BMJ Open

Validation of billing code combinations to identify cardiovascular magnetic resonance imaging scans in Ontario, Canada: a retrospective cohort study

Idan Roifman,[1,2,3] Feng Qiu,[3] Kim A Connelly,[1,4] Graham A Wright,[1] Michael Farkouh,[2,5] Laura Jimenez-Juan,[1] Harindra C Wijeysundera[1,2,3]

Check for updates

[1]Department of Medicine, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada
[2]Department of Medicine, Institute of Health Policy Management and Evaluation, Toronto, Ontario, Canada
[3]Department of Medicine, Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada
[4]Department of Medicine, St. Michael's Hospital, Toronto, Ontario, Canada
[5]Department of Medicine, Toronto General Hospital, Toronto, Ontario, Canada

Correspondence to
Dr Idan Roifman;
idan.roifman@sunnybrook.ca

ABSTRACT

Objectives Cardiovascular magnetic resonance (CMR) imaging is the gold-standard test for the assessment of heart function. Despite its importance, many jurisdictions lack specific billing codes that can be used to identify patient receipt of CMR in administrative databases, limiting the ability to perform 'big data' CMR studies. Our objective was to identify the optimal billing code combination to identify patients who underwent CMR using administrative data in Ontario.

Design Retrospective cohort study.

Setting Quaternary care academic referral centre in Ontario, Canada.

Participants We tested all billing code combinations in order to identify the optimal one to determine receipt of CMR. The reference gold standard was a list of all cardiothoracic magnetic resonance scans performed at Sunnybrook Health Sciences Centre between 1 January 2014 and 31 December 2016, verified by chart audit. We assessed the diagnostic performance (accuracy, sensitivity, specificity, positive predictive value and negative predictive value) for all code combinations.

Results Our gold-standard cohort consisted of 2339 thoracic MRIs that were performed at Sunnybrook Health Sciences Centre from 1 January 2014 to 31 December 2016. Of these, 2139 (91.5%) were CMRs and 200 (8.5%) were chest MRIs. We identified the most accurate billing combination for the determination of patient receipt of CMR. This combination resulted in an accuracy of 95.3% (95% CI 94.4% to 96.2%), sensitivity of 97.4% (95% CI 96.6% to 98.1%), specificity of 86.4% (95% CI 83.1% to 89.6%), positive predictive value of 96.9% (95% CI 96.1% to 97.6%) and negative predictive value of 88.4% (95% CI 85.4% to 91.5%).

Conclusions Our study is the first to verify the ability to accurately identify patient receipt of CMR using administrative data, facilitating more robust population-based CMR studies in the future.

Strengths and limitations of this study

► Cardiovascular magnetic resonance (CMR) imaging is the gold-standard test to assess heart structure and function.

► Our study validated a combination of billing codes that can identify patient receipt of CMR in administrative databases in Ontario with an accuracy of approximately 95.3%.

► This validated combination of codes can be used by researchers in the future in order to identify those who underwent CMR and conduct CMR-related population-based 'big data' studies.

► Only data from a single centre in Ontario were used, thus potentially limiting generalisability. Further, the reported positive predictive values are very high due to 91.5% of the scans being true CMRs. If there was a setting in which there would be a higher prevalence of non-cardiac chest MRIs, the positive predictive value would be expected to be lower.

INTRODUCTION

Cardiovascular magnetic resonance (CMR) imaging is an increasingly used and important clinical imaging modality and is considered the gold standard in the assessment of the structure and function of the human heart.[1–5] Despite the growth of CMR-related technology, there are no published population-based studies using CMR. Such studies are important in order to assess real-world outcomes related to this emerging technology as well as to understand health services issues such as access and cost. The main barrier to conducting such studies is the lack of a reliable mechanism to track those who undergo CMR on a population basis. In Canada, for example, most jurisdictions lack a unique billing code for CMR, thus limiting the accuracy of tracking. One of those jurisdictions is Ontario, the largest Canadian province and home to approximately 14 million people. In Ontario, CMRs are billed with a generic code for 'thoracic MRI' which includes both CMRs and chest magnetic resonance scans.[6 7] Further, there are two additional billing codes that are commonly used for

CMR scans; one for 'cardiac gating' (monitoring the patient's ECG during the CMR scan) and the other for repeated MRI sequences that are sometimes performed as part of a comprehensive CMR scan.[7] Administrative medical coding data are known to have inaccuracies. Therefore, it is important to validate these data using medical chart review as the gold-standard comparator. It is currently unknown if different combinations of the three billing codes can identify those who undergo CMR accurately. The objective of this paper was to identify the optimal billing code combination to identify patients who undergo CMR using administrative data in Ontario while using a gold-standard reference cohort composed of CMR scans that were confirmed based on a chart audit.

METHODS
Administrative databases
Given Ontario's universal healthcare, administrative databases in Ontario contain data on all patients in the population who undergo assessment and billing by a physician. In fact, the Ontario Health Insurance Plan database captures approximately 98% of physician billings in the province.[8]

Candidate billing codes in for CMR imaging
As previously mentioned, there is no specific billing code for CMR imaging in Ontario. Candidate codes were selected based on examination of local billing patterns and after extensive consultation with CMR specialists in six large referral centres across Ontario. Since CMR is a niche cardiac diagnostic test in Canada, these six centres perform that vast majority of all CMR scans in the province. All patients scanned are billed code X441 ('MRI thorax-multislice sequence'). This is the generic code for thoracic MRI which includes both cardiovascular and chest MRIs. In addition, there are two other related billing codes that may also be billed. One of these is X445: ('MRI thorax-repeat another plane, different pulse sequence'). This code is billed when scanning requires the performance of additional MRI sequences in different anatomical planes. This is often the case when the scan is performed to address clinical questions regarding cardiac anatomy. Finally, X486 ('addition of cardiac gating') may also be billed. This code is used when ECG recordings and gating are necessary while the patient is being scanned. In practice, this is required to improve image quality of movie images (cine images) of the beating heart and to monitor patients who may be at risk of arrhythmia. We sought to identify the diagnostic properties (accuracy, sensitivity, specificity, negative predictive value and positive predictive value) for the following combination of codes:
A. X441+X445.
B. X441+X486.
C. X441+X445+X486.

The gold-standard reference
All thoracic MRI (CMR and chest MRIs) scans performed during a 3-year period, between 1 January 2014 and 31 December 2016, at Sunnybrook Health Sciences Centre were included in this study and were deemed the gold-standard reference. Scans were identified using a list obtained from the records of the department of medical imaging. Next, via chart review, each scan on this list was checked by the principal author (IR), a trained CMR specialist, to ensure that the characterisation of cardiovascular versus chest MRI was indeed correct. There were no discrepancies between the list obtained by medical imaging and the chart review. Note that individual patients may have been represented twice within this cohort if they had multiple thoracic MRI scans. The unit of analysis was individual MRI scans, not patients.

The approach we employed to ascertain the validity of administrative data definitions using data obtained from chart review as the gold standard is similar to that used to validate administrative data definitions and algorithms for a number of medical conditions including myocardial infarction, rheumatoid arthritis, Alzheimer's disease and multiple sclerosis.[9–13]

Statistical analysis
Sensitivity, specificity, accuracy, positive and negative predictive values were then calculated for each of the three combination of codes listed above, using the validated list obtained from the Department of Medical Imaging as the gold-standard reference. Sensitivity was calculated as true positives/true positives plus true negatives, specificity was calculated as false negatives/false negatives plus false positives, positive

Table 1 Characteristics of the gold-standard patient cohort of patients undergoing cardiothoracic MRIs at Sunnybrook Health Sciences Centre from 1 January 2014 to 31 December 2016

n	2339
Age in years (mean±SD)	59±17
Sex (% female)	36
Active smokers (%)	14
Diabetes mellitus (%)	23
Dyslipidaemia (%)	41
Hypertension (%)	53
Outpatient (%)	87
Indication for cardiovascular MRI (%)	
Non-ischaemic cardiomyopathy	30
Viability	12
Arrhythmogenic right ventricular cardiomyopathy	11
Pulmonary veins	21
Aortic/vascular	16
Research (not included in the analyses)	9

Table 2 Diagnostic performance of the different thoracic magnetic resonance code combinations for the identification of cardiovascular magnetic resonance scans

	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	Accuracy (95% CI)
X441+X445	99.1 (98.7 to 99.6)	0.1 (0.0 to 0.1)	81.1 (79.4 to 82.7)	11.1 (0.0 to 25.6)	80.5 (78.9 to 82.1)
X441+X486	97.4 (96.6 to 98.1)	86.4 (83.1 to 89.6)	96.9 (96.1 to 97.6)	88.4 (85.4 to 91.5)	95.3 (94.4 to 96.2)
X441+X445+X486	97.2 (96.4 to 97.9)	86.4 (83.1 to 89.6)	96.8 (96.1 to 97.6)	87.6 (84.5 to 90.7)	95.1 (94.2 to 96.0)

All numbers are expressed as percentages.

predictive value was calculated as true positives/true positives plus false positives and negative predictive value was calculated as false negatives/false negatives plus true negatives. Accuracy was calculated as true positives+true negatives/true positives+true negatives+false positives+false negatives. A 7-day window was allowed for lag time between the date of the scan on the gold-standard list and submission of billing codes to the ministry of health and long-term care, for payment. A sensitivity analysis using a window of <24 hours was also performed. Finally, subgroup analyses were performed according to inpatient/outpatient status. Binomial exact 95% CIs (by the Clopper-Pearson method) were calculated for all measures of diagnostic performance. All analyses were conducted using SAS V.9.4 (SAS Institute).

Patient and public Involvement
Patients were not involved in this study.

RESULTS
Our gold-standard cohort consisted of 2339 cardiothoracic MRIs that were performed at Sunnybrook Health Sciences Centre from 1 January 2014 to 31 December 2016. Of these, 2139 (91.5%) were cardiovascular MRIs and 200 (8.5%) were chest MRIs. Patient characteristics of those undergoing cardiothoracic MRIs are summarised in table 1. All clinical patients undergoing cardiothoracic MRI (ie, excluding those with research scans) received code X441.

The results of our analyses are found in table 2. The optimal combination of codes was found to be X441+X486. This combination resulted in an accuracy of 95.3%. The combination of all three codes (X441, X486 and X445) had comparable results, with an accuracy of

95.1%. In contrast the combination of X441 and X445 produced inferior results with an accuracy of only approximately 80.5% and unacceptably low specificity and negative predictive value numbers.

There was no substantive change when using a window of <24 hours between the billing of the individual codes with the best combination remaining X441+X486 and the worst performing combination remaining X441+X445 (see table 3). There were also no clinically important differences when examining subgroups of patients by inpatient/outpatient status with the best performing combination continuing to be X441+X486 (see table 4A,B).

DISCUSSION
In the first study to examine the optimal mechanism to identify patients who underwent CMR using administrative data, we were able to identify the best and worst performing combinations of billing codes. The combination of the billing codes X441+X486 was optimal for identifying patients who underwent CMR, doing so with an accuracy of approximately 95.3%. This combination was marginally more accurate than the combination of X441+X445+X486, which provided an accuracy of 95.1%. In contrast, the combination of X441+X445 produced the worst results, with an accuracy of approximately 80.5%.

There has been a plethora of work validating diagnostic algorithms for a variety of cardiovascular indications including stroke, atrial fibrillation, cardiovascular procedures and major adverse cardiovascular events using administrative data.[4 8 11 12 14–19] In contrast, there has been relatively little work attempting to validate patient receipt of cardiovascular imaging using administrative data. A

Table 3 Diagnostic performance of the different thoracic magnetic resonance code combinations for the identification of cardiovascular magnetic resonance (CMR) scans: sensitivity analysis using a window of <24 hours between performance of the CMR and submission of billing code for payment

	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	Accuracy (95% CI)
X441+X445	99.1 (98.7 to 99.6)	0.5 (0 to 1.1)	81.7 (80.0 to 83.2)	11.1 (0.0 to 25.6)	81.1 (79.5 to 82.7)
X441+X486	97.3 (96.6 to 98.1)	89.0 (86.0 to 92.0)	97.5 (96.8 to 98.2)	88.1 (85.0 to 91.2)	95.8 (95.0 to 96.6)
X441+X445+X486	97.1 (96.3 to 97.9)	89.0 (86.0 to 92.0)	97.5 (96.8 to 98.2)	87.3 (84.1 to 90.5)	95.6 (94.8 to 96.4)

All numbers are expressed as percentages

Table 4A Diagnostic performance of the different thoracic magnetic resonance code combinations for the identification of cardiovascular magnetic resonance scans among inpatients

	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Accuracy
X441+X445	98.7 (97.2–100.0)	62.5 (0.0–14.6)	88.2 (84.2–92.2)	40.0 (0.0–82.94)	87.3 (83.2–91.4)
X441+X486	97.4 (95.3–99.4)	81.3 (67.9–94.8)	97.4 (95.3–99.4)	81.3 (67.7–94.8)	95.4 (92.8–98.0)
X441+X445+X486	97.4 (95.3–99.4)	81.3 (67.8–94.8)	97.4 (95.3–99.4)	81.3 (67.8–94.8)	95.4 (92.8–98.0)

All numbers are expressed as percentages.

recent study from Ontario demonstrated that in inpatients who suffered a stroke, administrative data alone were able to identify those who had MRIs or CTs of the brain in over 90% of cases.[17] In contrast, in that study, administrative data were only able to identify transthoracic echocardiograms in approximately 40% of inpatients.[17]

We are not aware of any studies that examined the validity of identifying CMR using administrative data. CMR is an increasingly used technology and is now considered to be the gold standard for the assessment of right and left ventricular function, viability and flow. However, despite this status, CMR does not have a unique billing code in Ontario or most other Canadian jurisdictions.[7 20 21] This has been perceived to be a barrier to identifying those who underwent CMR using administrative data, thus limiting the extent of 'big data' population-based CMR studies.

IMPORTANCE

Our study is the first to demonstrate that CMRs can be accurately tracked using a simple combination of billing codes with high accuracy. Given the similarity of billing codes for CMR across Canada, our approach will be translatable to other jurisdictions. By using our validated coding combinations, one can identify patients who underwent this important diagnostic test in order to perform robust population-based outcomes studies related to CMR.

LIMITATIONS

This study must be interpreted in the context of its limitations. First, specificity and the related negative predictive value were modest even for the best performing of our algorithms. This is due to the fact that we have very few true and false negatives in our cohort. With that said, the accuracy for the best of our coding combinations

exceeded 95% using the cohort of patients who underwent thoracic MRIs. Second, the Institute for Clinical Evaluative Sciences has strict policies on data confidentiality and minimising the risk of identification of subjects. In this context, patients who were analysed in a de-identified manner cannot be reidentified at a later time. Identification would allow us to access medical charts and determine the patients' demographic characteristics. Due to these restrictions, we were unable to discern the patient characteristics of the approximately 5% of CMR scans in which the billing code combinations were inaccurate. Third, the gold-standard cohort was obtained from a single centre in Ontario. However, this is unlikely to affect the validity of our results given the similarities in billing practices across the province and the fact that the billing codes themselves are identical across the province. Fourth, while these results are generalisable to Canada, they may not be generalisable to other countries who use different, more specific CMR codes. Fifth, the reported positive predictive values are very high due to 91.5% of the scans being true CMRs. If there was a setting in which there would be a higher prevalence of non-cardiac chest MRIs, the positive predictive value would be expected to be lower. Finally, miscoding of administrative data, transcription errors by physicians or clerical staff submitting billing claims and in the case of those physicians who work under a salary model and 'shadow bill', improper documentation of diagnostic billing codes may all have affected the accuracy of the administrative data. Physicians in Ontario who are reimbursed under a salary structure often are required to send patient billings to the Ministry of Health. These billings are not used for directly reimbursing the physician but rather to ensure that the salary provided is relatively commensurate with the services provided. These 'shadow' billings are not counted by the Ontario Ministry of Health and thus are

Table 4B Diagnostic performance of the different thoracic magnetic resonance code combinations for the identification of cardiovascular magnetic resonance scans among outpatients

	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Accuracy
X441+X445	99.2 (98.8–99.6)	0.0 (0.0–0.0)	90.7 (89.4–92.1)	0.0 (0.0–1.1)	90.1 (88.7–91.5)
X441+X486	97.3 (96.5–98.1)	77.8 (71.4–84.0)	97.8 (97.0–98.5)	74.6 (68.1–81.2)	95.5 (94.5–96.5)
X441+X445+X486	97.1 (96.3–97.9)	77.7 (71.4–84.4)	97.7 (97.0–98.5)	72.9 (66.3–79.4)	95.3 (94.3–96.3)

All numbers are expressed as percentages.

not traceable in administrative databases. However, the percentage of physicians who shadow bill in Ontario is <2% and our study found high accuracy for administrative data definitions using a gold-standard reference model based on chart review.

CONCLUSIONS

Our study is the first to report the ability to identify patient receipt of CMR using administrative data with high accuracy. These findings are important as they provide a validated method to identify those who undergo this diagnostic test in order to perform population-based 'big data' studies in the future.

REFERENCES

1. Society for Cardiovascular Magnetic Resonance, 2017. CMS appropriate use criteria program for advanced diagnostic imaging services https://scmrsite-ymcom/news/349512/CMS-Appropriate-Use-Criteria-Program-for-Advanced-Diagnostic-Imaging-Services.htm.
2. Ronan G, Wolk MJ, Bailey SR, *et al*. ACCF/AHA/ASE/ASNC/HFSA/HRS/SCAI/SCCT/SCMR/STS 2013 multimodality appropriate use criteria for the detection and risk assessment of stable ischemic heart disease: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Heart Association, American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance, and Society of Thoracic Surgeons. *J Nucl Cardiol* 2014;21:192–220.
3. Roifman I, Zia MI, Zavodni A, *et al*. Evolution of right ventricular function post-acute ST elevation myocardial infarction. *J Magn Reson Imaging* 2014;40:709–14.
4. Roifman I, Wijeysundera HC, Austin PC, *et al*. Temporal trends in the utilization of noninvasive diagnostic tests for coronary artery disease in ontario between 2008 and 2014: a population-based study. *Can J Cardiol* 2017;33:279–82.
5. Patel MR, White RD, Abbara S, *et al*. 2013 ACCF/ACR/ASE/ASNC/SCCT/SCMR appropriate utilization of cardiovascular imaging in heart failure: a joint report of the American College of Radiology Appropriateness Criteria Committee and the American College of Cardiology Foundation Appropriate Use Criteria Task Force. *J Am Coll Cardiol* 2013;61:2207–31.
6. Chan BTB SS. *Supply and Utilization of General Practitioner and Family Physician Services in Ontario*: ICES Investigative Report Institute for Clinical Evaluative Sciences, 2005.
7. Ontario Ministry of Health and Long-Term Care. *Schedule of Benefits for Physician Services under the Health Insurance Act*, 2014.
8. Singh SM, Webster L, Calzavara A, *et al*. Validation of algorithms to identify invasive electrophysiology procedures using administrative data in Ontario, Canada. *Med Care* 2017;55:e44–e50.
9. Widdifield J, Ivers NM, Young J, *et al*. Development and validation of an administrative data algorithm to estimate the disease burden and epidemiology of multiple sclerosis in Ontario, Canada. *Mult Scler* 2015;21:1045–54.
10. Jaakkimainen RL, Bronskill SE, Tierney MC, *et al*. Identification of physician-diagnosed Alzheimer's disease and related dementias in population-based administrative data: a validation study using family physicians' electronic medical records. *J Alzheimers Dis* 2016;54:337–49.
11. Tu K, Mitiku T, Guo H, *et al*. Myocardial infarction and the validation of physician billing and hospitalization data using electronic medical records. *Chronic Dis Can* 2010;30:141–6.
12. Tu K, Mitiku T, Lee DS, *et al*. Validation of physician billing and hospitalization data to identify patients with ischemic heart disease using data from the Electronic Medical Record Administrative data Linked Database (EMRALD). *Can J Cardiol* 2010;26:e225–e228.
13. Widdifield J, Bernatsky S, Paterson JM, *et al*. Accuracy of Canadian health administrative databases in identifying patients with rheumatoid arthritis: a validation study using the medical records of rheumatologists. *Arthritis Care Res* 2013;65:1582–91.
14. Hall R, Mondor L, Porter J, *et al*. Accuracy of administrative data for the coding of acute stroke and TIAs. *Can J Neurol Sci* 2016;43:765–73.
15. Hughes C. Diagnosis coding for hypertension and heart failure. *Fam Pract Manag* 2017;24:34.
16. Mazzali C, Paganoni AM, Ieva F, *et al*. Methodological issues on the use of administrative data in healthcare research: the case of heart failure hospitalizations in Lombardy region, 2000 to 2012. *BMC Health Serv Res* 2016;16:234.
17. Porter J, Mondor L, Kapral MK, *et al*. How reliable are administrative data for capturing stroke patients and their care. *Cerebrovasc Dis Extra* 2016;6:96–106.
18. Quan H, Khan N, Hemmelgarn BR, *et al*. Validation of a case definition to define hypertension using administrative data. *Hypertension* 2009;54:1423–8.
19. Tu K, Nieuwlaat R, Cheng SY, *et al*. Identifying patients with atrial fibrillation in administrative data. *Can J Cardiol* 2016;32:1561–5.
20. Ontario Ministry of Health. *Ontario Ministry of Health and Long-Term Care: IntelliHEALTH ONTARIO*, 2015.
21. Ontario Ministry of Health and Long Term Care. *Bulletin 4540 - 2008 Physician Services Agreement – Changes Effective*:2011.