Infection Control & Hospital Epidemiology (2019), **40**, 1144–1150
doi:10.1017/ice.2019.184

Original Article

Using data linkage methodologies to augment healthcare-associated infection surveillance data

Seungwon Lee MPH[1], Paul Ronksley PhD[1,5], John Conly MD[2,3,5,6], Stephanie Garies MPH[1,7], Hude Quan MD, PhD[1,5], Peter Faris PhD[4], Bing Li MA[4] and Elizabeth Henderson PhD[1]

[1]Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada, [2]Department of Medicine, University of Calgary, Calgary, Alberta, Canada, [3]Alberta Health Services Infection Prevention and Control, Calgary, Alberta, Canada, [4]Alberta Health Services Research Facilitation Analytics (DIMR), Calgary, Alberta, Canada, [5]O'Brien Institute of Public Health, University of Calgary, Calgary, Alberta, Canada, [6]The Calvin, Phoebe and Joan Snyder Institute for Chronic Diseases, University of Calgary, Calgary, Alberta, Canada and [7]Department of Family Medicine, University of Calgary, Calgary, Alberta, Canada

Abstract

Background and objectives: The landscape of antimicrobial resistance (AMR) surveillance is changing rapidly. The primary objective of this study was to assess the benefit of linking population-based infection prevention and control surveillance data on methicillin-resistant *Staphylococcus aureus* (MRSA) to hospital discharge abstract data (DAD). We assessed the value of this novel data linkage for the characterization of hospital-acquired (HA) and community-acquired MRSA (CA-MRSA) cases.

Methods: Incident inpatient MRSA surveillance data for all adults (≥18 years) from 4 acute-care facilities in Calgary, Alberta, between April 1, 2011, and March 31, 2017, were linked to DAD. Personal health number (PHN) and gender were used to identify specific individuals, and specimen collection time-points were used to identify specific hospitalization records. A third common variable on admission date between these databases was used to validate the linkage process. Descriptive statistics were used to characterize HA-MRSA and CA-MRSA cases identified through the linkage process.

Results: A total of 2,430 surveillance records (94.6%) were successfully linked to the correct hospitalization period. By linking surveillance and administrative data, we were able to identify key differences between patients with HA- and CA-MRSA. These differences are consistent with previously reported findings in the literature. Data linkage to DAD may be a novel tool to enhance and augment the details of base surveillance data.

Conclusion and recommendations: This is the first Canadian study linking a frontline healthcare-associated infection AMR surveillance database to an administrative population database. This work represents an important methodological step toward complementing traditional AMR surveillance data practices. Data linkage to other data types, such as primary care, emergency, social, and biological data, may be the basis of achieving more precise data focused around AMR.

(Received 7 February 2019; accepted 7 June 2019; electronically published 29 July 2019)

The amount and types of antimicrobial resistance (AMR) have rapidly increased over the last few years.[1,2] In response, the World Health Organization (WHO) adopted the Global Action Plan on Antimicrobial Resistance in May 2015 and began working on the Global Antimicrobial Resistance Surveillance System (GLASS) to promote standardized AMR surveillance worldwide.[3] Furthermore, WHO recently published a high priority list of AMR pathogens on February 27, 2017.[4] These efforts represent attempts to harmonize aspects of global surveillance data.[5,6]

These published action plans stressed the importance of research and innovation, as well as accessibility to data in a timely manner. The use of surveillance information at the individual level provides patient safety-related intervention measures such as isolation. Although aggregated surveillance data can be used to track patient safety performance, access to detailed and diverse data related to patient health status and outcomes is often limited. Most surveillance records do not contain information on patient outcomes, comorbidities, and other pertinent information (eg, antibiotic usage) that could be used to inform prevention strategies on AMR. Chart review, which is time-consuming and expensive, remains the most commonly used strategy to obtain granular information on positive cases with AMR.

Administrative health data, which provide secondary information collected while delivering health care, have the potential to enhance surveillance efforts, and these data remain a relatively untapped resource in the setting of surveillance for HAIs. Drees *et al*[7] provided an overview of advantages and disadvantages of using administrative and surveillance database in healthcare epidemiology and antimicrobial stewardship research. These authors also provided examples of linkage work previously conducted on specific topics of central-line–associated bloodstream infections, surgical site infections, and ventilator-associated pneumonia.[7] To date, no

Author for correspondence: Dr Elizabeth Henderson, E-mail: ehenders@ucalgary.ca.
PREVIOUS PRESENTATION. This study was presented at the 2018 International Population Data Linkage Conference on September 13, 2018, in Banff, Alberta, Canada.
Cite this article: Lee S, *et al.* (2019). Using data linkage methodologies to augment healthcare-associated infection surveillance data. *Infection Control & Hospital Epidemiology*, 40: 1144–1150, https://doi.org/10.1017/ice.2019.184

published studies have linked administrative data to population-level AMR surveillance data in a Canadian setting.

Recognizing this data gap and the need to harmonize surveillance data, we sought to link incident inpatient methicillin-resistant *Staphylococcus aureus* (MRSA) surveillance data obtained by the infection prevention and control (IPC) department with provincial administrative hospitalization records as a means of augmenting traditional IPC surveillance systems and to inform patient safety and practice. We used MRSA as a representative AMR pathogen for analysis in our study because we anticipated an abundant number of cases compared to other pathogens. Specifically, we use this linked data source to characterize hospital-acquired and community-acquired MRSA cases within the province of Alberta, Canada.

Methods

Study population

The province of Alberta spans 661,848 km^2 with a current population of 4.1 million residents.[8] All health care in Alberta is delivered by Alberta Health Services (AHS). The provincial AHS is further divided into 5 zones: Calgary, Edmonton, Central, North, and South. The Calgary zone provides healthcare delivery to 1.5 million residents with ~1.4 million within the Calgary metropolitan region. The 4 adult acute-care facilities providing healthcare delivery within Calgary were analyzed. The number of acute-care beds at these institutions at the time of this study ranged from 269 to 1,095 (mean, 629). Pediatric patients (≤18 years old) were excluded.

Data sources

The provincial standardized in-hospital surveillance system, known as ProvSurv, is maintained by the Alberta Health Services Infection Prevention and Control Provincial Surveillance team. ProvSurv captures infection prevention and control (IPC) surveillance data from hospitals across Alberta and gives frontline IPC professionals the ability to follow and manage individual patient cases in any provincial hospital.[9] This provincial surveillance system is unique in Canada. Cultures are collected from a patient following established provincial standards.[10] All positive laboratory results, including those for AMR pathogens, are reported into the electronic medical record system as soon as the laboratory results become available. These positive culture results reported in patients admitted to AHS acute-care facilities undergo daily medical chart reviews by the IPC team. Surveillance case definitions (eg, HA vs CA vs HCA; colonization vs infection) for AMR pathogens (eg, MRSA) are applied and appropriate interventions (eg, isolation) are initiated and entered into ProvSurv system. The compiled cases further undergo data quality validation every quarter. These centralized rigorous case detection, classification, and validation processes ensure that each incident patient is counted only once from the admitted in-hospital population.

The Discharge Abstract Database (DAD) is a population-based administrative database that captures clinical and demographic information on hospital discharges. The DAD contains up to 25 diagnosis codes per hospital admission using the *International Classification of Diseases, Tenth Revision, Canada* (ICD-10-CA).[11] These data are submitted to the Canadian Institute for Health Information (CIHI), which sets national standards and training to ensure high-quality coding. The provincial data for the DAD are held by Alberta Health Services Analytics.

Table 1. Example Data Elements in ProvSurv Surveillance System

Variable	Description
Organism	Identified pathogen
Gender	Gender (coded male or female)
Classification	Hospital acquired, community acquired, or healthcare community associated
Encounter_type	Describes facility type (eg, inpatient, outpatient, long-term care, etc)
Encounter_date	Admission date of patient's hospitalization
Encounter_site	Facility name
Service	Hospital service
Culture_date	Date that culture was collected
Culture_site	Site of swab culture
Severity	Denotes colonization or infection
Decision	Indicates whether isolation protocol initiated

Table 2. Example Data Elements in Discharge Abstract Database (DAD)

Variable	Description
Inst	Institution
Ageadmit	Age at admission
Sex	Gender
Admitdate	Admission date
Instfrom	Transfer from institution no.
Admitcat	Admit category
Disp	Discharge disposition
Docsvc	Provider service
Doctype	Provider type
Dxcluster 1–25	Diagnosis cluster (1 being main cluster)
Dxcode 1–25	Diagnosis code (1 being main condition using up most resources in this admission)
Proccode	CCI intervention code, proccode1 being the main

Note. CCI, Charlson comorbidity index.

Data cleaning

Individual person-specific incident MRSA ProvSurv Surveillance records from 2011 to 2017 were extracted by the AHS IPC Provincial Surveillance Team. Corresponding data from the DAD were pulled from 2009 to 2017 to calculate the number of prior hospitalizations within 1 year from the earliest surveillance record. We inspected the extracted data elements to fully understand these databases. Additionally, we checked whether collected variables were coded consistently between 2 databases (eg, dates in similar format, clinical information coded in string, age reported in integer, or isolation status coded in Boolean format). Inconsistent variables (eg, gender, M vs male and F vs female) were recoded into similar formats and types. SAS version 9.4 software (SAS Institute, Cary, NC) was used for data cleaning and for statistical analyses. The data dictionaries of variables in these 2 databases, along with their short descriptions, are provided in Tables 1 and 2.

https://doi.org/10.1017/ice.2019.184 Published online by Cambridge University Press

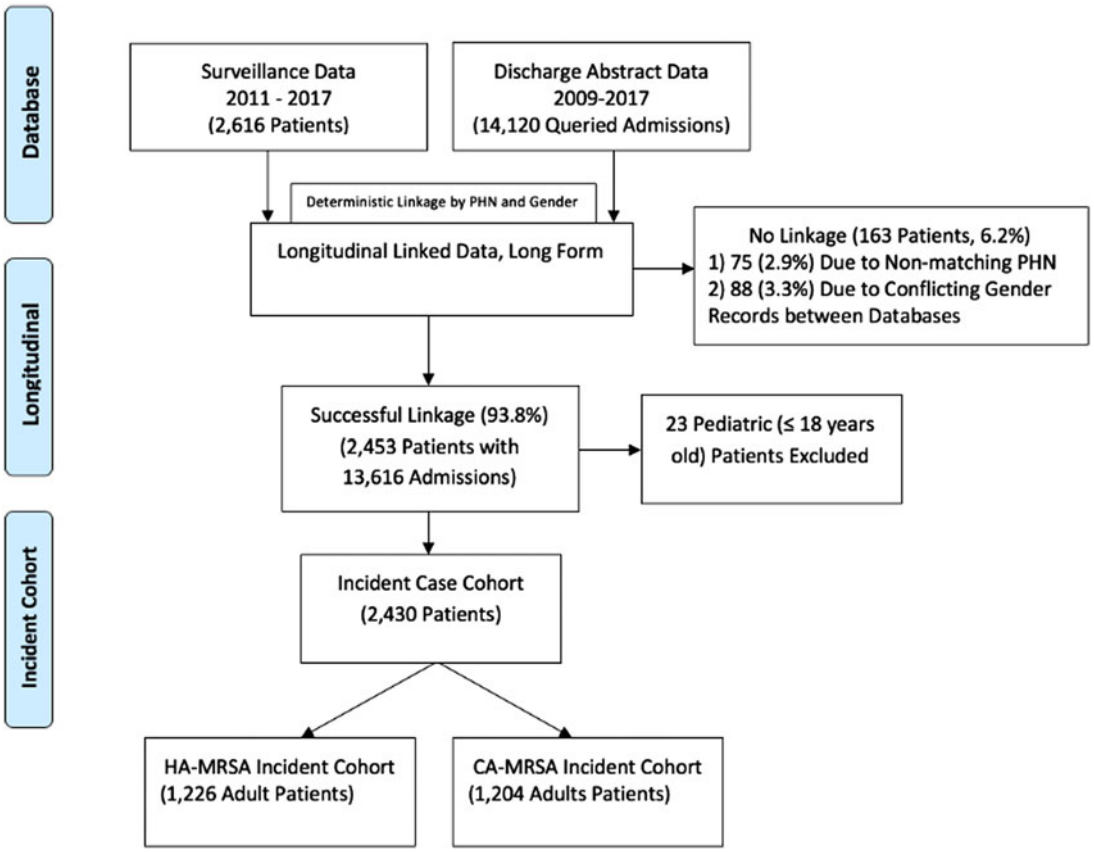

Fig. 1. Flow diagram depicting data linkage steps. The study population was admitted adult patients (>18 years old) who were identified with first-time incident MRSA in 4 acute-care facilities in Calgary.

Data linkage

Data from the DAD and ProvSurv were deterministically linked using a unique personal health number (PHN). The Alberta PHN is a unique lifetime identifier assigned to each individual within the province of Alberta when the person becomes eligible for basic coverage with the Alberta Health Care Insurance Plan (AHCIP).[12] Alberta's health insurance is structured as a single-payer system. Thus, only 1 unique PHN is assigned to 1 individual for accessing all aspects of healthcare services in Alberta. Initially, the 2 databases were linked using PHN only, then the linkage method was tested using both PHN and gender. Nonmatched records were dropped from the analysis. A longitudinal cohort was initially formed; each individual had multiple hospital admissions. We used date information to match incident MRSA records to a specific hospital admission record. The surveillance record contained swab sample collection date and positive result date for MRSA. This information was used to identify the DAD admission record where the incident MRSA case arose. We compared the admission dates to validate whether the correct admission period was linked. Furthermore, we explored the potential of developing a targeted MRSA screening algorithm using this linked data. We hypothesized that additional clinical granularity provided by the coded administrative data would allow this exploration.

Surveillance pathogen and case definitions

In Alberta, hospital-acquired MRSA (HA-MRSA) and community-acquired MRSA (CA-MRSA) onsets are classified based on the timing of specimen collection in relation to the timing of hospital admission (eg, HA-MRSA cases are those in which the specimen from which isolated MRSA was collected >48 hours after admission).[13–15] Healthcare-associated MRSA (HCA-MRSA) was defined as case patients identified with MRSA within 48 hours of admission who had had interactions (eg, indwelling catheter, resident of long-term care, etc) with the facility within 12 months, outside the 14-day facility attribution time period. These incident case definitions used in Alberta were developed and approved by the provincial surveillance committee and are analogous[16] to surveillance definitions used elsewhere. MRSA surveillance incidence data from April 1, 2011, to March 31, 2017, were used. The exposures associated with HCA-MRSA may occur at outpatient settings; thus, linked data may have been incomplete for this category.

Previous literature indicated that those who are at greater risk of HA-MRSA tend to be exposed to healthcare more frequently and are hospitalized longer.[17] Those typically diagnosed with CA-MRSA tend to be younger and healthier. Thus, these 2 groups were treated as separate subpopulations in our study and our choice of analytical methods reflected these differences.

Statistical analysis

Descriptive statistics were conducted for HA-MRSA and CA-MRSA cases. These categories were further described as infection (clinical disease) and colonization (no active disease) based on discrete definitions. Demographic and clinical characteristics (eg, age, sex, and Charlson comorbidities) were reported, along with hospitalization information including length of stay, number

https://doi.org/10.1017/ice.2019.184 Published online by Cambridge University Press

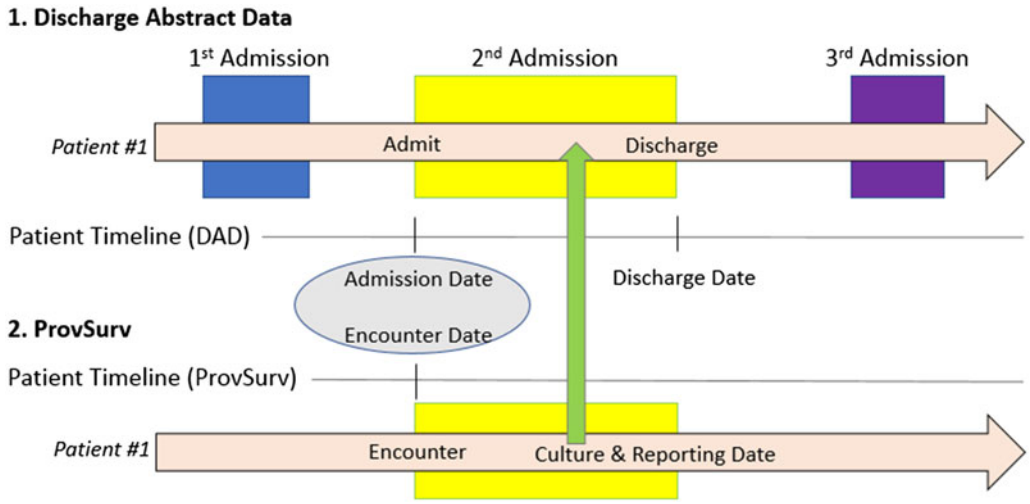

Fig. 2. Diagram illustrating the data linkage and validation process. Personal health number (PHN and gender) ensured that these DAD and ProvSurv records belong to patient 1. Patient 1 has multiple DAD records and a ProvSurv record. Of 3 DAD historical in-hospital admissions records illustrated here, the second admission is the record that needs to be linked to the ProvSurv record. If these records are from the same admission period, then culture date and reporting date from ProvSurv essentially should lie between the admission date and discharge date from DAD (central arrow). The admission date (DAD) and encounter date (ProvSurv) were compared for validation. This process ensured that correct DAD hospitalization record was linked to ProvSurv record. We used SAS version 9.4 software for this study, but other software such as Python programming language (Python Software Foundation, https://www.python.org/) and R (R Core Team, Vienna, Austria) could also be used.

of hospitalizations in the prior year, discharge disposition, and the most responsible diagnoses.

The Charlson comorbidity score and index were calculated using a well-established and validated coding algorithm.[18] Sample collection time, laboratory positive culture report time and testing rationales behind incident case samples (abstracted from medical chart review) are part of the obtained surveillance records and were summarized as counts. SAS version 9.4 software (SAS Institute, Cary, NC) was used for statistical analyses. A P value cutoff of 0.05 was used to compare infection and colonization categories.

The University of Calgary's Conjoint Health Research Ethics Board (CHREB) approved this study.

Results

We identified 3 shared variables (PHN, gender, and admission date) between these 2 databases. The linkage methodology (Fig. 1) revealed records with inaccurately coded gender (n = 75) or PHN (n = 88). Therefore, the approach that used both PHN and gender, which resulted in a smaller number of linked records, was used to ensure data quality. Of 2,616 records, 2,453 (93.8%) were successfully linked to the DAD. Also, 23 pediatric patients were excluded from the study, bringing the total number of patients in this study to 2,430. The admission date and encounter date matched for all cases (100%) between linked cases. In total, 1,226 (50.4%) patients were identified with HA-MRSA, and 1,204 (49.5%) patients were identified with CA-MRSA in the 4 adult acute-care facilities in the linked records. Figures 1 and 2 depict the overall study design and linkage steps undertaken, respectively.

Descriptive statistics for HA-MRSA and CA-MRSA categories are shown in Tables 3 and 4. The initial analysis demonstrated those HA-MRSA and CA-MRSA characteristics followed that previously reported in the literature (Table 3). The HA-MRSA population was more likely to be older ($P < .001$), had a higher Charlson comorbidity score (mean, 1.7 (±1.9 SD) compared to 1.0 (±1.6) in the CA-MRSA group) and index, and had more comorbidities. The HA-MRSA group was also more likely to have had prior hospitalizations compared to the CA-MRSA group (Table 4).

The stratification by clinical presentation (infection vs colonization) in HA-MRSA and CA-MRSA did not impact these observations (Supplementary Tables 1 and 2 online). In HA and CA-MRSA categories, colonization accounted for 73.5% (901 of 1,226) of the identified HA-MRSA cases and 78.1% (940 of 1,204) of the CA-MRSA cases. Of the 325 HA-MRSA cases with first time infection, 54 patients (16.6%) died on first hospital admission compared to 93 of 901 patients (10.3%) with HA-MRSA colonization ($P < .05$). The mortality rates in CA-MRSA infection and colonization were 8.0% (21 of 264) and 7.2% (68 of 940), respectively ($P = .66$).

Supplementary Table 3 provides the sampling rationale, stratified by body sites for 2,430 patients. Clinical isolates were identified as the most common means for identifying HA (315 of 325) and CA-MRSA (249 of 264) infections, respectively. This categorization was not as evident in HA (108 of 901) and CA colonization (61 of 940) categories.

We were unable to develop a targeted MRSA screening algorithm using this linked data because administrative data did not provide sufficient clinical granularity.

Discussion

To our knowledge, this is the first Canadian study linking a frontline HAI AMR surveillance database to a population-based administrative database to analyze HA-MRSA and CA-MRSA cases. The World Health Organization's Global Action Plan on Antimicrobial Resistance noted that there is an important knowledge gap on incidence, prevalence, and geographical patterns that needs to be identified and made accessible in a timely manner to inform local, national, and regional actions.[19] Our findings augment the current IPC surveillance practice efficiency by making relevant clinical information available in a timely manner via the linkage with the DAD data. Furthermore, this work offers a glimpse into the intricate relationship between patient health and the healthcare system through additional information otherwise not readily

https://doi.org/10.1017/ice.2019.184 Published online by Cambridge University Press

Table 3. Hospital-Acquired Methicillin-Resistant *Staphylococcus aureus* (HA-MRSA) and Community-Acquired MRSA (CA-MRSA) Patient Characteristics Identified in 4 Calgary Acute-Care Facilities Between April 2011 and March 2017

Characteristic	HA-MRSA Overall (N = 1,226), No. (%)	CA-MRSA Overall (N = 1,204), No. (%)	*P* Value
Age, mean y (SD)	70.3 (17.2)	60.9 (20.8)	<.001
Age category			
19–45 y	119 (9.7)	314 (26.1)	<.001
46–64 y	272 (22.2)	339 (28.2)	<.001
65+y	835 (68.1)	551 (45.8)	<.001
Gender, % male	644 (52.5)	692 (57.5)	.014
Most frequent comorbidities (Charlson), No. (%)			
Congestive heart failure	161 (13.1)	78 (6.5)	<.001
Dementia	158 (12.9)	73 (6.1)	<.001
Chronic pulmonary disease	148 (12.1)	88 (7.3)	<.001
Diabetes, complication	239 (19.5)	137 (11.4)	<.001
Charlson score, mean (SD)	1.7 (1.9)	1.0 (1.6)	<.001
Comorbidities index, No. (%)			
0	426 (34.8)	683 (56.7)	<.001
1	235 (19.2)	228 (18.9)	.851
2	253 (20.6)	153 (12.7)	<.001
3	154 (12.6)	67 (5.6)	<.001
4+	158 (12.8)	73 (6.1)	<.001

Note. SD, standard deviation.

Table 4. Hospital-Acquired Methicillin-Resistant *Staphylococcus aureus* (HA-MRSA) and Community-Acquired MRSA (CA-MRSA) Hospital Outcome Characteristics Identified in 4 Calgary Acute-Care Facilities Between April 2011 and March 2017

Characteristic	HA-MRSA Overall (N = 1,226)	CA-MRSA Overall (N = 1,204)	*P* Value
Median LOS (IQR)	33.5 (62)	7 (12)	
Range	4–422	1–187	
No. of Hospitalizations in previous 1 year, no. (%)			
0	191 (15.6)	343 (28.5)	<.001
1	409 (33.4)	650 (54.0)	<.001
2	305 (24.5)	117 (9.7)	<.001
3	153 (12.5)	54 (4.5)	<.001
4+	168 (13.7)	40 (2.1)	<.001
Disposition of index, no. (%)			
Discharge/ Transfer	1,079 (88.0)	1115 (92.6)	<.001
Death	147 (12.0)	89 (7.4)	<.001
Top 5 most responsible diagnoses, no. (%)			
1	Congestive heart failure 60 (4.9)	*S. aureus* sepsis 37 (3.1)	
2	Rehabilitation care 54 (4.4)	Cellulitis of lower limb 30 (2.5)	
3	COPD with LRTI 47 (3.8)	Congestive heart failure 29 (2.4)	
4	Urinary tract infection 35 (2.9)	COPD with acute exacerbation 29 (2.4)	
5	Delirium on dementia 32 (2.6)	Sepsis, unspecified 28 (2.3)	

Note. COPD, chronic obstructive pulmonary disease; LRTI, lower respiratory tract infection.

available for standard IPC surveillance using CA- and HA-MRSA as an example.

A multidisciplinary team with training and skills in (1) infection prevention and control, (2) surveillance practices, and (3) infectious diseases and microbiology, and (4) programming and data science should be assembled for conducting this type of database linkage. In the age of "big data," data linkage is becoming a necessary process to answer clinical and research questions. This team should have a thorough understanding of contained database elements and how these databases reflect the context of the research question being asked. For example, we understood that the infection control surveillance database contained HAI hospitalization encounters, but we recognized that this database lacked the detailed granular health information for each patient. In DAD, we recognized that this database contained coded hospitalization information but lacked surveillance details.

The 2 types of linkage processes are deterministic and probabilistic linkage,[20] and these processes have different techniques and assumptions.[20] We chose to proceed with deterministic linkage because we had 3 shared variables: PHN and gender, with admission or encounter date for validation. Retrospectively, we used these linkage process to assess the error rate of the input surveillance data. Among the 163 patients (6.2%) of the records that did not link properly, 75 (2.9%) were due to nonmatching PHN and 88 (3.3%) were due to conflicting gender records between the databases. We determined that these latter findings were errors associated with the locally collected surveillance records and foreign visitor status outside Alberta.

Foreign visitors would not be covered under Alberta's insurance plan and would not have had a PHN recorded, which explains why these cases were not linked. We felt this error rate was acceptable taking into account daily clinical practice activities and the relatively low numbers of cases that did not match. A fulsome review of linkage quality is beyond the scope of this paper; we wanted to demonstrate the utility of the linkage with administrative data. A future study should assess various linkage methodologies (ie, comparing performances of deterministic vs probabilistic performance), and we expect this to be a substantial undertaking.

We were successful in linking these 2 databases, but we faced several challenges. DAD, through PHN, contained all historical hospitalizations pertinent to a single individual. Ensuring that each surveillance record matched the correct hospitalization episode required recoding of data and working with time-related information contained in these databases. During data cleaning, we confirmed that the same variable (eg, admission date) had been

https://doi.org/10.1017/ice.2019.184 Published online by Cambridge University Press

coded in different formats between these databases. Therefore, data cleaning became a crucial step in our methodology. One of our priorities was ensuring that combined dataset made clinical and temporal sense. The end product was actionable enriched incident MRSA surveillance data that could be explored further.

Data linkage to the untapped rich information stored in administrative databases such as the DAD, which contains information on outcomes and comorbidities, may serve as a methodology to assess patient safety and performance indicators to complement traditional infection prevention and control surveillance practices.[21] An advantage of utilizing a single, province-wide surveillance system is that overcounting of incident cases is avoided. Linkage work, illustrated in our study, may provide inferences on patient and facility characteristics, enhancing the details of AMR surveillance data and contributing to the understanding of local epidemiology. It may also lead to the development of prevention strategies for patients and facilities with elevated rates.

In a follow-up question, we explored the feasibility of developing a targeted screening algorithm using these linked data. We soon realized that this targeted algorithm could not be answered with the data we obtained. Combining administrative data greatly enhanced the value of surveillance data, but it still did not contain sufficient clinical granularity for addressing the latter question. Improving details of administrative data may require applying natural language processing on free-text data held in electronic medical records. Other pertinent information such as socioeconomic status (eg, homelessness) and clinical practice information (eg, hand hygiene) were also missing. Developing a targeted screening strategy is analogous to creating a risk adjustment algorithm that accounts for patient comorbidities,[22] MRSA exposure history, social risk factors[23] (eg, injectable drug-use, homelessness), and patient–staff interaction.

This algorithm, based on input variables or features, was designed to determine whether an individual needed to be screened. It required a substantial linkage of existing databases ranging from surveillance, primary care, administrative, and others. The availability of big data is transforming healthcare practices across the world. Data linkage to other data types is the basis of taking the first step toward achieving a precision medicine approach to enhance existing infection control surveillance. Developing a targeted screening algorithm is an example of precision medicine achieved through health data science in the context of MRSA.

The results of this study should be interpreted in light of several limitations. We only linked 2 population-based databases that focused on admitted patients, and we lacked information on those who were not admitted. Additionally, CA-MRSA cases that were identified in the community and resolved without hospitalization would not be captured in either database. Furthermore, exposures associated with HCA-MRSA may occur at outpatient settings, and linked data may be incomplete for this category. Our findings may not be generalizable outside our region, and results may differ in other parts of the country.

In summary, the findings from our linkage study represent an important step toward complementing traditional surveillance data practices on AMR. We used MRSA as an example pathogen in this study, but the linkage methodology we employed could be used on any AMR pathogen. Importantly, AMR surveillance strategies are diverse and highly variable. Health systems may not be using the same case definitions and their referencing laboratory standards may not be similar. The data entered into surveillance systems reflect these divergent data collection strategies. Identifying robust data collection strategies, assessing data quality, developing appropriate data linkage process, and assessing data linkage quality are all critical topics that require substantial research efforts. However, without first obtaining granular details in a timely manner, it is difficult to conduct root cause analysis to identify issues.

Data linkages to other health data have the potential to improve the variability and the lack of standardization. These data linkages may represent a new paradigm in improving and augmenting the HAI surveillance data practices. We urge other jurisdictions and health systems to consider adopting data linkage into their surveillance practices. Although this pursuit may require a skilled work force capable of working with administrative data, it has the potential to generate efficient and systematic evidence for the development of focused intervention strategies targeting the growing issue of AMR.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/ice.2019.184

Acknowledgments. The authors thank the Alberta Health Services Infection Prevention and Control Surveillance for providing data for this research.

Financial support. No financial support was provided relevant to this article.

Conflicts of interest. All authors report no conflicts of interest relevant to this article.

References

1. Wilson APR, Livermore DM, Otter JA, *et al.* Prevention and control of multi-drug-resistant gram-negative bacteria: recommendations from a Joint Working Party. *J Hosp Infect* 2016;92:S1–S44.
2. Vasoo S, Barreto JN, Tosh PK. Emerging issues in gram-negative bacterial resistance: an update for the practicing clinician. *Mayo Clin Proc* 2015;90:395–403.
3. Tornimbene B, Eremin S, Escher M, Griskeviciene J, Manglani S, Pessoa-Silva CL. WHO global antimicrobial resistance surveillance system early implementation. *Lancet Infect Dis* 2018;18:241–242.
4. WHO publishes list of bacteria for which new antibiotics are urgently needed. World Health Organization website. https://www.who.int/news-room/detail/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed. Published 2017. Accessed March 16, 2018.
5. Antibiotic/antimicrobial resistance (AR/AMR)—biggest threats and data. Centers for Disease Control and Prevention website. https://www.cdc.gov/drugresistance/biggest_threats.html. Published 2018. Accessed April 14, 2019.
6. Amaratunga K, Tarasuk J, Tsegaye L, Archibald CP. Advancing surveillance of antimicrobial resistance: summary of the 2015 CIDSC report. *Can Commun Dis Rep* 2016;42:232–237.
7. Drees M, Gerber JS, Morgan DJ, Lee GM. Research methods in healthcare epidemiology and antimicrobial stewardship: use of administrative and surveillance databases. *Infect Control Hosp Epidemiol* 2016;37:1278–1287.
8. Census profile, 2016 census: Alberta profile. Statistics Canada website. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/Page.cfm?Lang=E&Geo1=PR&Code1=48&Geo2=&Code2=&Data=Count&SearchText=Alberta&SearchType=Begins&SearchPR=01&B1=All&GeoLevel=PR&GeoCode=48. Published 2017. Accessed March 16, 2018.
9. Provincial wide surveillance. Alberta Health Services website. https://www.albertahealthservices.ca/assets/healthinfo/ipc/hi-ipc-provincial-surveillance.pdf. Published 2018. Accessed March 17, 2018.
10. Methicillin-resistant *Staphylococcus aureus* (MRSA) culture. Calgary Laboratory Services website. http://www.calgarylabservices.com/lab-services-aguide/microbiology/Test/Tests/MethicillinResistantStaphylococcusAureus-MRSA.aspx. Published 2018. Accessed October 22, 2018.

11. Discharge abstract database (DAD) metadata. Canadian Institute for Health Information website. https://www.cihi.ca/en/discharge-abstract-database-metadata. Published 2018. Accessed April 5, 2018.
12. Alberta health care insurance plan (AHCIP). Government of Alberta website. https://www.alberta.ca/ahcip.aspx. Published 2019. Accessed April 6, 2019.
13. Naimi TS, LeDell KH, Como-Sabetti K, *et al.* Comparison of community and health care–associated methicillin-resistant *Staphylococcus aureus* infection. *JAMA* 2003;290:2976–2984.
14. Morrison MA, Hageman JC, Klevens RM. Case definition for community-associated methicillin-resistant Staphylococcus aureus. *J Hosp Infect* 2006;62:241.
15. MRSA: information for clinicians. Centers for Disease Control and Prevention website. https://www.cdc.gov/mrsa/community/clinicians/index.html. Published 2016. Accessed April 5, 2018.
16. Methicillin-resistant *Staphylococcus aureus* (MRSA) provincial surveillance protocol. Alberta Health Services website. https://www.albertahealthservices.ca/assets/healthinfo/ipc/hi-ipc-sr-mrsa-surveillance-protocol.pdf. Published 2018. Accessed April 5, 2018.
17. Andreassen AES, Jacobsen CM, de Blasio BF, White R, Kristiansen IS, Elstrøm P. The impact of methicillin-resistant S. *aureus* on length of stay, readmissions and costs: a register based case-control study of patients hospitalized in Norway. *Antimicrob Resist Infect Control* 2017;6:74.
18. Quan H, Sundararajan V, Halfon P, *et al.* Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130–1139.
19. Tackling antimicrobial resistance and antimicrobial use: a pan-Canadian framework for action. Government of Canada website. https://www.canada.ca/en/health-canada/services/publications/drugs-health-products/tackling-antimicrobial-resistance-use-pan-canadian-framework-action.html#a3.1. Published 2017. Accessed April 6, 2018.
20. Dusetzina SB, Tyree S, Meyer AM, *et al.* Linking data for health services research: a framework and instructional guide, an overview of record linkage methods. Agency for Healthcare Research and Quality (US) website. https://www.ncbi.nlm.nih.gov/books/NBK253312/. Published 2014. Accessed April 6, 2019.
21. Virnig BA, McBean M. Administrative data for public health surveillance and planning. *Annu Rev Public Health* 2001;22:213–230.
22. Lukovac E, Koluder-Cimic N, Hadzovic-Cengic M, Baljic R, Hadzic A, Gojak R. Analysis of comorbidity of the patients affected by Staphylococcal bacteremia/sepsis in the last ten years. *Materia Socio Medica* 2012;24:13.
23. Campbell KM, Vaughn AF, Russell KL, *et al.* Risk factors for community-associated methicillin-resistant *Staphylococcus aureus* infections in an outbreak of disease among military trainees in San Diego, California; 2002. *J Clin Microbiol* 2004;42:4050–4053.

https://doi.org/10.1017/ice.2019.184 Published online by Cambridge University Press