

# Administrative health databases for addressing emerging issues in adults with CHD: a systematic review

## Original Article

**Cite this article:** Cohen S, Gilutz H, Marelli AJ, Iserin L, Benis A, Bonnet D, Burgun A. (2018) Administrative health databases for addressing emerging issues in adults with CHD: a systematic review. *Cardiology in the Young* 28: 844–853. doi: 10.1017/S1047951118000446

Received: 23 November 2017  
Revised: 19 February 2018  
Accepted: 20 February 2018

### Key words:

Adult with CHD; administrative health database; epidemiology; outcomes

### Author for correspondence:

S. Cohen, INSERM, U1138, Centre de Recherche des Cordeliers, Equipe 22, 15 rue de l'École de Médecine, Paris 75006, France. Tel: +33 1 44 27 63 93; Fax: +33 1 44 27 64 21 E-mail: sarah.cohen.hegp@gmail.com

Sarah Cohen<sup>1,2,3</sup>, Harel Gilutz<sup>4</sup>, Ariane J. Marelli<sup>2</sup>, Laurence Iserin<sup>3</sup>, Arriel Benis<sup>5</sup>, Damien Bonnet<sup>6,7</sup> and Anita Burgun<sup>1,8</sup>

<sup>1</sup>INSERM-UMRS 1138 Team 22, Cordeliers Research Center, Paris Descartes University, Paris, France, <sup>2</sup>McGill Adult Unit for Congenital Heart Disease Excellence, Montreal, Quebec, Canada, <sup>3</sup>Adult Congenital Heart Disease Unit, Cardiology Department, Reference Center for Complex Congenital Heart Diseases (M3C), European Georges Pompidou Hospital, Assistance Publique des Hôpitaux de Paris (AP-HP), Paris, France, <sup>4</sup>School of Medicine, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel, <sup>5</sup>Faculty of Technology Management, Holon Institute of Technology (HIT), Holon, Israel, <sup>6</sup>Department of Paediatric Cardiology, Centre de référence des Malformations Cardiaques Congénitales Complexes (M3C), Hôpital Necker-Enfants Malades, AP-HP, Paris, France, <sup>7</sup>Paris Descartes University Sorbonne Paris Cité, Paris, France and <sup>8</sup>Department of Medical Informatics and Public Health, European Georges Pompidou Hospital, AP-HP, Paris, France

### Abstract

The need for population-based studies of adults with CHD has motivated the growing use of secondary analyses of administrative health data in a variety of jurisdictions worldwide. We aimed at systematically reviewing all studies using administrative health data sources for adult CHD research from 2006 to 2016. Using PubMed and Embase (1 January, 2006 to 1 January, 2016), we identified 2217 abstracts, from which 59 studies were included in this review. These comprised 12 different data sources from six countries. Of these, 55% originated in the United States of America, 28% in Canada, and 17% in Europe and Asia. No study was published before 2007, after which the number of publications grew exponentially. In all, 41% of the studies were cross-sectional and 25% were retrospective cohort studies with a wide variation in the availability of patient-level compared with hospitalisation-level episodes of care; 58% of studies from eight different data sources linked administrative data at a patient level; and 37% of studies reported validation procedures. Assessing resource utilisation and temporal trends of relevant epidemiological and outcome end points were the most reported objectives. The median impact factor of publication journals was 4.04, with an interquartile range of 3.15, 7.44. Although not designed for research purposes, administrative health databases have become powerful data sources for studying adult CHD populations because of their large sample sizes, comprehensive records, and long observation periods, providing a useful tool to further develop quality of care improvement programmes. Data linkage with electronic records will become important in obtaining more granular life-long adult CHD data. The health services nature of the data optimises the impact on policy and public health.

As a result of improved survival of CHD patients,<sup>1</sup> the population of adults with CHD is growing and ageing, with an estimated prevalence of three per 1000 adults.<sup>2</sup> Adults accounted for two-thirds of the entire CHD population in 2010 in Quebec, Canada.<sup>3</sup> Residual anomalies and late-onset complications have increased the need for hospitalisations and health care utilisation.<sup>4,5</sup> As CHD patients age, they develop an increased risk of life-long cardiac<sup>6</sup> and cerebrovascular events. Moreover, care gaps, common in adult CHD, have increased the risk of cardiovascular events.<sup>7–9</sup>

As a byproduct of a patient care, vast quantities of information are collected and stored in administrative health databases for the purposes of registration, billing, or record-keeping. The reuse of patient data for research has gained considerable importance in non-adult CHD, as well as adult CHD, populations. Follow-up of CHD populations can be traced for decades using health administrative databases. Multiple years of data permit studying change over time for numerous variables. In addition, such data usually include vital statistics, physician visits, hospital discharge abstracts, pharmaceutical prescriptions, and claims data routinely field by physicians. Administrative health data sources have thus emerged as an important source of population-based analyses for adult CHD patients in order to guide public health policy and resource allocation in industrialised countries.

Data reuse, also called secondary use of data, refers to studies whose purpose is not directly related to the initial reason for collecting data or to the care of the individual patient who is the subject of the health information. Such comprehensive and broad data sources, although not

initially developed for the study of disease distribution or disease trends, offer the opportunity for population-level analyses. Data on population health differ between countries in terms of availability, size, and content. Denmark, for example, has gathered a wide range of data variables on all its citizens, including very comprehensive data on a patient trajectory within their health system.<sup>10</sup> Such data sources are typically collected on national or state-wide levels. These data sources have become increasingly used in adult CHD populations worldwide, in developed countries, and where at least some portion of the population benefits from government-funded health insurance.

Against this backdrop, we carried out a systematic literature review to identify all the studies based on secondary use of administrative health data sources that provided new knowledge on adult CHD. Our purpose was to review the outcomes covered, the data source characteristics, and the strengths and limitations of administrative data sources used to address knowledge gaps in the adult CHD populations.

## Materials and methods

### Systematic search

This systematic review focuses on studies reusing administrative health data repositories. The review is in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).<sup>11</sup>

### Search strategy

We conducted a comprehensive search of PubMed and Embase for relevant peer-reviewed publications from January 1, 2006, to January 1, 2016. The search strategy was developed by H.G. and S.C. with the help of a reference librarian. A comprehensive list of MeSH terms and keywords was used to query Medline and Embase (Supplementary Table 1). The search strategy also included screening of reference lists of relevant publications – the “snowball” search technique.

### Eligibility criteria

Publications were selected if they met the following criteria: the study relied on administrative health data of any kind – e.g., expenditures, hospital discharges, claims, national survey, and death certificates – in the methodology, either for initiation of the research or for follow-up, regardless of outcomes – electronic medical record was not considered as “administrative health data”; the study population comprised adult CHD patients ( $\geq 18$  years) or included both adults and children but distinguished the two groups in the results and presented specific comments referring to adult CHD; and the study was published in a peer-reviewed journal in English.

CHD was defined according to the criteria of Mitchell et al<sup>12</sup> already used for several reviews<sup>2</sup> – that is, “a gross structural abnormality of the heart or intrathoracic great vessels that is actually or potentially of functional significance”. Thus, we excluded publications dealing with non-structural lesions – such as cardiomyopathies and congenital arrhythmias – ductus arteriosus in premature infants, mitral valve prolapse, or isolated bicuspid aortic valve. We included patients with Marfan syndrome when they presented a complication that was “functionally significant” or required an invasive intervention as they have commonalities: the relatively low prevalence, the absence of curative treatment,

the need for cardiac surgery, and the importance of life-long follow-up specialised care.

### Study selection

Supplementary Figure 1 represents the study selection process. The PubMed and Embase searches yielded 2217 publications. After exclusion of 29 duplicates, the titles and abstracts of 2188 records were screened and assessed according to the following exclusion criteria: the adult population was not specifically studied or mentioned; the data source used non-exhaustive or comprised a registry of patients volunteering to be included. As such, studies from the CONCOR registry or from tertiary centres with high-volume care of CHD, such as the databases from the CHD Program at the University Hospitals Leuven, Belgium, or from the Royal Brompton Hospital in London, United Kingdom, were excluded based on these criteria. Two of the authors, H.G. and S.C., independently read the first 50 abstracts to harmonise the search. Disagreements were resolved by consensus meetings. In case a database was used in several studies, all the corresponding articles were considered for review. Finally, 197 full-text publications were independently selected for eligibility assessment by both authors – H.G. and S.C. At this stage, 10 publications were added to the 197 after searching the reference lists of relevant publications. After a detailed review of the full text of these 207 eligible publications, 59 were finally included in this systematic review (Supplementary Fig 1).

### Data extraction

Data were extracted using a standardised collection form. Attention was given to the study characteristics: year of publication, data source and coverage, follow-up duration, definition of CHD diagnoses, study design, population included and its characteristics, exclusion criteria, objectives of the study (classified into categories), potential bias, and journal impact factor obtained from the Journal Citation Reports (Thomson Reuters, New York, New York, United States of America). We then grouped the selected articles according to the database used and briefly described each one: name, coverage, sponsoring organisation, data sources, and available data.

## Results

### Description of administrative health data sources used

To date, 59 studies relied on secondary use of administrative health databases to describe specific issues associated with adult CHD patients. Most of them originated in the United States of America ( $n = 32$ ; 55%) and Canada ( $n = 17$ ; 28%). Only four (7%) were from Europe and six (10%) from Asia (Supplementary Fig 2). In some countries, publications were derived from several administrative databases. Canadian publications originated mostly from the province-wide Quebec CHD database, in 15 out of 17 Canadian studies. American studies were derived from one federal database, the California Office of Statewide Health Planning, and from five national bases: the Nationwide Inpatient Sample, the multiple cause of death (MCOB) public-use data file, the Pediatric Health Information System, the United Network for Organ Sharing, and the University Health System Consortium Clinical Database/Resource Manager.

The 59 studies were derived from 12 different data sources from six countries. Only two of them were CHD-specific – the

Quebec CHD Database and the Danish Register of CHD – and were created by merging several national or provincial data sources before extracting CHD patients' data. Table 1 shows the characteristics of each source. A unique personal identifier provided for every inhabitant is described in Denmark, in Canada, and in the United States of America with the United Network for Organ Sharing. In the Database from the Canadian Institute for Health Information, excluding Quebec, multiple hospitalisations for the same patient can be tracked with the use of this unique patient identifier. Further, in Quebec and Denmark, using this encrypted number, data from different administrative sources, for example, physicians' claims database, hospital discharge summary database, death registry, and prescription registry, may be cross-linked for each patient to provide a global and longitudinal overview of a patient's history. Thus, such databases contain patient-level data, whereas the Nationwide Inpatient Sample, for example, contains hospitalisation-level data. Among the 59 publications included, 58% (n = 34) studies derived from eight data sources linked administrative data at a patient level, whereas 42% (n = 25) studies derived from two data sources contained hospitalisation-level information (Table 1). Supplementary Results give further details on data sources from Quebec, the United States of America, Taiwan, and Denmark, which were particularly productive in reusing administrative databases in the field of adult CHD in recent years.

Validation procedures and/or specific algorithms that merge several administrative data sources were used to increase the internal validity of CHD diagnoses in 37% (n = 22) of the included studies from three data sources: Denmark, Quebec, and Taiwan.

### Study characteristics related to methodology

Among the 59 publications included in this review, 24 used a cross-sectional design,<sup>4,13–35</sup> 15 used a cohort study design,<sup>36–50</sup> nine were non-analytical descriptive,<sup>3,5,51–57</sup> and one used a time-series analysis.<sup>58</sup> In six studies, the authors constructed a matched control cohort<sup>6,59–63</sup> using a non-CHD control population when available.<sup>6,59–62</sup> Four studies combined cohort study and nested case-control study to assess the cumulative incidence or prevalence of an outcome, such as atrial arrhythmia, coronary artery disease, and stroke, as well as its predictors (Supplementary Table 2).<sup>8,44,64,65</sup> Because of the inherent nature of the data sources, all studies were retrospective. Study characteristics are summarised in Supplementary Table 2.

### Study objectives and end points

Although all of these publications dealt with the adult CHD field, the questions they answered were not the same, as shown in Supplementary Table 3. Overall, five (8.5%) studies addressed demographics of the adult CHD population, such as prevalence, incidence, and sex ratio,<sup>3,34,50,51,53,66</sup> and 19 (32.2%) assessed specific issues in the adult CHD population. These were cardiovascular – atrial arrhythmia,<sup>37,48,64</sup> pulmonary hypertension,<sup>63</sup> coronary artery disease,<sup>65</sup> heart failure,<sup>25</sup> and stroke<sup>44,47,60</sup> – or non-cardiovascular – maternal issues,<sup>20,34,43</sup> pneumonia,<sup>61</sup> non-alcoholic cirrhosis,<sup>24</sup> cancer,<sup>45</sup> dementia, gastrointestinal bleed, and chronic kidney disease.<sup>66</sup> As the data were routinely derived from health insurance claims, resource utilisation and cost/cost-effectiveness were reported in, respectively, 26 and 12 articles. Resource utilisation included lengths of stay in hospitalisation

and/or ICU, total hospital charges, emergency department visits, outpatient visits (to general practitioner or cardiologist), number of procedures, and number of admissions. In 16 articles (24%), the authors focused on patterns of care, comparing specialised versus non-specialised centres,<sup>8,13,29,57</sup> teaching versus non-teaching hospitals,<sup>14,54</sup> adult versus paediatric hospitals,<sup>27</sup> or even analysing the impact of hospital volume<sup>14,16,28,30,32,46,54,56</sup> on outcomes. Finally, these databases covered long periods of follow-up, which made it possible to assess temporal trends in half of the articles (n = 30, Supplementary Table 3).

### Publication trends and journal impact factor

No qualifying study was published before 2007. From 2006 to 2015, the number of articles grew exponentially, going from three articles in 2007 to 16 from four different countries in 2015 (Fig 1).

Journal impact factors for all 59 articles ranged from 0.825 to 17.047, with a median of 4.04 and an interquartile range of 3.15, 7.44. The impact factors according to the source of publications are shown in Table 2. Denmark, Quebec, and the United States of America – non-Nationwide Inpatient Sample – published articles in very-high-impact-factor (impact factors >13.5) journal data sources (Fig 2a). Figure 2b represents the cumulative number of articles and their impact factor according to the year of publications. In 2015, 11 articles (18.6%) of the 59 were published in very-high-impact-factor journals.

### Discussion

Although generally developed for registration and billing purposes, administrative health data are rich sources of information and were extensively used over the past decade in adult CHD research in a growing number of countries. This systematic review highlights the differences between data sources used with a wide variation in the availability of patient-level compared with hospitalisation-level episodes of care or in the availability of internal validation of diagnoses. In addition, the characteristics of the studies are influenced by the structure of the data sources from which they were derived. Thus, such data sources were particularly used for assessing resource utilisation and temporal trends of outcome end points over long observation periods.

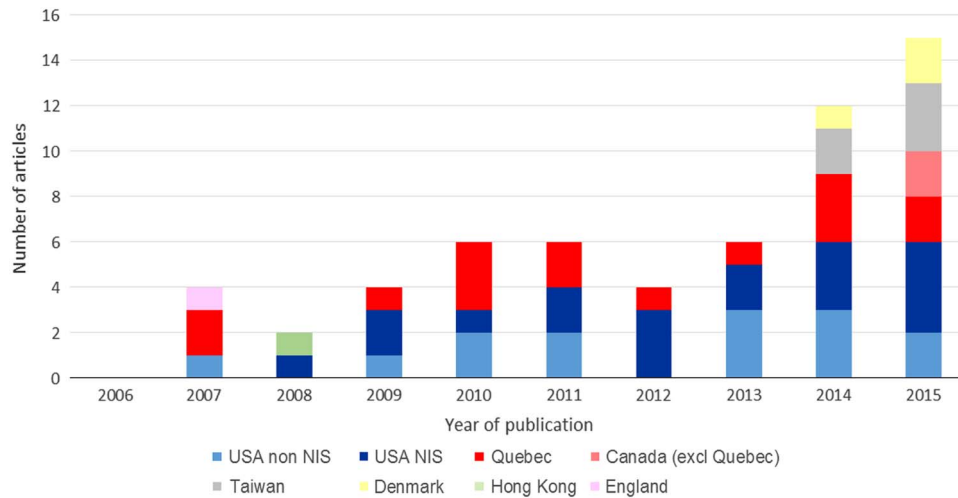
In this review, we excluded all registries assembled as primary data sources even if they covered large adult CHD populations. For example, since November 2001 in the Netherlands, adult CHD patients have been recruited in tertiary and secondary medical centres and included in the nationwide CONCOR registry.<sup>67</sup> Thus, this registry makes it possible to study only a survival cohort of patients and comprises neither very mild CHD in patients lost to medical surveillance nor critical CHD that led to death before enrolment. Moreover, the CONCOR registry has an inclusion bias because of its recruitment in tertiary and secondary medical centres. The same limitations exist in the CHD database of Leuven, a Belgian tertiary medical centre,<sup>1</sup> or in the Royal Brompton Hospital in London.<sup>68</sup> On the other hand, by their nature, these sources provide highly detailed data for clinical assessment. Here, we focused on studies relying on administrative health databases to provide new knowledge on adult CHD. In areas in which access to health care is universal, as in Taiwan, Quebec, or Denmark, the entire population is covered by the data collection, and non-participation is minimal. This is a paramount difference compared with registries or clinical databases in which participation is voluntary and therefore biased. Owing to their

**Table 1.** Description of administrative health data sources used in the included articles.

Name	Geographic area covered	Sponsoring organisation	Brief description (data sources, types of data)	Unit of analysis	Available data			
					General inpatient data (demographics, diagnoses, procedures, dates of stay)	Outpatient data	Out-hospital death	In-hospital death
Quebec CHD Database <sup>3,4,8,35,37,39-41,44,51,58,63-66</sup>	Province of Quebec, Canada	McGill Adult Unit Congenital Heart Disease Excellence	Database created by merging information on all patients with a CHD diagnosis from the three province-wide administrative databases: the physicians' services and drug claims database, the hospital discharge summary database; and the Quebec Health Insurance Board. It does not contain non-CHD population	Patient	✓ and physicians' characteristics hospital records hospital characteristics	✓	✓	✓
California OSHPD and OSHPD-Ambulatory Surgery Data <sup>13,29,30,56,57</sup>	California, USA	State of California (Health care Cost and Utilization Project partner, sponsored by the Agency for Health care Research and Quality) <a href="http://www.oshpd.ca.gov/">http://www.oshpd.ca.gov/</a>	Data set containing deidentified information on all inpatient discharges from non-federal acute care hospitals in California, principal diagnosis, and up to 24 secondary diagnoses. The OSHPD-Ambulatory Surgery Data file is a comprehensive, public data set of outpatient surgery encounters consisting of one record for each time a patient is treated in a licensed ambulatory surgery centre in California	Hospitalisation (no linkage between hospitalisations in a same patient)	✓ and race, ethnicity payer status: public insurance, private insurance, and other patient's ZIP code admission source	No	NA	✓
Hospital Episode Statistics database <sup>52</sup>	England (NHS hospitals only)	National Health Services <a href="http://content.digital.nhs.uk/hes">http://content.digital.nhs.uk/hes</a>	Database holding data for England on the care provided by NHS hospitals (including to patients treated privately) and for NHS hospital patients treated elsewhere. It contains details of all admissions, outpatient appointments, and emergency attendances. Private hospitals are not covered	Hospitalisation	✓ and surgical procedures	✓ Only NHS hospitals	No	✓
Hong Kong Health Authority's Clinical Data Analysis and Reporting System (CDARS) <sup>53</sup>	Hong Kong (Public hospital only)	Hong Kong Health Authority	Comprehensive, prospectively entered, centralised, computerised database for patients admitted to all public Hong Kong hospitals	Patient	✓ and surgical procedures prescription drugs	N/A	N/A	✓
Nationwide Inpatient Sample <sup>5,14,15,18-21,24,25,27,28,32-34,43,46,54,62</sup>	USA	Managed under the Health care Cost and Utilization Project. Sponsored by the Agency for Health care Research and Quality (AHRQ) <a href="https://www.hcup-us.ahrq.gov/nisoverview.jsp">https://www.hcup-us.ahrq.gov/nisoverview.jsp</a>	The largest publicly available all-payer inpatient care database in the United States of America. Stratified sample designed to approximate a 20% sample of discharges from US community (non-federal, short-term, general, and specialty) hospitals. Sampling weights help provide national estimates	Hospitalisation (no linkage between the hospitalisations in a same patient)	✓ and race, ethnicity patient's ZIP code and median household income fetal outcomes hospital characteristics physician identification payer status: public insurance, private insurance, and other admission characteristics discharge status (transfer, discharged to home, left against medical advice) cost	No	No	✓
United Network for Organ Sharing <sup>23,36,38</sup>	USA	U.S. Department of Health and Human Services <a href="https://optn.transplant.hrsa.gov/data/about-data/">https://optn.transplant.hrsa.gov/data/about-data/</a>	It contains all national data on the candidate waiting list, organ donation, matching, and transplantation, through electronic data collection forms	Patient	Information at the time of listing Information from the initial transplant admission Information at each visit after transplant: vital status, cause of death, graft status, employment status, clinical Information -information at the time of organ donation: donor demographics, co-morbidities, infectious disease status, cause of death	N/A	✓	✓
Pediatric Health Information System <sup>16,17,22,26</sup>	USA					No	No	✓

Table 1. (Continued)

Name	Geographic area covered	Sponsoring organisation	Brief description (data sources, types of data)	Unit of analysis	Available data			
					General inpatient data (demographics, diagnoses, procedures, dates of stay)	Outpatient data	Out-hospital death	In-hospital death
		Children's Hospital Association <a href="http://www.childrenshospitals.org/">http://www.childrenshospitals.org/</a>	Database of clinical (hospital discharge) and financial data from 49 tertiary-care paediatric hospitals in the US affiliated to an alliance of free-standing paediatric hospitals, the Children's Hospital Association	Patient (data can be linked across encounters within the same hospital)	✓ and race insurance type discharge status (discharged to home, rehabilitation facility, nursing facility, or inpatient death) cost medications			
University Health System Consortium Clinical Database Resource Manager database <sup>31</sup>	USA	National Alliance of medical centers <a href="https://www.vizientinc.com/">https://www.vizientinc.com/</a>	Collects discharge data from a National Alliance of 120 academic medical centres and 308 affiliated community hospitals across the USA (~90% of the nation's non-profit academic medical centres)	Hospitalisation	✓ and inpatient procedures severity of index score admission type (ICU admission) discharge location cost provider/hospital characteristics	No	No	✓
Multiple Cause of Death Public-Use Data File <sup>55</sup>	USA	National Center for Health Statistics <a href="https://www.cdc.gov/nchs/nvss/deaths.htm">https://www.cdc.gov/nchs/nvss/deaths.htm</a>	Files derived from compiled data collected from all death certificates issued in the USA for deaths occurring in the USA. The data files included demographic and geographic information on all decedents, the underlying cause of death, and up to 20 conditions listed as contributing causes of death. Deaths occurring outside the country in US citizens and members of the US Armed Forces are excluded	Patient	No	No	✓ demographic geographic information underlying cause of death contributing causes of death	
Discharge Abstract Database of the Canadian Institute for Health Information <sup>49,50</sup>	Canada (less Quebec)	Canadian Institute for Health Information <a href="https://www.cihi.ca/en/types-of-care/hospital-care/acute-care/dad-metadata">https://www.cihi.ca/en/types-of-care/hospital-care/acute-care/dad-metadata</a>	Database of hospitalisation records from all acute care hospitals in Canada (less Quebec), capturing data on admission and discharge dates, patient demographic variables, diagnoses, procedures, and discharge disposition	Patient (unique anonymous patient identification number)	✓ and diagnostic and therapeutic procedures	N/A	N/A	✓
National Health Insurance Research Database <sup>6,42,45,47,48</sup>	Taiwan (99% Taiwan's population is covered)	National Health Insurance Research <a href="http://nhird.nhri.org.tw/en/index.htm">http://nhird.nhri.org.tw/en/index.htm</a>	Includes the entire claims data from all medical facilities contracted with the NHI (outpatient, inpatient, emergency, dental, traditional Chinese medicine services, drug prescriptions), demographic data, for all enrollees in Taiwan	Patient (unique scrambled patient identification number)	✓ and hospital's characteristics reimbursement fees type of visit treatment codes and medications dental care traditional Chinese medicine services	✓	✓	✓
Danish Public Registries <sup>59-61</sup>	Denmark	Danish National Health Service	Nationwide public registries may be cross-linked using a unique personal identification number provided for every inhabitant since 1968. It comprises Danish Civil registration system: dates of birth, immigration, and death Danish National Patient Registry: data on all hospital admissions in Denmark, dates of admission/discharge, surgical procedures and diagnoses Danish National Prescription Registry	Patient (unique patient identification number)	✓ and surgical procedures prescription drugs	✓ Since 1995	✓ Including cause of death	✓ Including cause of death



**Figure 1.** Number of articles on adults with CHD patients using administrative databases published per year and per country. NIS = Nationwide Inpatient Sample.

**Table 2.** Journal impact factors according to the source of publications.

Country	No. of articles (%)	Impact factor, median (IQR)
Quebec	15 (25.7)	14.16 (5.74; 14.62)
Denmark	3 (5.1)	5.69 (4.48; 11.37)
USA non-NIS	14 (23.7)	4.43 (3.26; 5.86)
United Kingdom	1 (1.7)	4.96 (NA)
USA NIS	18 (30.5)	3.40 (2.31; 4.85)
Canada (excluding Quebec)	2 (3.4)	3.11 (3.11; 3.11)
Taiwan	5 (8.5)	1.88 (1.21; 3.15)
Hong Kong	1 (1.7)	1.26 (NA)
Total	59	4.04 (3.15; 7.44)

NIS = Nationwide Inpatient Sample

comprehensiveness and their large size, these databases can generate sample sizes usually not available in single or even multi-institutional databases. This is particularly helpful and relevant when studying rare diagnoses or procedures.

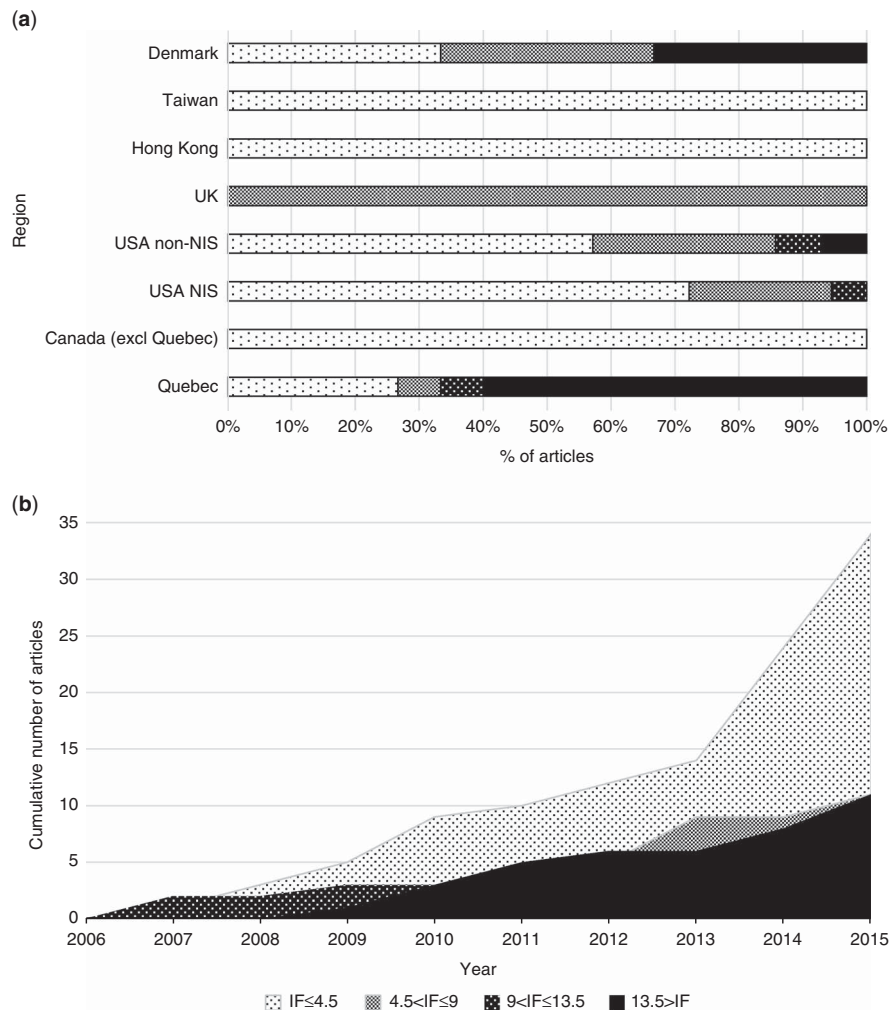
As shown, research based on administrative data sources can be done at a patient level or at a hospitalisation level depending on the characteristics of the data source. The use of a consistent set of identifiers in administrative health databases allows researchers to build histories of individuals. In the Quebec CHD Database, the three available province-wide administrative databases were linked using patients' unique encrypted health care insurance numbers.<sup>51</sup> In Denmark, the number assigned to all residents at birth or upon immigration is included in all national registers, whereas the Danish Government gathered nearly 200 databases, on everything from medical records to socio-economic data on jobs and salaries.<sup>10</sup> Going even further, the Danish Biobank Register stores more than 22 million samples from 5.4 million individuals in national administrative registries at an individual level. Conversely, in other countries, patient identifiers may change over time, making longitudinal analyses more difficult or impossible. Indeed, the unit of measure in the Nationwide Inpatient Sample system is hospital stay, not the patient timeline. The Nationwide Inpatient Sample does not

identify individual patients, and recurrent hospitalisations appear as distinct observations.<sup>69</sup>

This difference of level of analysis is one of the factors that appear to be reflected in the journal impact factor of published studies: Quebec, Denmark, and the United States non-Nationwide Inpatient Sample reach the highest impact factors where patient-level analyses are possible. Another contributing factor to high-impact publications appears to be related to the use of validation procedures to migrate raw data from administrative data sources to “clean” data assembled in the form of a database (Quebec). Publications in high-impact-factor journals reflect enhanced study quality and rigour, with the increase in the cumulative impact factor of published studies over time being an encouraging indicator of the growing quality of studies being produced in adult CHD research with such data sources.

The absence of shared identifiers between the administrative health database and other data sources – for example, a clinical research file – prevents record linkage among heterogeneous data sources. Thus, outpatient data or vital status is not available in United States databases, except for United Network for Organ Sharing, if death did not occur at hospital. In addition, there is usually no link with outpatient clinical data.<sup>69</sup> Methodologies using probabilistic linkage based on variables such as admission date, discharge date, patient sex, and patient date of birth have been developed to merge information from different origins. For example, this methodology has been used to merge information from a registry, the Society of Thoracic Surgeons (STS) Congenital Heart Surgery Database, with a paediatric administrative data set, the Pediatric Health Information Systems (PHIS).<sup>70</sup> Similarly, the availability of direct identifiers allowed linkage of the Pediatric Cardiac Care Consortium data with the National Death Index and the United Network for Organ Sharing, thereby providing significant information regarding the long-term outcomes after surgical procedures.<sup>71</sup> However, until now, none of these methods has been used to assess specific issues in adults.

As CHD is associated with life-long co-morbidities, and also benefits from life-long specialised care, longitudinal studies across the lifespan are essential.<sup>66</sup> In fact, as shown in this systematic review, data are usually available over a large number of years, facilitating longitudinal studies in which unique identifiers can be followed up over time. Follow-up can be traced for years or



**Figure 2.** Distribution of journal impact factors (IF) according to (a) the source of the publication and (b) the year of publication. NIS = Nationwide Inpatient Sample.

decades in order to analyse CHD-specific surgical outcomes, adverse events, or co-morbidities or practice patterns. Using the Quebec CHD Database from 1990 to 2005, Marelli et al showed a significant increase in referrals to specialised adult CHD centres following the introduction of clinical guidelines.<sup>8</sup> This change in clinical practice was independently associated with reduced mortality. Such extensive historical data are important for actionable policy-driven decision-making. This is underscored with studies that monitor health care utilisation, cost of disease burden,<sup>4,5,49,50,56</sup> and inadequate access to services,<sup>72,73</sup> ideally designed with administrative data sources.<sup>74</sup> Thus, administrative health databases are a powerful tool to assess patient management and outcomes and to further develop quality of care improvement programmes.

Diagnostic validity has been an important criticism of CHD studies using administrative data sources.<sup>69</sup> Indeed, in this review, only 37% of studies reported validation procedures. Overall, in this review, all published articles used International Classification of Diseases codes (Eighth, Ninth, or Tenth Revision) to identify CHD patients, which often lack sufficient detail to adequately characterise specific CHD phenotypes or procedures; for example, there is no International Classification of Diseases code for a Norwood procedure. The lack of granularity in the coding schemes – for example, detailed anatomic diagnoses or

procedures – and the lack of standardised definitions may give a coarse overview of the diagnoses or the patient's clinical status. Hence, researchers are limited to investigating broad classes of defects such as severe CHD,<sup>3,51</sup> simple CHD,<sup>59</sup> univentricular,<sup>19,22,24,26,31,33</sup> or valvular diseases.<sup>58</sup> More rarely, CHD with an unequivocal definition as coarctation of the aorta,<sup>28,47,65</sup> or, for example tetralogy of Fallot,<sup>48</sup> have been published specifically. In some studies, lesion-specific algorithms substantially enhance the quality of the work relating to atrial septal defect to distinguish it from persistent foramen ovale.<sup>41</sup> Even when the relevant code exists, however, there may be errors due to the coding process. Physicians may lack expertise in the International Classification of Diseases terminology. In jurisdictions where coding is done by administrative personnel, coding errors may occur owing to staff's limited medical knowledge or because of poor documentation in the medical record, leading to variations in the quality of administrative data on diagnosis.<sup>75,76</sup> In Quebec, authors minimised misclassification bias by using all available data for a given subject, including inpatient, outpatient, procedural, and provider information. From these, they developed an algorithm, and tested it by manually auditing almost a third of the files.<sup>51</sup> In Taiwan, to minimise misclassification bias, at least two corresponding codes were required for confirmation if CHD codes originated from outpatient data, also using procedural codes to enhance

diagnostic accuracy.<sup>6,42</sup> In Denmark, included patients' hospital records were validated to secure a correct diagnosis and status.<sup>60,61</sup>

As administrative databases also contain diagnostic codes for co-morbid conditions, they are a source of misclassification. For example, the studies presented in this review have examined conditions including dementia, gastrointestinal bleed, chronic kidney disease,<sup>66</sup> stroke,<sup>44,47</sup> or coronary artery disease.<sup>65</sup> Depending on the clinical question, inherent limitations in this type of data include the lack of accurate assessment of unmeasured confounders including smoking status, alcohol, drug abuse, and obesity<sup>45,47</sup> or absence of detail on left- versus right-sided heart failure and specifics of prosthetic materials.<sup>44</sup> Similarly, family history, lifestyle factors, and drug prescription information were available only in a few nationwide integrated data systems.<sup>60,61</sup>

Records in administrative health databases only include data for individuals who use the services during the period of interest. Those without access to care, those who failed to encounter the health care system during the study period, or those who may have migrated may thus not be captured. In areas in which access to care is universal as in Taiwan, Quebec, or Denmark, and with long follow-up periods, this bias is minimal.<sup>51,61</sup> Conversely, in countries where health care is supported by private insurance, the information extracted from the administrative health databases may be influenced by access to care and insurance status, socioeconomic level, and ethnicity, thus limiting the generalisability of findings to other countries with different structures for access to care.<sup>3,35,46,64</sup> Recently, based on this knowledge, Gilboa et al<sup>77</sup> estimated the CHD prevalence across all age groups in the United States of America by extrapolating from the population of Quebec and applying a race-ethnicity adjustment factor.

### Limitations

Some limitations of the present systematic review need to be discussed. First, it should be noted that some details on the databases used in some of these studies are not available in the published articles but only available on websites. Moreover, specificities of each administrative health database depend on the specificities of each health care system, which may not be extensively described in each study. Second, the analyses were made at the article level, not at a database level. We recognise that most of the sources have led to several articles. Thus, they may be over-represented when they originated from research groups that are highly productive in terms of publications. However, we did not have source-level data available for analysis. We carried out a systematic review to study the contribution of using a methodology related to administrative health data sources in adult CHD research, so that conducting a meta-analysis or statistical analyses could not be applied. Finally, we reported an increasing number of studies using an administrative health database in the field of adult CHD around the industrialised world, but such studies do not capture the growing population of adult CHD patients in underdeveloped countries.<sup>78</sup>

In conclusion, this systematic literature review focuses on the secondary use of administrative health data sources for adult CHD research purposes in industrialised countries. With the increasing access and use of these data sources, understanding their features and limitations is critical to ensure appropriate interpretation and extends beyond the scope of adult CHD. Although not designed for research purposes, such data sources can be particularly useful for the assessment of population-level

epidemiology, outcomes, and health services research over long observation periods, providing a powerful tool to further develop quality of care improvement programmes. Study quality is enhanced with validation procedures, unique identifiers over time, and comprehensive data capture. Prevailing limitations include diagnostic accuracy in specific subgroups, unmeasured confounders, and lack of clinically relevant patient-level data. Geographic variations in health insurance limit generalisability between jurisdictions. In the future, efforts to standardise diagnostic coding will facilitate data pooling, integration, and reuse of existing data at a supranational level to compare and aggregate results where relevant. Interoperability, quality control, validation, and merging with clinical data sources would optimise the specificity and validity of study findings. Data are increasingly covering a variety of modalities, including administrative databases, electronic medical records, clinical registries, research data sets, monitoring systems, and biobanks. Harmonising data collection will improve the translational potential of adult CHD research.

**Supplementary material.** To view supplementary material for this article, please visit <https://doi.org/10.1017/S1047951118000446>

**Acknowledgements.** The authors thank Sophie Guiquerro for her help in building the search algorithm and Sarah Zohar for her comments.

**Financial Support.** S.C. was supported by the French Federation of Cardiology, Institut Servier, and the Fondation pour la Recherche Médicale. A.J.M. is supported by the Canadian Institute of Health Research, the Heart and Stroke Foundation of Canada, and the Fonds de recherche du Québec – Santé.

**Conflicts of Interest.** None.

**Ethical Standards.** The authors assert that all procedures contributing to this study comply with the ethical standards of relevant national guidelines on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008, and has been approved by the institutional ethics committee of the Queens University of Belfast.

### References

1. Moons P, Bovijn L, Budts W, Belmans A, Gewillig M. Temporal trends in survival to adulthood among patients born with congenital heart disease from 1970 to 1992 in Belgium. *Circulation* 2010; 122: 2264–2272.
2. van der Bom T, Bouma BJ, Meijboom FJ, Zwiderman AH, Mulder BJ. The prevalence of adult congenital heart disease, results from a systematic review and evidence based calculation. *Am Heart J* 2012; 164: 568–575.
3. Marelli AJ, Ionescu-Ittu R, Mackie AS, Guo L, Dendukuri N, Kaouache M. Lifetime prevalence of congenital heart disease in the general population from 2000 to 2010. *Circulation* 2014; 130: 749–756.
4. Mackie AS, Pilote L, Ionescu-Ittu R, Rahme E, Marelli AJ. Health care resource utilization in adults with congenital heart disease. *Am J Cardiol* 2007; 99: 839–843.
5. Opatowsky AR, Siddiqi OK, Webb GD. Trends in hospitalizations for adults with congenital heart disease in the U.S. *J Am Coll Cardiol* 2009; 54: 460–467.
6. Lin YS, Liu PH, Wu LS, Chen YM, Chang CJ, Chu PH. Major adverse cardiovascular events in adult congenital heart disease: a population-based follow-up study from Taiwan. *BMC Cardiovasc Disord* 2014; 14: 38.
7. Mackie AS, Ionescu-Ittu R, Therrien J, Pilote L, Abrahamowicz M, Marelli AJ. Children and adults with congenital heart disease lost to follow-up: who and when? *Circulation* 2009; 120: 302–309.
8. Mylotte D, Pilote L, Ionescu-Ittu R, et al. Specialized adult congenital heart disease care: the impact of policy on mortality. *Circulation* 2014; 129: 1804–1812.



9. Yeung E, Kay J, Roosevelt GE, Brandon M, Yetman AT. Lapse of care as a predictor for morbidity in adults with congenital heart disease. *Int J Cardiol* 2008; 125: 62–65.
10. Frank L. Epidemiology. When an entire country is a cohort. *Science* 2000; 287: 2398–2399.
11. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015; 349: g7647.
12. Mitchell SC, Korones SB, Berendes HW. Congenital heart disease in 56,109 births. Incidence and natural history. *Circulation* 1971; 43: 323–332.
13. Gurvitz MZ, Inkelas M, Lee M, Stout K, Escarce J, Chang RK. Changes in hospitalization patterns among patients with congenital heart disease during the transition from adolescence to adulthood. *J Am Coll Cardiol* 2007; 49: 875–882.
14. Opotowsky AR, Landzberg MJ, Kimmel SE, Webb GD. Percutaneous closure of patent foramen ovale and atrial septal defect in adults: the impact of clinical variables and hospital procedure volume on in-hospital adverse events. *Am Heart J* 2009; 157: 867–874.
15. Karamlou T, Diggs BS, Ungerleider RM, Welke KF. Adults or big kids: what is the ideal clinical environment for management of grown-up patients with congenital heart disease? *Ann Thorac Surg* 2010; 90: 573–579.
16. Kim YY, Gauvreau K, Bacha EA, Landzberg MJ, Benavidez OJ. Resource use among adult congenital heart surgery admissions in pediatric hospitals: risk factors for high resource utilization and association with inpatient death. *Circ Cardiovasc Qual Outcomes* 2011; 4: 634–639.
17. Kim YY, Gauvreau K, Bacha EA, Landzberg MJ, Benavidez OJ. Risk factors for death after adult congenital heart surgery in pediatric hospitals. *Circ Cardiovasc Qual Outcomes* 2011; 4: 433–439.
18. Rodriguez FH 3rd, Moodie DS, Parekh DR, et al. Outcomes of hospitalization in adults in the United States with atrial septal defect, ventricular septal defect, and atrioventricular septal defect. *Am J Cardiol* 2011; 108: 290–293.
19. Karamlou T, Diggs BS, Welke K, et al. Impact of single-ventricle physiology on death after heart transplantation in adults with congenital heart disease. *Ann Thorac Surg* 2012; 94: 1281–1287; discussion 1287–1288.
20. Opotowsky AR, Siddiqi OK, D'Souza B, Webb GD, Fernandes SM, Landzberg MJ. Maternal cardiovascular events during childbirth among women with congenital heart disease. *Heart* 2012; 98: 145–151.
21. Baraona F, Gurvitz M, Landzberg MJ, Opotowsky AR. Hospitalizations and mortality in the United States for adults with Down syndrome and congenital heart disease. *Am J Cardiol* 2013; 111: 1046–1051.
22. Collins RT 2nd, Fram RY, Tang X, Robbins JM, Sutton MS. Impact of anatomical subtype and medical comorbidities on hospitalizations in adults with single ventricle congenital heart disease. *Int J Cardiol* 2013; 168: 4596–4601.
23. Gelow JM, Song HK, Weiss JB, Mudd JO, Broberg CS. Organ allocation in adults with congenital heart disease listed for heart transplant: impact of ventricular assist devices. *J Heart Lung Transplant* 2013; 32: 1059–1064.
24. Krieger EV, Moko LE, Wu F, et al. Single ventricle anatomy is associated with increased frequency of nonalcoholic cirrhosis. *Int J Cardiol* 2013; 167: 1918–1923.
25. Rodriguez FH 3rd, Moodie DS, Parekh DR, et al. Outcomes of heart failure-related hospitalization in adults with congenital heart disease in the United States. *Congenit Heart Dis* 2013; 8: 513–519.
26. Collins RT 2nd, Fram RY, Tang X, Robbins JM, St John Sutton M. Hospital utilization in adults with single ventricle congenital heart disease and cardiac arrhythmias. *J Cardiovasc Electrophysiol* 2014; 25: 179–186.
27. Bhatt AB, Rajabali A, He W, Benavidez OJ. High resource use among adult congenital heart surgery admissions in adult hospitals: risk factors and association with death and comorbidities. *Congenit Heart Dis* 2015; 10: 13–20.
28. Bhatt P, Patel NJ, Patel A, et al. Impact of hospital volume on outcomes of endovascular stenting for adult aortic coarctation. *Am J Cardiol* 2015; 116: 1418–1424.
29. Fernandes SM, Chamberlain LJ, Grady S Jr., et al. Trends in utilization of specialty care centers in California for adults with congenital heart disease. *Am J Cardiol* 2015; 115: 1298–1304.
30. Hayward RM, Dewland TA, Moyers B, et al. Device complications in adult congenital heart disease. *Heart Rhythm* 2015; 12: 338–344.
31. Seckeler MD, Moe TG, Thomas ID, et al. Hospital resource utilization for common noncardiac diagnoses in adult survivors of single cardiac ventricle. *Am J Cardiol* 2015; 116: 1756–1761.
32. Singh V, Badheka AO, Patel NJ, et al. Influence of hospital volume on outcomes of percutaneous atrial septal defect and patent foramen ovale closure: a 10-years US perspective. *Catheter Cardiovasc Interv* 2015; 85: 1073–1081.
33. Tabatabai S, DeFaria Yeh D, Stefanescu A, Kennedy K, Yeh RW, Bhatt AB. National trends in hospitalizations for patients with single-ventricle anatomy. *Am J Cardiol* 2015; 116: 773–778.
34. Thompson JL, Kuklina EV, Bateman BT, Callaghan WM, James AH, Groteg CA. Medical and obstetric outcomes among pregnant women with congenital heart disease. *Obstet Gynecol* 2015; 126: 346–354.
35. Beausejour Ladouceur V, Lawler PR, Gurvitz M, et al. Exposure to low-dose ionizing radiation from cardiac procedures in patients with congenital heart disease: 15-year data from a population-based longitudinal cohort. *Circulation* 2016; 133: 12–20.
36. Patel ND, Weiss ES, Allen JG, et al. Heart transplantation for adults with congenital heart disease: analysis of the United network for organ sharing database. *Ann Thorac Surg* 2009; 88: 814–821; discussion 821–822.
37. Bernier M, Marelli AJ, Pilote L, et al. Atrial arrhythmias in adult patients with right- versus left-sided congenital heart disease anomalies. *Am J Cardiol* 2010; 106: 547–551.
38. Karamlou T, Hirsch J, Welke K, et al. A United Network for Organ Sharing analysis of heart transplantation in adults with congenital heart disease: outcomes and factors associated with mortality and retransplantation. *J Thorac Cardiovasc Surg* 2010; 140: 161–168.
39. Khairy P, Ionescu-Ittu R, Mackie AS, Abrahamowicz M, Pilote L, Marelli AJ. Changing mortality in congenital heart disease. *J Am Coll Cardiol* 2010; 56: 1149–1157.
40. Kotowycz MA, Therrien J, Ionescu-Ittu R, et al. Long-term outcomes after surgical versus transcatheter closure of atrial septal defects in adults. *JACC Cardiovasc Interv* 2013; 6: 497–503.
41. Mylotte D, Quenneville SP, Kotowycz MA, et al. Long-term cost-effectiveness of transcatheter versus surgical closure of secundum atrial septal defect in adults. *Int J Cardiol* 2014; 172: 109–114.
42. Chen TH, Hsiao YC, Cheng CC, et al. In-hospital and 4-year clinical outcomes following transcatheter versus surgical closure for secundum atrial septal defect in adults: a national cohort propensity score analysis. *Medicine (Baltimore)* 2015; 94: e1524.
43. Hassan N, Patenaude V, Oddy L, Abenhaim HA. Pregnancy outcomes in Marfan syndrome: a retrospective cohort study. *Am J Perinatol* 2015; 30: 123–130.
44. Lanz J, Brophy JM, Therrien J, Kaouache M, Guo L, Marelli AJ. Stroke in adults with congenital heart disease: incidence, cumulative risk, and predictors. *Circulation* 2015; 132: 2385–2394.
45. Lee YS, Chen YT, Jeng MJ, et al. The risk of cancer in patients with congenital heart disease: a nationwide population-based cohort study in Taiwan. *PLoS One* 2015; 10: e0116844.
46. Maxwell BG, Stepan J, Cheng A. Complications of catheter-based electrophysiology procedures in adults with congenital heart disease: a national analysis. *J Cardiothorac Vasc Anesth* 2015; 29: 258–264.
47. Wu MH, Chen HC, Kao FY, Huang SK. Risk of systemic hypertension and cerebrovascular accident in patients with aortic coarctation aged <60 years (from a National Database Study). *Am J Cardiol* 2015; 116: 779–784.
48. Wu MH, Lu CW, Chen HC, Chiu SN, Kao FY, Huang SK. Arrhythmic burdens in patients with tetralogy of Fallot: a national database study. *Heart Rhythm* 2015; 12: 604–609.
49. Islam S, Yasui Y, Kaul P, Mackie AS. Hospital readmission of patients with congenital heart disease in Canada. *Can J Cardiol* 2016; 32: 987.e987–987.e914.

50. Islam S, Yasui Y, Kaul P, Marelli AJ, Mackie AS. Congenital heart disease hospitalizations in Canada: a 10-year experience. *Can J Cardiol* 2016; 32: 197–203.
51. Marelli AJ, Mackie AS, Ionescu-Ittu R, Rahme E, Pilote L. Congenital heart disease in the general population: changing prevalence and age distribution. *Circulation* 2007; 115: 163–172.
52. Billett J, Majeed A, Gatzoulis M, Cowie M. Trends in hospital admissions, in-hospital case fatality and population mortality from congenital heart disease in England, 1994 to 2004. *Heart* 2008; 94: 342–348.
53. Chan YC, Ting CW, Ho P, Poon JT, Cheung GC, Cheng SW. Ten-year epidemiological review of in-hospital patients with Marfan syndrome. *Ann Vasc Surg* 2008; 22: 608–612.
54. Karamlou T, Diggs BS, Ungerleider RM, McCrindle BW, Welke KF. The rush to atrial septal defect closure: is the introduction of percutaneous closure driving utilization? *Ann Thorac Surg* 2008; 86: 1584–1590; discussion 1590–1591.
55. Gilboa SM, Salemi JL, Nembhard WN, Fixler DE, Correa A. Mortality resulting from congenital heart disease among children and adults in the United States, 1999 to 2006. *Circulation* 2010; 122: 2254–2263.
56. Lu Y, Agrawal G, Lin CW, Williams RG. Inpatient admissions and costs of congenital heart disease from adolescence to young adulthood. *Am Heart J* 2014; 168: 948–955.
57. Maxwell BG, Maxwell TG, Wong JK. Decentralization of care for adults with congenital heart disease in the United States: a geographic analysis of outpatient surgery. *PLoS One* 2014; 9: e106730.
58. Ionescu-Ittu R, Mackie AS, Abrahamowicz M, et al. Valvular operations in patients with congenital heart disease: increasing rates from 1988 to 2005. *Ann Thorac Surg* 2010; 90: 1563–1569.
59. Videbaek J, Laursen HB, Olsen M, Hofsten DE, Johnsen SP. Long-term nationwide follow-up study of simple congenital heart disease diagnosed in otherwise healthy children. *Circulation* 2016; 133: 474–483.
60. Nyboe C, Olsen MS, Nielsen-Kudsk JE, Hjortdal VE. Atrial fibrillation and stroke in adult patients with atrial septal defect and the long-term effect of closure. *Heart* 2015; 101: 706–711.
61. Nyboe C, Olsen MS, Nielsen-Kudsk JE, Johnsen SP, Hjortdal VE. Risk of pneumonia in adults with closed versus unclosed atrial septal defect (from a nationwide cohort study). *Am J Cardiol* 2014; 114: 105–110.
62. Maxwell BG, Wong JK, Kin C, Lobato RL. Perioperative outcomes of major noncardiac surgery in adults with congenital heart disease. *Anesthesiology* 2013; 119: 762–769.
63. Lowe BS, Therrien J, Ionescu-Ittu R, Pilote L, Martucci G, Marelli AJ. Diagnosis of pulmonary hypertension in the congenital heart disease adult population impact on outcomes. *J Am Coll Cardiol* 2011; 58: 538–546.
64. Bouchardy J, Therrien J, Pilote L, et al. Atrial arrhythmias in adults with congenital heart disease. *Circulation* 2009; 120: 1679–1686.
65. Roifman I, Therrien J, Ionescu-Ittu R, et al. Coarctation of the aorta and coronary artery disease: fact or fiction? *Circulation* 2012; 126: 16–21.
66. Afilalo J, Therrien J, Pilote L, Ionescu-Ittu R, Martucci G, Marelli AJ. Geriatric congenital heart disease: burden of disease and predictors of mortality. *J Am Coll Cardiol* 2011; 58: 1509–1515.
67. van der Velde ET, Vriend JW, Mannens MM, Uiterwaal CS, Brand R, Mulder BJ. CONCOR, an initiative towards a national registry and DNA-bank of patients with congenital heart disease in the Netherlands: rationale, design, and first results. *Eur J Epidemiol* 2005; 20: 549–557.
68. Brida M, Dimopoulos K, Kempny A, et al. Body mass index in adult congenital heart disease. *Heart* 2017; 103: 1250–1257.
69. Khera R, Krumholz HM. With great power comes great responsibility: big data research from the national inpatient sample. *Circ Cardiovasc Qual Outcomes* 2017; 10: pii: e003846.
70. Pasquali SK, Jacobs JP, Shook GJ, et al. Linking clinical registry data with administrative data using indirect identifiers: implementation and validation in the congenital heart surgery population. *Am Heart J* 2010; 160: 1099–1104.
71. Spector LG, Menk JS, Vinocur JM, et al. In-hospital vital status and heart transplants after intervention for congenital heart disease in the Pediatric Cardiac Care Consortium: completeness of ascertainment using the national death index and united network for organ sharing datasets. *J Am Heart Assoc* 2016; 5: e003783.
72. Marelli AJ, Gurvitz M. From numbers to guidelines. *Prog Cardiovasc Dis* 2011; 53: 239–246.
73. Marelli AJ, Therrien J, Mackie AS, Ionescu-Ittu R, Pilote L. Planning the specialized care of adult congenital heart disease patients: from numbers to guidelines; an epidemiologic approach. *Am Heart J* 2009; 157: 1–8.
74. Riehle-Colarusso TJ, Bergersen L, Broberg CS, et al. Databases for congenital heart defect public health studies across the lifespan. *J Am Heart Assoc* 2016; 5: pii: e004148.
75. Cronk CE, Malloy ME, Pelech AN, et al. Completeness of state administrative databases for surveillance of congenital heart disease. *Birth Defects Res A Clin Mol Teratol* 2003; 67: 597–603.
76. Frohnert BK, Lussky RC, Alms MA, Mendelsohn NJ, Symonik DM, Falken MC. Validity of hospital discharge data for identifying infants with cardiac defects. *J Perinatol* 2005; 25: 737–742.
77. Gilboa SM, Devine OJ, Kucik JE, et al. Congenital heart defects in the United States: estimating the magnitude of the affected population in 2010. *Circulation* 2016; 134: 101–109.
78. Webb G, Mulder BJ, Aboulhosn J, et al. The care of adults with congenital heart disease across the globe: current assessment and future perspective: a position statement from the International Society for Adult Congenital Heart Disease (ISACHD). *Int J Cardiol* 2015; 195: 326–333.