C. MacDonald, K. Langlois, D.A. Marshall, Symptom onset, diagnosis and management of osteoarthritis, Health Rep. 25 (2014) 10–17.

[6] L.B. Murphy, J.J. Sacks, C.G. Helmick, T.J. Brady, K.E. Barber, J.M. Hootman, et al., Arthritis prevalence: which case definition should be used for surveillance? Comment on the article by jafarzadeh and felsons, Arthritis Rheum. 71 (2019) 172−175.

[7] D.A. Marshall, R. Shakib, S. Bertazzon, J.E. Seidell, A.B. Patel, M. Nazr, et al., Geographic variation in osteoarthritis prevalence in Alberta: a spatial analysis approach 103 (2019) 112−121.

[8] J.A. Kopec, J.M. Bertholot, C. Le Petit, J. Aghajanian, E.C. Sayre, J. Cibere, et al., Descriptive epidemiology of osteoarthritis in british columbia, Canada, J. Rheumatol. 34 (2007) 386−392.

[9] J.A. Kopec, E.C. Sayre, J. Cibere, W.M. Flanagan, J. Aghajanian, A.H. Anis, et al., Trends in physician-diagnosed osteoarthritis incidence in an administrative database in british columbia, Canada, 1996-1997 through 2003-2004, Arthritis Rheum. 59 (2008) 929-934.

[10] M.M. Rahman, C.H. Goldsmith, A.H. Anis, J.A. Kopec, Osteoarthritis incidence and trends in administrative health records from british columbia, Canada, J. Rheumatol. 41 (2014) 1147−1154.

[11] J.G.K. Sun, L.W. Svenson, N.R. Bell, C. Frank, Estimating osteoarthritis incidence from population-based administrative health care databases, Ann. Epidemiol. 17 (2006) 51−56.

[12] D.A. Marshall, C. Barnabe, K.V. MacDonald, C. Maxwell, D. Mosher, T. Waylayk, et al., Estimating the burden of osteoarthritis to plan for the future, Arthritis Care Res. 67 (2015) 1379−1386.

[13] D.P.G. Yu, J. Bedson, K.P. Jordan, Annual consultation incidence of osteoarthritis estimated from population-based health care data in england, Rheumatology 54 (2015) 2051-2060.

[14] K.E. Powell, R.J. Presley, D. Tolsona, S. Harris, K.J. Mertz, K. Viel, et al., Administrative data as a tool for arthritis surveillance: estimating prevalence and utilization of services, J. Publ. Health Manag. Pract. 9 (2003) 291−298.

[15] L.R. Harrold, W. Straus, S.E. Andrade, J.L. Reed, J. Cernieux, B.E. Lewis, J.H. Gurwitz, Challenges of estimating health service utilization for osteoarthritis patients on a population level, J. Rheumatol. 29 (2002) 1931−1936.

[16] S.M. Cadorette, L. Wong, An introduction to health care administrative data, Can. J. Hosp. Pharm. 68 (2015) 232−237.

[17] L. Lix, J. Mann, Defining and Validating Chronic Diseases: an Administrative Data Approach – an Update with Icd-10-Ca, Winnipeg, 2008.

[18] L. Lix, C. Burchill, C. Metge, N. McKeen, D. Moore, R. Bond, Defining and Validating Chronic Diseases: an Administrative Data Approach, Winnipeg, 2006.

[19] M.M. Rahman, C.H. Goldsmith, A.H. Anis, J. Cibere, Validation of administrative osteoarthritis diagnosis using a clinical and radiological population-based cohort, International Journal of Rheumatology 2016 (2016) 1−7.

[20] M. Fortin, S. Sanche, J. Almirall, Self-reported versus health administrative data: implications for assessing chronic illness burden in populations. A cross-sectional study, CMAJ Open 5 (2017) E729−E733.

[21] A. Wright, A.B. McCoy, T.T. Hickman, D.S. Hilaire, D. Borbolla, W.A. Bowes 3rd, et al., Problem list completeness in electronic health records: a multi-site study and assessment of success factors, Int. J. Med. Inf. 84 (2015) 784−790.
