Lifetime Prevalence of Congenital Heart Disease in the General Population
from 2000 to 2010

Running title: Marelli et al.; Prevalence of Congenital Heart Disease 2000-2010

Ariane J. Marelli, MD, MPH1; Raluca Ionescu-Ittu, PhD1; Andrew S. Mackie, MD, SM2; Liming Guo, MSc1; Nandini Dendukuri, PhD3; Mohammed Kaouache, PhD1

1McGill Adult Unit for Congenital Heart Disease Excellence, McGill University Health Centre, Montreal, Canada; 2Division of Cardiology, Stollery Children’s Hospital & Dept of Pediatrics, University of Alberta, Edmonton, Canada; 3Division of Clinical Epidemiology, McGill University Health Centre, Montreal, Canada

Address for Correspondence:
Ariane J. Marelli, MD, MPH
Professor of Medicine, McGill University
Director, McGill Adult Unit for Congenital Heart Disease Excellence, Royal Victoria Hospital
687 Pine Avenue West, Room H4.33
Montreal, Quebec H3A 1A1
Tel: 514-934-1934, x43153
Fax: 514-934-4475
Email: ariane.marelli@mcgill.ca

Journal Subject Code: Cardiovascular (CV) surgery:[41] Pediatric and congenital heart disease, including cardiovascular surgery
Abstract

**Background**—Our objective was to obtain contemporary life-time estimates of CHD prevalence using population based data sources up to year 2010.

**Methods and Results**—The Quebec CHD database contains 28 years of longitudinal data on all individuals with CHD from 1983-2010. Severe CHD was defined as tetralogy of Fallot, truncus arteriosus, transposition complexes, endocardial cushion defects, and univentricular hearts. We used latent class Bayesian models combining case definitions from physician claims, hospitalization, and surgical data to obtain point and interval prevalence estimates (95% CI) of CHD in the first year of life, in children (<18 years of age) and in adults. We identified 107,559 CHD patients from 1983 to 2010. Prevalence of CHD in the first year of life was 8.21/1000 live-births (7.47, 9.02) from 1998 to 2005. In 2010, overall prevalence of CHD was 13.11/1000 (12.43, 13.81) in children and 6.12 /1000 (5.69, 6.57) in adults. The CHD prevalence increased by 11% in children and 57% in adults from 2000-2010. The prevalence in the severe CHD subgroup increased by 19% (17%, 21%) in children and 55% (51%, 62%) in adults. By 2010, adults accounted for 66% of the entire CHD population.

**Conclusions**—With an increase of more than 50% in CHD prevalence since 2000, by 2010 adults accounted for two-thirds of patients with severe and other forms of CHD in the general population. Our findings should inform allocation of resources and the planning of workforce needs for the predominantly adult CHD population.

**Key words:** congenital heart disease, adult congenital heart disease, epidemiology, health services research, prevalence, severe congenital heart disease
Introduction

Congenital heart disease (CHD) is diagnosed in 8.14 of 1,000 or close to 1% of births in the United States (US). Advances in medical and surgical therapy have increased the survival of CHD patients. It is now recognized that CHD is associated with life-long comorbidity impacting health services utilization and costs. The impact of ongoing disease burden includes atrial arrhythmias, pulmonary hypertension and repeated need for surgery resulting in significant increases in health services utilization during childhood, transition years, adulthood and in the geriatric age group. Not surprisingly in the US in 2004, birth defects accounted for more than 139,000 hospitalizations of which 46,500 were for cardiac and circulatory anomalies. In the same year, whereas hospital costs associated with all birth defects totalled $2.6 billion, an approximate $1.4 billion was directly attributable to cardiovascular anomalies. Therefore, although cardiovascular birth defects accounted for only 34% percent of stays for birth defect hospitalizations, they comprised more than half the costs. Thus, as the demographic distribution of disease changes, policy makers need accurate estimates of the growing numbers of CHD patients.

Using administrative databases with universal health coverage, in Quebec, Canada, first-time estimates of time trends from 1985-2000 in the CHD prevalence in the general population were published. A 22% increase in children and 85% increase in adults with severe CHD was documented from 1985 to year 2000 with the number of adults accounting for just over 50% of the total CHD population by the year 2000.

Our goal in this study was to provide contemporary population-based estimates for CHD prevalence across the life-span. Our specific objectives were to estimate the lifetime prevalence of CHD; to compare the number of adults to the number of children with CHD in the Quebec
population from 2000 to 2010 and to estimate the change in prevalence of severe CHD in
Quebec from 2000 to 2010 in adults compared to children. We also sought to determine the
prevalence of CHD in infants. In addition to doing so for the year 2010, in order to be consistent
with the study period of the Centers for Disease Control in the Atlanta region (CDC), we
estimated the prevalence of CHD in the first year of life from 1998-2005.¹

Methods

Data sources

In Quebec, Canada every individual is assigned a unique Medicare number in the first year of
life used to track all diagnoses and health services rendered and systematically recorded until
death. Administrative databases include the physicians’ services and drug claims, hospital
discharge summary databases and the Quebec Health Insurance Board and Death Registry.³
These contain demographic and vital statistics data as well as all ICD-9 and 10 (as of 2006)
diagnostic codes and procedure codes recorded with or without a hospitalization in Quebec since
1983.

To create the Quebec Congenital Heart Disease Database from 1983-2000, information
on all patients with a CHD diagnosis from either of the three data sources were merged and
cleaned on site at the McGill Adult Unit for Congenital Heart Disease Excellence. Patients were
identified with CHD if they had at least one diagnostic code for CHD and/or a CHD specific
surgical procedure and diagnostic algorithms were developed as published to optimize the
extraction of valid CHD diagnoses.³ To increase the specificity of the algorithms, on the original
dataset from 1983-2000 consisting of 61,386 subjects, manual audits were performed
independently by an adult congenital heart disease specialist (AJM) and a pediatric cardiologist
(ASM). These were carried out on random samples of 19,073 (or 31%) of included and excluded subjects in the raw data set to adjudicate discrepant diagnoses and adjust algorithms accordingly.3

Data collection

For the current study, using the same patient encrypted unique identifiers used for the data from 1983-2000, data updates were requested from the same and only administrative data source in Quebec up to the year 2010. Because a subject carries a diagnosis of CHD since birth, he/she entered each birth cohort based on his/her age, regardless of the calendar year of diagnosis. To minimize misclassification and increase sensitivity of using CHD diagnostic codes to capture CHD patients, we used all available data for a given subject over a 28-year observation period, including inpatient, outpatient, procedural and provider information. This was done by cross-referencing all available province-wide administrative databases. Patients who had more than 15 years of observation with only a single contact with the health care system billed as a severe or unspecified CHD diagnosis where excluded. New data was thus obtained on existing and new CHD cases diagnosed from 1983-2010 on a total of 157,395 subjects (Figure 1) in whom we applied our previously published algorithms3 to arrive at the total number of subjects with a CHD diagnosis in Quebec by the year 2010. All data used in this study were de-identified and the study was approved by the Quebec Commission for Access to Information which regulates access to confidential data in Quebec.

The updated Quebec CHD database therefore contains comprehensive longitudinal, demographic, diagnostic, and therapeutic records of all patient-linked encounters with the healthcare system from January 1, 1983 to December 31, 2010 (inclusive) for all Quebec residents identified with CHD linked with a singular scrambled identifier over a patient’s life.
Severe CHD was defined as tetralogy of Fallot (TOF), truncus arteriosus (TA), transposition complex (TGA), endocardial cushion defects, univentricular heart and hypoplastic left heart syndrome as previously published. All remaining diagnoses were considered “other” CHD lesions. The prevalence analysis in 2010 consisted of the number of patients who were alive with CHD in Quebec as of mid-year 2010 (numerator) and the mid-year 2010 Quebec general population (denominator) which was 7,929,365 and comprised of 1,523,722 children and 6,405,643 adults. For the prevalence analysis in 2005, the Quebec population in mid-2005 consisted of 1,550,513 children and 6,030,679 adults. For the prevalence estimate in infants, the total number of infants alive with CHD measured at mid-year of each year from 1998-2005 and was divided by 589,570 or the total number of infants alive in the Quebec population during the same years. These dates were chosen to overlap with the observation period of data from the CDC. The same analysis was also done for 2010 using 87,760 as the denominator for the Quebec infant population. Age in a specific year was defined as the age, in years, as of July 1 of that year. Infants were defined as subjects <1, children as <18 and adults as ≥ 18 years of age respectively.

Statistical methods

Prevalence during a specific year was defined as the number of cases who were alive as of July 1 of that specific year and who were diagnosed before the end of that specific year, divided by the size of the corresponding mid-year Quebec population. Prevalence was reported as the number of cases per 1000 individuals. The prevalence ratio of two different calendar years corresponded to the ratio of their corresponding prevalence estimates. The proportion of adults and children with CHD was defined as the number of adults and children divided by the total number of CHD patients alive in that year.
We obtained data on 61,386 subjects from 1983-2000 and new data on a total of 157,395 subjects by the year 2010. We determined that we could not reasonably perform manual audits on the new data readjusting our algorithms based on the longer observation period. Since there is no perfect measure of CHD prevalence, we used Bayesian latent class models to account for uncertainty around the point estimates. The ICD-9 codes as well as the surgical procedure codes are both imperfect measures of the underlying true CHD status, which is latent. Adopting the terminology used for evaluation of diagnostic tests, we describe the ICD-9 codes as having high sensitivity and low specificity, while the surgical procedure codes have high specificity but low sensitivity. Separate latent class models were used to predict prevalence of different types of lesions, as well as “all lesions”, “severe lesions” and all “other lesions”. This type of model has been used for the estimation of disease prevalence in the absence of a gold-standard test, and similar models have been applied to estimate the prevalence of various outcomes using imperfect measures from administrative databases.

In order to obtain a meaningful estimate of the prevalence, some information external to the data (referred to as prior information) must be provided on at least two parameters. To ensure that prior data will not be used twice to derive prior distribution and the likelihood, we used the sensitivity of procedure codes estimated using data over a first 14 year of follow up period (1983-1996) to elicit a prior distribution for the sensitivity of procedure codes over a second 14 year period (1996-2010), such that the two periods did not overlap. The 95% confidence interval for the sensitivity estimate obtained from the earlier cohort (data from January 1, 1983 to December 31, 1996) was used to define the lower and upper bounds of the prior distribution. Based on our previous work using the Quebec CHD database from 1983-2000 as described above, our prior knowledge of the CHD specific surgical procedure codes suggests
the specificity is as close to 100% as possible.

Each individual in the mid-year Quebec population of 2010 was thus assigned to one of four cells of a 2x2 table on the basis of the results of diagnoses of CHD by the two ascertainment methods. Likelihood was then constructed based on the multinomial distribution of the four categories in the 2x2 table. The unknown parameters in the model are the CHD prevalence (our primary parameter of interest), and the sensitivity and specificity of the two imperfect measures. The same methodology was used to obtain prevalence rates for adults and children in 2005 and for infants in the first year of life from 1998-2005. The prevalence of severe CHD was estimated in a similar way, but using the ICD-9 codes for severe CHD in the numerator as defined above and as previously published.3 Age specific prevalence was estimated with the prior obtained from the corresponding age categories.

All descriptive analyses were performed using SAS software (SAS institute Inc SAS/STAT Software 9.2, Cary, NC). Estimations of the latent class models, such as the prevalence and its 95% credible interval, were carried out using WinBUGS software 1.4.3.21 Details on the variables used for the latent class models and technical details on the programs used are described in the Statistical Appendix.

Results

Figure 1 illustrates the derivation of the 107,559 subjects in the Quebec health claims databases that had either a CHD diagnosis and/or congenital cardiac procedure code from 1983 to 2010. After application of our algorithms and the latent class modelling to account for uncertainty about the true CHD status or prevalence of CHD, it was estimated that 39,051 adults and 20,011 children with CHD were alive in mid-year 2010 in Quebec (95% CIs 36,307 to 41,922 and
18,973 to 21,079, respectively) as shown in Table 1.

Life-time prevalence of CHD

The prevalence of CHD across age groups is presented in Figure 2. A prevalence of 13.11 (12.43, 13.81) per 1,000 children and 6.12 (5.69, 6.57) per 1,000 adults was documented in 2010. A subgroup analysis was performed in children to identify the birth prevalence in the first year of life from 1998-2005 and in 2010. We obtained a rate of 8.21, 95% CI (7.47,9.02) for years 1998-2005 and 8.12, 95% CI (7.59,8.87) per 1,000 infants for 2010.

Prevalence by lesion for 2010 is summarized in Table 1. The prevalence of severe CHD was 1.76 (1.68, 1.84) per 1,000 children and 0.62 (0.56, 0.68) per 1,000 adults. Conotruncal anomalies (TOF, truncus arteriosus, TGA) were present in 0.98 per 1,000 children and 0.33 per 1,000 adults. The prevalence of “other” CHD was 11.34 (10.67, 12.04) per 1,000 children and 5.50 (5.07, 5.95) per 1,000 adults. All shunt lesions including unspecified defects of septal closure had prevalence values of 9.19 per 1,000 children and 1.60 per 1,000 adults. Valve lesions including congenital aortic stenosis or insufficiency, anomalies of the pulmonary artery or valve, congenital mitral or tricuspid valve disease, and Ebstein’s anomaly had a prevalence of 1.28 per 1,000 children and 1.73 per 1,000 adults.

Change in proportion of adults and children in years 2000, 2005 and 2010

The number of adults and children (95% CI) with any CHD is presented in Figure 3. For children, the number with any CHD (95% CI) remained relatively stable with 18,913 (18,586, 19,241) in 2000 and 20,011 (18,973, 21,079) in 2010. For adults however, the number with any CHD went from 22,291 (19,181, 25,402) in 2000 to 39,051 (36,307, 41,922) in 2010 (Figure 3, Panel A). Similarly, while the number of children with a severe lesion increased slightly from 2,355 (2,283, 2,427) in 2000 to 2,686 (2,564, 2,809) in 2010, a substantial increase was observed
in the number of adults with severe CHD which went from 2,275 (2,131, 2,419) in 2000 to 3,956 (3,573, 4,339) in 2010 (Figure 3, Panel B).

The proportion of subjects with CHD who were adults increased from 54% (51%, 57%) in 2000 to 66% (64%, 68%) (Figure 3, Panel A). Similarly, in 2010, adults represented 60% (57%, 62%) of all subjects with severe CHD as compared to 49% (47%, 51%) in 2000 (Figure 3, Panel B).

**Changing prevalence in adults and children in years 2000, 2005, and 2010**

From 2000 to 2010 we noted, in both children and adults an increase in the prevalence of all CHD and severe CHD, however, a much larger increase was observed in adults compared to children. The prevalence of CHD increased by 57% (48%, 71%) in adults compared to 11% (5%, 17%) in children, and as a consequence, adults represented two thirds of the CHD population in 2010. Similarly, an increase in the prevalence of severe lesions of 55% (51%, 62%) was estimated in adults, which is larger than the 19% (17%, 21%) increase estimated in children (Figure 4).

**Prevalence ratios of CHD population stratified by age in year 2010 compared to year 2000**

We analyzed how the increase in prevalence ratios of all and severe lesions behaved within specific age categories as presented in Figure 5, Panels A and B. Prevalence ratios for 2010/2000 indicate that in those with severe disease, the largest increase was observed among those of 26 years of age or older with a prevalence ratio of 1.84 (1.67, 2.03) in those 26 to 40 years of age and 1.79 (1.68, 1.89) amongst those 41 years of age or older. Prevalence ratios for total CHD also followed the same trends with the largest increase observed in adults above the age of 26, although the rise was not as sharp in young adults 18-25 years of age.
Discussion

This study is unique in providing contemporary population-based estimates of the life-time prevalence of CHD based on longitudinal follow-up. In 2010, adults represented 66% of the total number of people alive with CHD in Quebec who had severe and other forms of CHD. Therefore, adults with CHD now constitute a significant majority of CHD patients accessing the health care system in the general population compelling us to reshape the delivery of health care services for patients with CHD. From 2000 to 2010, the prevalence of severe CHD rose at faster rates for adults compared to children consistent with ongoing improvement in care and survival.

In the Quebec CHD database, as presented in this study, we observed a prevalence of CHD in infancy, remarkably close to the most commonly reported rate of birth prevalence of CHD in the US. ¹ This finding supports the use of our data to estimate the number of adults with CHD in the US and underscores the fact that population based empirical measurements can and should be sought for in other jurisdictions.

From 2000-2010, the increase in prevalence of severe CHD in adults was 55% (51%, 62%) compared to an 85% increase previously observed from 1985-2000.³ Despite a slower rate of rise, adults with CHD now constitute 60% of the severe CHD population and 66% of persons with other CHD. Consistent with this observation, the largest increase in prevalence stratified by age occurred for those 18-40 years of age from 2000-2010 compared to those in the 13-17 age range from 1985-2000. As a result of this larger increase in adults compared to children, the median age of persons alive with severe CHD in the Quebec population increased by 14 years from 1985 to 2010. The median age was estimated in a previous publication by Marelli et al.³ to be 11 years in 1985 and 17 years in 2000, while in this study we found that the median age of persons with severe CHD in 2010 was 25 years in 2010. The sequential increase in age of the
most rapidly growing segment of the CHD population suggests that the next decades will witness
an increasingly older population of patients not only with severe CHD but with the comorbidity
that is expected to add to the disease burden.\textsuperscript{13}

Improved care, decreased mortality and/or improved diagnosis over the life-span are
likely contributors to the observation of an increasing prevalence of CHD. We and others have
found a decrease in mortality of CHD patients as documented in Canada, the US and other
industrialized nations.\textsuperscript{4, 6, 22} This decrease in mortality over time is likely an important
contributing factor to an increasing pool of prevalent CHD patients, especially for patients with
severe CHD. For patients with mild forms of CHD such as aortic and mitral valve disease, the
rise in prevalence from childhood to adulthood is consistent with improvement in diagnostic
techniques and/or presentation in adulthood with an increased likelihood of being captured with
longer observation periods. For example, bicuspid aortic valve is often not detected in childhood
as the physical finding (a click) is subtle and until stenosis and/or regurgitation develops, often
not until adulthood, the diagnosis is not made. The reported rise in adults of mitral or tricuspid
valve disease could be due to progressive atrioventricular valve regurgitation following
endocardial cushion defect repair. Recently we have observed a significant decrease in mortality
associated with the delivery of specialized care for adults with severe and other forms of CHD.\textsuperscript{23}
These findings taken together provide the evidence base needed to improve quality of care for
adults with CHD\textsuperscript{24} in order to improve outcomes.

In the Quebec CHD database we observed a prevalence of CHD in infancy of 8.21/1,000
from 1998-2005 and 8.12/1,000 for the year 2010. Our results show that the 2010 CHD
prevalence in infancy is lower than the prevalence in childhood. The most likely explanation for
this, is related to ascertainment of mild CHD that is often diagnosed later in childhood. Therefore
it is not surprising that with up to 18 years of follow-up in children we capture higher rates of total CHD prevalence that includes both severe and milder forms of CHD. The analysis from 1998-2005 was performed for comparative purposes with other analyses of birth prevalence. While our methodology is different from Reller et al., our findings are within the same range as the rate of 8.14/1,000 published by the CDC in the US and similar to the most commonly reported rate of birth prevalence of CHD in industrialized nations that clusters around 8/1,000.

Mid-year was used for infants because Statistics Canada, from which the denominators were taken, reports population estimates at July 1 of each year. This insured that the same inclusion criteria are applied to the numerator and denominator of prevalence rate estimates in infants. To the extent that infant mortality is low in the current era, the mid-year population of infants can be expected to be a reasonable proxy for the (live) births in a given year. The similarity in birth prevalence rates between our data and the US data lends credibility to the extrapolation of our findings to jurisdictions where advances in medical and surgical therapy are expected to yield similar results in survival.

Population based prevalence rates may be extrapolated to arrive at estimates of the number of people alive with CHD. The challenge in measuring prevalence in children and adults is in obtaining a meaningful denominator. Because of universal health coverage in Canada and the use of a unique identifier throughout a subject’s life, the Quebec CHD database is one of the only data sources available that relates prevalence measurements to a comprehensive population denominator. Extrapolating our current study findings to a 34 million people, we estimate that there are a total of 257,138 people alive with CHD in Canada. Indexed to the population by age, we expect that in 2010, there were 166,428 and 90,710 adults and children with CHD respectively in Canada. This constitutes a sharp increase compared to our 2000
findings at which time there were 96,324 CHD adults and 84,868 CHD children for a Canadian population of approximately 31 million inhabitants, reflecting a near 70% increase in the number of adults alive with CHD. The predominantly adult CHD patient population has now been documented to fall squarely within the jurisdiction of adult health care providers.

In the US, the 32nd Bethesda Conference estimated that in year 2000, the total number of adults living with CHD in the United States was 800,000.26 Consistent with these numbers, extrapolated Canadian data to the US for the year 2000 estimated that for a country population of an approximate 281 million inhabitants, there were 855,334 adults and 859,573 children with CHD.15 Although, there are no available published estimates for prevalence rates of CHD beyond birth for 2010 for comparative purposes in the United States, our data are in line with the expected increase of the CHD population over time.2 Our findings suggest the adult population of CHD patients in the US is approximately 1.5 million and underscores the need for data supporting more robust US based population estimates.

Our study should be interpreted in light of the study design. To minimize misclassification bias arising from the use of administrative databases, we maximized accuracy of CHD diagnosis by developing algorithms that cross-referenced diagnoses and procedures from multiple sources (previously described), and by validation through a systematic manual review process applied to randomly chosen samples of 31% of the original 1983-2000 group of patients.3 Even so, some degree of misclassification may be present particularly with milder forms of CHD such as mitral valve anomalies diagnosed in adulthood. To optimize generalizability, we used the entire population of Quebec, Canada’s second largest province, as our study population and we compared the CHD prevalence in infants and adults from this study with those published in the US. The CHD birth prevalence in this study was in line with
estimates of the CHD birth prevalence in the US.\textsuperscript{1} Although our methodology was different from that of the CDC in that we did not include stillborn infants and our data collection methods differed, the fact that the reported rate in Quebec is similar to the CDC rates in the US is reassuring. To minimize ascertainment bias arising from longer observation periods using administrative data sources, we used Bayesian latent class methodology that accounts for uncertainty in the data about the true CHD status of a subject. This is particularly important for milder forms of CHD where detection can increase with enhanced imaging even in the absence of symptoms over the course of a life-time. In contrast, all severe forms of CHD had prevalence rates that are lower in adults than in children. We reported a prevalence of severe CHD in children of 1.7/1,000. This is not dissimilar to the prevalence of CHD requiring cardiac catheterization in the first year of life of 1.5/1,000 reported by Fyler\textsuperscript{28} over three decades ago. This reassures us that detection of severe disease in children is robust and illustrates the spectrum of factors that complicate empirical measurement of CHD across a patient’s life. These include disease severity, ascertainment technique, age at diagnosis and cohort effect. Thus in this analysis we have applied a methodology that highlights the difference between ‘CHD diagnosis’ as can be measured empirically in large population-based data sources and ‘CHD prevalence’ which is our best estimate of true disease status. Misinformation bias may occur due to migration rates; emigration and immigration could lead to over or under estimation of prevalence rates respectively, by modifying the denominator. However, the Quebec population is relatively stable. In 2010, there was a net migration into the province of 44,215 people on a population of 7.9 million people\textsuperscript{29} resulting in a possible error of 0.6% in our prevalence estimate. This would account for misclassification of a maximum of 645 of the 107,559 subjects with CHD diagnosis codes (Figure 1). Finally, the accuracy of our prevalence estimates depends on the extent to
which the source population for the Quebec health system reflects the total Quebec population.

While Quebecers may theoretically seek care outside of the province, there are few incentives for doing so because this would involve out-of-pocket costs for services that are covered by universal health insurance in Quebec. Furthermore, even if some patients did elect to pay for procedures performed outside Quebec, subjects would still be captured in our database if they had any health care encounter in Quebec with a CHD diagnosis at any time during our long observation period.

Notwithstanding these limitations, this is one of the only data sources available, enabling contemporary estimates for the number of people alive with CHD across the life-span in the general population. By 2010, adults accounted for nearly two-thirds of patients with severe and other forms of CHD. These findings highlight an emerging public health issue. There is no longer any doubt that care of well over half of the CHD population falls squarely in the arena of adult medicine. Our study is expected to inform policy makers both in terms of the organization of health services delivery and workforce allocation to meet the needs of this population whose demographics have changed. Our findings underscore the need for longitudinal data sources in the US that will enable more granular calculations of the CHD population across the lifespan and more contemporary population-based estimates in jurisdictions where such data can be made available.

Acknowledgments: The authors would like to acknowledge Matthew Oster MD MPH, Suzanne Gilboa PhD, James Kucik PhD, and Owen Devine PhD from the Centers for Disease Control and Prevention for informing our efforts to incorporate uncertainty around our prevalence estimates. We would like to acknowledge Amy Verstappen, immediate past-President of the Adult Congenital Heart Association for her ongoing advocacy of our need for data. We acknowledge Dr. Louise Pilote for supporting our work and her initial contributions towards the creation of the
Quebec Congenital Heart Disease database. Finally, we would like to acknowledge the staff at the Régie de l’assurance maladie du Québec for their helpfulness in assisting us with the laborious and meticulous extractions required for the reliable raw data to be sent to the MAUDE Unit for the creation of the Quebec CHD database.

**Conflict of Interest Disclosures:** Dr. Ariane Marelli received peer-reviewed public funding that supports her research from the Canadian Institute of Health Research, the Heart and Stroke Foundation of Canada and the Fonds de la recherche en santé du Quebec. Dr. Marelli is a research scholar of the Fonds de la recherche en santé du Québec.

**References:**


Table 1. Prevalence of severe and other congenital heart disease in Québec in the year 2010

<table>
<thead>
<tr>
<th>Severe lesions</th>
<th>Children alive in 2010 in QC, Canada</th>
<th>Adults alive in 2010 in QC, Canada</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (95%CI)</td>
<td>Prevalence per 1000 (95%CI)</td>
</tr>
<tr>
<td>All congenital heart lesions</td>
<td>20,011 (18,973, 21,079)</td>
<td>13.11 (12.43, 13.81)</td>
</tr>
<tr>
<td></td>
<td>39,051 (36,307, 41,922)</td>
<td>6.12 (5.69, 6.57)</td>
</tr>
<tr>
<td>Tetralogy of Fallot or Truncus arteriosus</td>
<td>901 (863, 939)</td>
<td>0.59 (0.57, 0.61)</td>
</tr>
<tr>
<td>Transposition complex</td>
<td>595 (580, 626)</td>
<td>0.39 (0.38, 0.41)</td>
</tr>
<tr>
<td>Endocardial cushion defect</td>
<td>809 (763, 840)</td>
<td>0.53 (0.50, 0.55)</td>
</tr>
<tr>
<td>Univentricular heart</td>
<td>229 (214, 229)</td>
<td>0.15 (0.14, 0.15)</td>
</tr>
<tr>
<td>Hypoplastic left heart syndrome</td>
<td>137 (137, 153)</td>
<td>0.09 (0.09, 0.10)</td>
</tr>
<tr>
<td>All Severe lesions</td>
<td>2,686 (2,564, 2,809)†</td>
<td>1.76 (1.68, 1.84)</td>
</tr>
<tr>
<td>Other lesions</td>
<td>7,464 (7,006, 7,953)</td>
<td>4.89 (4.59, 5.21)</td>
</tr>
<tr>
<td>Atrial septal defect</td>
<td>5,114 (4,793, 5,449)</td>
<td>3.35 (3.14, 3.57)</td>
</tr>
<tr>
<td>Patent ductus arteriosus</td>
<td>1,328 (1,252, 1,420)</td>
<td>0.87 (0.82, 0.93)</td>
</tr>
<tr>
<td>Unspecified defect of septal closure</td>
<td>122 (107, 122)</td>
<td>0.08 (0.07, 0.08)</td>
</tr>
<tr>
<td>Aortic coarctation</td>
<td>412 (397, 443)</td>
<td>0.27 (0.26, 0.29)</td>
</tr>
<tr>
<td>Congenital aortic stenosis or insufficiency</td>
<td>748 (382, 1,114)</td>
<td>0.49 (0.25, 0.73)</td>
</tr>
<tr>
<td>Anomalies of the pulmonary artery or valve</td>
<td>992 (654, 1,330)</td>
<td>0.65 (0.43, 0.87)</td>
</tr>
<tr>
<td>Congenital mitral or tricuspid valve disease</td>
<td>183 (92, 275)</td>
<td>0.12 (0.06, 0.18)</td>
</tr>
<tr>
<td>Ebstein’s anomaly</td>
<td>31 (15, 46)</td>
<td>0.02 (0.01, 0.03)</td>
</tr>
<tr>
<td>Anomalies of great veins</td>
<td>15 (15, 31)</td>
<td>0.01 (0.01, 0.02)</td>
</tr>
<tr>
<td>Unspecified congenital anomalies</td>
<td>870 (443, 1,282)</td>
<td>0.57 (0.29, 0.84)</td>
</tr>
<tr>
<td>Unknown congenital anomalies</td>
<td>15 (15, 31)</td>
<td>0.01 (0.01, 0.02)</td>
</tr>
<tr>
<td>All other lesions</td>
<td>17,325 (16,287, 18,378)†</td>
<td>11.34 (10.67, 12.04)</td>
</tr>
</tbody>
</table>

DOI: 10.1161/CIRCULATIONAHA.113.008396
Figure Legends:

**Figure 1.** Identification of subjects with a congenital heart disease diagnosis in 6,405,643 adults and 1,523,722 children in Québec from 1983-2010. *49,836 subjects were excluded in accordance with previously published algorithms.*³ This included subjects: never hospitalized for CHD and without CHD specific surgery and with no CHD diagnosis made by a cardiovascular specialist; those whose CHD diagnoses were not made by a cardiovascular specialist or a primary care physician or echocardiographer; those whose CHD specific surgical procedure billed by a non-cardiovascular surgical specialist, without an ICD-9/10 CHD diagnosis. Patients who had more than 15 years of observation with only a single contact with the health care system billed as a severe or unspecified CHD diagnosis where excluded. †The subgroup with ICD-9/10 diagnoses for CHD (n=105,235) and the subgroup with CHD surgery codes (n=19,124) are not mutually exclusive: 16,800 subjects had both ICD-9/10 and surgery codes for CHD. The sum of the patients with ICD-9/10 diagnosis of CHD and those with a CHD diagnosis based on surgical codes minus those who had both constitutes the number of patients with a final CHD diagnosis in Québec between 1983 and 2010.

**Figure 2.** The life-time prevalence of congenital heart disease in children and adults in Québec in 2010.

**Figure 3.** The numbers and proportions of adult and children in Québec with all (Panel A) and severe (Panel B) CHD over time in 2000, 2005 and 2010. CHD= congenital heart disease.
Figure 4. Change in CHD prevalence of children and adults in Québec from 2000, 2005 and 2010 for patients with severe CHD. CHD= congenital heart disease

Figure 5. Change in prevalence ratios of CHD in Québec from 2000 to 2010 for all (Panel A) and severe (Panel B) CHD stratified by age. CHD= congenital heart disease
Figure 1

Subjects with ICD-9/10 congenital heart disease codes from 1983 to 2010
N=157,395

Subjects who were excluded according to our published algorithm
N=49,836*

Subjects with a diagnosis of CHD based on ICD-9/10 codes according to our published algorithm
N=105,235

Subjects with a diagnosis of CHD based on surgery codes according to our published algorithm
N=19,124

Subjects with CHD diagnosis or procedure codes from 1983 to 2010
N=107,559†
In 2010, the prevalence was 13.11, 95% CI = (12.43, 13.81) per 1000 children, and 6.12, 95% CI = (5.69, 6.57) per 1000 adults.
Figure 3A
Figure 3B
Figure 4

Prevalence of severe CHD per 1,000 population

PR 2000 vs 2010 = 1.19, 95% CI 1.17 – 1.21

PR 2000 vs 2010 = 1.55, 95% CI 1.51 – 1.62

Children
Adults
Figure 5A

Prevalence of overall CHD per 1,000 population

Age in years
PR (95%CI)

- Age 1-12: 13.4 (1.04, 1.20)
- Age 13-17: 10.0 (0.97, 1.18)
- Age 18-25: 10.7 (1.09, 1.23)
- Age 26-40: 7.02 (1.83, 2.00)
- Age 41+: 5.45 (1.58, 1.75)
Figure 5B
Lifetime Prevalence of Congenital Heart Disease in the General Population from 2000 to 2010
Ariane J. Marelli, Raluca Ionescu-Ittu, Andrew S. Mackie, Liming Guo, Nandini Dendukuri and Mohammed Kaouache

Circulation. published online June 18, 2014;
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2014 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/early/2014/06/18/CIRCULATIONAHA.113.008396

Data Supplement (unedited) at:
http://circ.ahajournals.org/content/suppl/2014/06/18/CIRCULATIONAHA.113.008396.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/
SUPPLEMENTAL MATERIAL

Statistical Appendix

A.1 Derivation of prior distributions

Based on our previous work we reasonably assumed that misclassification errors had been minimized in the CHD database up to the year 2000 since this data had been validated through manual audits independently adjudicated by 2 CHD experts. This was performed on 17,474 of the original 61,386 subjects in the original data over the course of 3 years with adjustment of the algorithms after each set of manual audits as previously published. We therefore sought a statistical method that would allow us to incorporate reasonable uncertainty around our estimates of CHD prevalence using validated prior information. Hence our choice of Bayesian methodology as rationalized in the Statistical Methods section of the manuscript.

Data over the first 14 years of follow up (from January 1, 1983 to December 31, 1996) was thus used to obtain an estimate of the sensitivity of procedure codes. This sensitivity was obtained as the number of patients alive as of midyear 1996 and who had a procedure code from January 1, 1983 to December 31, 1996 divided by the total number of patients alive as of midyear 1996 and who were in our CHD database. The 95% confidence interval for the sensitivity estimate obtained from the above-stated period was used to define the lower and upper bounds of the prior distribution for the sensitivity of procedure codes.

Our prior knowledge of the surgical procedure codes suggests the specificity was as close to 100% as possible.

For the remaining three parameters (sensitivity and specificity for the ICD-9 code and the prevalence), we used non-informative uniform prior distributions on the interval (0, 1), which means that a priori these parameters can take any value between 0 and 1.

A.2 Likelihood contributions for latent class models

The cross tabulation of the results of the ICD-9 code and the surgical procedure code can be presented in a 2x2 table as described in Table A.1.

<table>
<thead>
<tr>
<th>Table A1</th>
<th>Surgical code</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
</tr>
<tr>
<td>ICD-9</td>
<td>Y1 (p1)</td>
</tr>
<tr>
<td>code</td>
<td></td>
</tr>
<tr>
<td>p1+p2+p3+p4=1</td>
<td>Y3 (p3)</td>
</tr>
</tbody>
</table>

The likelihood based on the 2x2 table is then given by the multinomial distribution with cell probabilities p1, p2, p3 and p4.

Since there is no perfect measure of CHD prevalence, a latent class model is constructed. For a given latent class model, we denote by \( \pi \) the prevalence, by \( S_1 \) and \( C_1 \) the sensitivity and specificity, respectively, of the ICD-9 code, by \( S_2 \) and \( C_2 \) the sensitivity and specificity, respectively, of the surgical code for that latent variable. The likelihood contributions for a given latent class model are listed in the table below. In each line of the table we have the probability...
of all possible combinations of observed and latent data for 2 methods of ascertainment.

<table>
<thead>
<tr>
<th>No. of subjects</th>
<th>True CHD status</th>
<th>ICD-9 code</th>
<th>Surgical code</th>
<th>Likelihood contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>(\pi S_1 S_2)</td>
</tr>
<tr>
<td>X2</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>(\pi S_1 (1 - S_2))</td>
</tr>
<tr>
<td>X3</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>(\pi (1 - S_1) S_2)</td>
</tr>
<tr>
<td>X4</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>(\pi (1 - S_1) (1 - S_2))</td>
</tr>
<tr>
<td>Y1-X1</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>((1 - \pi) (1 - C_1) (1 - C_2))</td>
</tr>
<tr>
<td>Y2-X2</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>((1 - \pi) (1 - C_1) C_2)</td>
</tr>
<tr>
<td>Y3-X3</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>((1 - \pi) C_1 (1 - C_2))</td>
</tr>
<tr>
<td>Y4-X4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>((1 - \pi) C_1 C_2)</td>
</tr>
</tbody>
</table>

The probability of each cell in the 2x2 table is then the sum of two cases, i.e. true positive and true negative. For instance, \(p_1\), the probability of the two results being positive, is the sum of \(\pi S_1 S_2\) and \((1 - \pi) (1 - C_1) (1 - C_2)\).

### A.3 WinBUGS program

The Bayesian latent class model was estimated using the WinBUGS code below. We provide as an example the program for the model for estimating the prevalence in adults in 2010.

```{r}
y[1:K] ~ dmulti(p[1:K], n) # y is the number of patients in each cell of the two-by-two table  
# between surgical procedure codes and ICD-9 codes. K is the 
# number of cells in the two-by-two table. n is the total sample size.

p[1]<- pi*(S1*S2)+(1-pi)*((1-C1)*(1-C2))  
p[2]<- pi*(S1*(1-S2))+(1-pi)*((1-C1)*C2)  
p[3]<- pi*((1-S1)*S2)+(1-pi)*(C1*(1-C2))  
p[4]<- pi*((1-S1)*(1-S2))+(1-pi)*(C1*C2)

pi ~ dbeta(1, 1) # non-informative prior for CHD prevalence  
S1 ~ dbeta(1, 1) # non-informative prior for sensitivity of the ICD-9 codes  
C1 ~ dbeta(1, 1) # non-informative prior for specificity of the ICD-9 codes  
S2 ~ dunif(0.2404, 0.2768) # Prior information on sensitivity of procedures  
C2 <- 1 # Prior information on specificity of procedures.  
# Procedures assumed to be perfectly specific  
prev <- 1000*pi

```

# adults all list(n=6380957; K=4, y=c(6739, 41578, 3325, 6329315))
A.4 Methodology related to WinBUGS

The details of implementing WinBUGS and checking for convergence are as follows. We used five MCMC (Markov chain Monte Carlo) chains. For each chain, 20,000 iterations were used to estimate the parameters after 2,000 burn-in iterations. Convergence and stationary of the Markov Chains were assessed by visually inspecting the history plots for each parameter, and also using the Brooks-Gelman-Rubin method provided within WinBUGS. The point estimates of the parameters consisted of the median of the posterior samples and the 95% credible intervals were obtained by taking the 2.5\textsuperscript{th} and the 97.5\textsuperscript{th} percentiles.