Received Date: 28/06/2016

Revised Date: 29/08/2016

Accepted Date: 13/09/2016

Article Type: Contemporary Reviews

Subject category: Contemporary Reviews

Databases for Congenital Heart Defect Public Health Studies Across the Lifespan

Short Title: Databases for Congenital Heart Defects

Tiffany J. Riehle-Colarusso MD, MPH[1], Lisa Bergersen MD, MPH[2], Craig S. Broberg MD, MCR[3], Cynthia H. Cassell PhD[1], Darryl T. Gray MD, ScD[4], Scott D. Grosse PhD[5], Jeffrey P. Jacobs MD[6,7], Marshall L, Jacobs MD[6,7], Russell S. Kirby PhD, MS[8], Lazaros Kochilas MD, MSCR[9], Asha Krishnaswamy BEE, MS[1], Arianne Marelli MD, MPH[10], Sara K. Pasquali, MD, MHS[11], Thalia Wood MPH[12], Matthew E. Oster MD, MPH[1,9], for the Congenital Heart Public Health Consortium[*]

[1] Division of Congenital and Developmental Disorders, National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA

[2] Harvard Medical School, Department of Cardiology, Children's Hospital of Boston, Boston, MA

[3] Adult Congenital Heart Program, Knight Cardiovascular Institute, Oregon Health and Sciences University, Portland, OR

[4] Center for Quality Improvement and Patient Safety, Agency for Healthcare Research and Quality, Rockville, MD

[5] Office of the Director, National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA

[6] Division of Cardiovascular Surgery, Department of Surgery, Johns Hopkins All Children's Heart Institute, Johns Hopkins All Children's Hospital and Florida Hospital for Children, Saint. Petersburg, Tampa, and Orlando, FL
[7] Division of Cardiac Surgery, Department of Surgery, Johns Hopkins University, Baltimore, MD
[8] Department of Community and Family Health, College of Public Health, University of South Florida, Tampa, FL
[9] Children's Healthcare of Atlanta, Emory University School of Medicine, Atlanta, GA
[10] McGill Adult Unit for Congenital Heart Disease, Montreal, Québec
[11] Department of Pediatrics and Communicable Diseases, University of Michigan C.S. Mott Children's Hospital, Ann Arbor, MI
[12] Association of Public Health Laboratories, Silver Springs, MD
* Individual participants Congenital Heart Public Health Consortium have been listed in an Appendix at the end of the article.

Corresponding Author: Dr. Tiffany Riehle-Colarusso, Centers for Disease Control and Prevention, 4770 Buford Highway, Mailstop E-86, Atlanta, Georgia, 30341. E-mail: tcolarusso@cdc.gov; Phone: 404-498-3805, Fax: 404-409-3040

Key words: congenital heart defects, public health science, databases

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention, the Agency for Healthcare Research and Quality, or the Department of Health and Human Services.

INTRODUCTION

In a 2012 meeting at the Centers for Disease Control and Prevention (CDC), key experts and stakeholders identified public health knowledge gaps about congenital heart defects (CHDs), namely: prevalence of CHDs across the lifespan, long-term outcomes of persons with CHDs, and health services delivery for persons with CHDs.[1] These gaps, and strategies to address them, formed the basis of a CHD public health science agenda. The strategies included leveraging information in existing databases to examine the epidemiology, health outcomes, and health service utilization of the CHD population.[1] Many databases with CHD data exist and are managed by hospitals, specialty organizations, partnerships, public health, and other governmental entities. Researchers may be familiar with some databases but not others. Anyone planning studies to address public health knowledge gaps may benefit from an understanding of this complex constellation of databases.

The Congenital Heart Public Health Consortium (CHPHC) was formed in 2009 as a collaboration of stakeholders with the mission to prevent CHDs and improve outcomes for affected individuals.[2] The CHPHC created a database workgroup to increase awareness of opportunities to contribute to the public health science agenda for CHDs using existing databases. The workgroup, consisting of experts in various disciplines (cardiologists, surgeons, epidemiologists, health service researchers), identified databases located in Canada or the United States (U.S.) with information on CHDs from 1990 onward. The goals of this paper are to provide an overview of database types and list examples of databases which may be used to address CHD public health knowledge gaps. IRB approval was not deemed necessary for this review.

Database characteristics which may be important to consider when designing a study to address CHD public health knowledge gaps can be grouped into 3 main areas: a) population included, b) data content, and c) accessibility. The first area relates to aspects such as sample size, inclusion criteria, whether the database is population-based, and whether persons are followed for a period of time. The second relates to what variables are included (e.g, type and amount of clinical detail, information on resource utilization, or financial information), data collection mechanisms and coding, and data timeliness, accuracy, and completeness. The last area involves obtaining access to use the data, which may be costly, time-consuming, or restricted, and will vary depending on the database selected.

Using existing data is often more cost-effective and reasonable than gathering new data; however, research is limited to the data which are available and there is often no perfect dataset to answer a particular question. Features of particular databases vary in importance, depending on the research question. One database's strength in answering a question may be a limitation for another question. For example, a database may be population-based but have limited clinical detail. This database may be good for an overall prevalence estimate, but not as useful for analyzing treatment outcomes of a particular CHD phenotype. It is the role of the researcher to determine which characteristics are most important and find the appropriate database that will best inform the particular research question. This paper does not comment on the strengths or weaknesses of specific databases, but rather, presents general information and additional resources. Researchers may use this information to help determine the utility of existing databases for their particular CHD public health study.

DATABASE CATEGORIES AND EXAMPLES

We grouped examples of databases into categories based on type of data source (administrative healthcare, birth defect surveillance, clinical, survey, and vital records). We briefly describe each category below, with discussion of strengths and limitations to consider when addressing public health knowledge gaps. We also determined whether identified example databases had individuals with only CHDs (cardiac-specific databases) or had individuals with many conditions, including CHDs (general databases). Examples of cardiac-specific and general databases in each of these categories are listed in Tables 1-4. Some databases have more than one type of data sources and are therefore listed in Table 5 under a separate combined category heading (e.g., Administrative and Clinical). The tables provide a brief description of the database, sponsoring organization, years of data, and a URL link for further information. An asterisk denotes cardiac-specific databases. While basic information is provided on a variety of databases, researchers are encouraged to contact database hosts for further information to assess their utility. Also, because databases are constantly evolving, other databases not captured in these tables may be useful in addressing a particular question.

Administrative healthcare databases

Administrative healthcare databases are generally developed from facility records or health insurance claims for billing purposes and/or to document health care provided; they are typically not designed for research purposes. Most are not specific to CHDs, but still useful for research and public health investigations related to CHDs. We identified thirteen administrative healthcare databases (one of which is cardiac-specific) (Table 1), two administrative/clinical databases, and four administrative/survey databases (Table 5).

Facility-based administrative healthcare databases include all patients at a certain institution, regardless of payer, and may be able to identify a person over multiple encounters.

However, these databases do not have data on outside resources utilized by that individual. Facility-based databases usually include the nominal charges for the services provided, although the provision of hospital or aggregated department-specific cost-to-charge ratios allow the estimation of facility-perspective costs.[3] On the other hand, claims-derived administrative healthcare databases cover healthcare use by all enrollees in certain health plans, regardless of where the care is received, and can follow individuals for as long as they are plan beneficiaries. Claims databases typically include millions of enrollees, and by definition do not include non-enrollees and the uninsured. These excluded groups may be needed in a study, depending on the particular public health issue being addressed. Claims-based databases capture billed charges and actual payments made, including payments made by health plans and enrollees.

In general, administrative healthcare databases can provide large sample sizes, detailed resource utilization and financial information, and are often population-based to the extent they capture all patients in a geographic area or health plan. However, some persons may not use the healthcare system; thus, administrative healthcare databases may either over-represent sicker patients or exclude those without access to care. Another limitation of U.S. administrative healthcare databases is how data are coded. Typically, these databases use *International Classification of Disease version 9 or 10 Clinical Modification* (ICD-9-CM, ICD-10-CM) codes which often lack sufficient detail to adequately characterize specific CHD phenotypes or procedures. Hence, researchers may be limited to investigating broad classes of CHDs or procedures. Administrative databases may also be difficult to access, due to restrictions and license fees, and to use, due to their size and need for strong programmers or computational power.[3]

One example of multi-institutional facility-based databases is the Healthcare Cost and Utilization Project (HCUP) databases developed and managed by the Agency for Healthcare Research and Quality (AHRQ) through a public-private partnership. The cornerstone of HCUP is facility-level inpatient and hospital outpatient discharge data that include diagnoses and procedure codes, admission source, discharge status, patient demographics, expected payment source, total billed hospital charges, estimated costs, length of stay and specific hospital characteristics. Hospitals provide these data on all patients, including self-pay and uninsured patients, to state-level entities that create state-specific hospital discharge databases. Under Memoranda of Agreements these entities voluntarily share with AHRQ their files, which become part of HCUP. For 2013, the most current data year available, 48 states(accounting for 97% of the U.S. population) participated in HCUP.[4] The states decide which data elements are included in standardized State Inpatient Databases (SID) and whether AHRQ can release their files directly to users. For 2013, SID files for 28 states were available directly from AHRQ; files for the remaining states can potentially be obtained from the state-level organization.[4] Nationally representative databases based on aggregated SID data include the annual Nationwide/National Inpatient Sample (NIS) and the triennial Kids' Inpatient Sample (KID). Other HCUP databases which capture CHD care are listed in Table 1. Copies of the HCUP databases can be purchased; aggregated data from select HCUP databases are freely available on-line at the HCUPnet site (http://hcupnet.ahrq.gov). Several health service research studies have used HCUP data to assess data on incidence, outcomes, facility costs, and factors related to hospitalization for individuals with CHDs.[5-10]

Health insurance claims databases include public insurers and proprietary insurance databases, such as Truven Health's MarketScan® suite of databases. The MarketScan® research databases include commercial databases of employer-sponsored insurance, a Medicare database, and a Medicaid database representing claims from anonymized states who contract with Truven. MarketScan® data from 2005 were used to estimate health care use and costs for children with CHDs.[5] Over thirty states have created, or are in the process of creating, all-payer claims databases (APCD) that combine claims from within their state from private and public payers.[11, 12] Some states have APCD data available on request, which could be useful in assessing resource utilization and healthcare costs for persons with CHD, as well as surveillance of those with CHDs.

AHRQ has tools which states can use to improve quality of care for vulnerable populations. To help researchers answer specific health service questions, lists of databases with results for quality measures and databases from which measures could be calculated are available on-line (http://nhqrnet.ahrq.gov/inhqrdr/resources). This detailed compendium has information on over 100 databases and websites, including several listed in this paper (e.g., MarketScan®, HCUP, state APCDs, and Medical Expenditure Panel Survey (MEPS)), which can guide researchers to appropriate databases for a particular study question about CHDs.

Birth defects surveillance

Surveillance of infants with birth defects is a core public health activity. While the United States has no national birth defect surveillance system, most states maintain their own surveillance programs, which can vary by which entity conducts the surveillance (e.g., health department), objectives, case ascertainment method, age of children included, or defects included. Surveillance data can be used for epidemiologic investigations[13-15] or health services research.[16-18] We presented three examples of Birth Defect Surveillance databases (Table 2) and three in the combined category of Birth Defects Surveillance/Survey (Table 5). It is beyond the scope of this paper to list all birth defect programs. However, a list of programs with links can be found at the National Birth Defect Prevention Network (NBDPN) website (http://www.nbdpn.org/state_programs_and_related_lin.php). Researchers should contact specific birth defect surveillance programs to explore opportunities to analyze the state's data.

The strengths of birth defects surveillance databases are that they usually include a comprehensive, population-based birth cohort of infants with birth defects. The NBDPN was formed to addresses issues of surveillance, research, and prevention among U.S. birth defect programs.[19] The NBDPN has created surveillance guidelines to help standardize data collection.[19] Recently, the NBDPN developed data quality measures and tri-level performance criteria, focused on data completeness, timeliness, and accuracy, to assess strengths and weaknesses of programs.[20] This information will be used to develop and implement national data quality standards for birth defects surveillance. Many programs also use chart review to validate diagnoses, obtain data from several data sources, or use modified ICD-9-CM or ICD-10-CM codes which are more specific for birth defects; thus, the data quality may be quite high. However, surveillance databases have varied methodology, rarely have resource utilization or financial details unless linked to databases with that information, and usually do not have detailed clinical data on treatment course, unless it related to the diagnosis of the CHD. Furthermore, due to Health Insurance Portability and Accountability Act (HIPAA) regulations, access to identifiable data is restricted and governed by the birth defects program. Birth defects surveillance databases, unless linked to other databases to provide information beyond infancy, are not longitudinal. Although birth defects surveillance databases may not be able to address some clinical or outcomes questions, their strengths provide important information on the birth prevalence of CHDs.

One of the oldest birth defects surveillance programs is the Metropolitan Atlanta Congenital Defects Program (MACDP), maintained by the Centers for Disease Control and Prevention. Begun in 1967, MACDP collects information on birth defects in infants and children up to six years of age who were born to mothers residing in select metropolitan Atlanta counties.[21] Cases are identified by trained abstractors who actively search newborn hospitals, pediatric hospitals, and other clinical sources, and cases are linked to vital records from the Georgia Department of Public Health. Records are reviewed and those with a CHD diagnostic code are classified by physicians trained in pediatric cardiology, using standard clinical nomenclature derived from the Society of Thoracic Surgeons Congenital Heart Surgery Database (STS-CHSD).[22] MACDP data on CHDs have been extensively analyzed, resulting in publications on trends in prevalence and survival,[13, 15, 23] risk factors for CHDs,[24] and a comparison of administrative and clinical coding for CHDs.[25]

Birth defects surveillance programs monitor the CHD occurrence in their jurisdiction and contribute to CHD epidemiology. However, given the rarity of birth defects, there is often insufficient data in any one state to address some public health questions. The NBDPN also publishes pooled data from participating programs; in the 2012 annual report critical CHD surveillance data were highlighted[26] and the public health role in newborn screening for critical CHDs was discussed.[27] There is also a data repository with data submitted by several states for infants with birth defects born 1999-2007, which has been used to study the association of race/ethnicity with birth defects,[28] the survival of infants born with birth defects,[29, 30] and may be used to study other issues related to CHDs.

Clinical CHD databases or registries

Many databases with clinical information on persons with CHDs exist, including single and multi-institutional databases as well as specialty care registries and research datasets. These databases vary in years of data collected, type of data, inclusion criteria, and purposes for utility. Research datasets may have uniquely different characteristics from clinical registries. Many clinical databases are designed to track patient outcomes, to improve quality of care or for care benchmarking. However, since the early years of pediatric cardiac interventions, it was recognized that the experience of any one institution was limited, and collaboration between centers was necessary to have sufficient numbers to conduct meaningful outcomes analyses. In this paper, we grouped examples of multi-institutional clinical datasets, specialty care registries, and research datasets in the "clinical" category. We identified 15 databases sourced primarily from clinical practice (13 cardiac-specific ones) (Table 3), and two administrative/clinical databases, sourced from a combination of large administrative healthcare databases combined with clinical practice data (Table 5).

The strength of clinical databases to address public health knowledge gaps lies in their detailed information on diagnosis, treatment, and clinical outcomes. Multi-institutional clinical databases usually amass a large sample size over time, with diversity in CHD phenotypes, patient characteristics and geographic representation. Furthermore, clinical databases often use standard nomenclature and outcome measures, although the implementation of these standards may be inconsistent within or across institutions or databases, as recently documented.[31] Clinical databases may also have information on comorbidities and non-cardiac events, which is especially important for the older population. Clinical databases are useful, for example, when evaluating how clinical factors such as treatment or hospital course might influence the long-term outcomes of persons with a particular CHD phenotype. However, clinical databases may include only certain cohorts (e.g., only persons with a specific diagnosis or undergoing a certain type of interventions), with little or no longitudinal follow-up of only limited outcome variables, may not be representative of the study population, and may not include resource utilization or financial data. Accessing the data may also require special approval or fee-for-access. These limitations may be important if a researcher is interested in an entire population or patient characteristics which may not be consistently captured in clinical data (e.g., birth information).

Efforts are ongoing to enhance and improve clinical databases for CHDs. The Multi-Societal Database Committee for Pediatric and Congenital Heart Disease was established in 2005 to provide infrastructure for collaboration between healthcare professionals interested in the outcomes of persons with CHDs.[32] This committee is working to collaborate on use of common nomenclature, uniform core dataset information, evaluation of case complexity, developing a mechanism for verifying case completeness and accuracy, and standardization of protocols for longitudinal follow-up of persons with CHDs.[32] The outputs from this committee could help address not only questions related to treatment outcomes, but public health questions as well.

One example of a large clinical database with geographical and diagnostic diversity is the

STS-CHSD, founded in 1994 to support quality improvement in cardiothoracic surgery.[33] As of December 31, 2015, STS-CHSD contains 394,980 operations reported from 124 pediatric and congenital heart surgery hospitals in the United States and three centers in Canada. With penetrance of over 95% in the U.S., the data in STS-CHSD are representative of all United States pediatric and congenital heart surgeries.[33] Definitions of all terms and codes used in the STS-CHSD have been standardized and published, including the use of the International Pediatric and Congenital Cardiac Code (IPCCC).[34] The STS-CHSD employs data quality measures and produces regular reports to better understand outcomes, provide benchmarks, and improve quality of care.[32, 33, 35-38] Data from the STS-CHSD have also helped fill public health knowledge gaps. Application of the STS-CHSD nomenclature improved the quality of surveillance data[22] for subsequent population-based analyses, e.g., prevalence trends in CHDs,[13] CHD survival,[14, 15] and receipt of special education by those with CHD.[39] As with other clinical databases, aspects of the STS-CHSD may limit its utility to answer some public health questions (e.g., access to care).

Surveys

In surveys, individuals are usually sampled from a defined population and queried using a structured instrument (e.g., telephone questionnaire) to generate information on a representative sample with respect to a target population of interest (e.g., children <18 years of age). Data can be used to profile key issues in the population of individuals with CHDs to help set priorities for healthcare policy, develop programs, and improve services. The utility of survey data for answering CHD public health questions varies, depending on the survey design, sample composition and size, timeframe, and topics or questions included. In general, surveys that include persons with CHDs may be large overall (i.e. a nationally representative sample), but may have a small number of total or specific CHD phenotypes, which may limit utility of the database. We identified several examples of databases with survey information which may be useful in public health studies of CHDs: five general survey databases (Table 4), four administrative/survey databases (Table 5), and three birth defect surveillance/survey databases (Table 5).

A strength of the identified surveys is that they ask the person or their proxy (e.g., a parent) about a broad range of topics relevant to public health (i.e., medical and non-medical exposures, resource utilization, demographics, socio-economic data, care coordination, continuity of care, barriers to care). Data important for understanding public health aspects of CHDs, such as self-reported information on quality of life or pregnancy exposures, may be available in survey data and not in other types of data sources. However, survey information is self-reported, often retrospective, and may have varying degrees of validity and recall bias. Data from surveys are typically cross-sectional – providing information about the population at one point in time – which may limit generalizability of research findings. Surveys typically lack identifiers that could otherwise be used for linking with other databases. While the survey may be conducted repeatedly, it is usually on a different sample each time as very few surveys re-contact participants to obtain longitudinal data.

Two main sources of national population-based data are the Decennial Census and the American Community Survey (ACS). The Decennial Census has been conducted since 1790 as required by the U.S. Constitution. Most households receive a short questionnaire and prior to 2010 one in every six households received a more detailed long questionnaire on socioeconomics. After 2000, the Census Bureau redesigned the census and the socioeconomic questionnaire became the ACS. The ACS surveys households monthly and provides yearly information to communities in 1-, 3-, and 5-year reports.[40] Data and tools to use this data from these surveys are publically available. The Census and ACS can be useful denominator and comparison data in studies of CHD population. Furthermore, these data can be linked to other databases to study community level factors influencing health and outcomes of persons with CHDs.

The National Survey of Children with Special Health Care Needs (NS-CSHCN) is a telephone survey sponsored by the federal Maternal and Child Health Bureau, designed to periodically sample the U.S. population to identify children < 18 years of age with special healthcare needs.[41] Telephone numbers were randomly dialed to identify households with one or more children < 18 years of age. Trained interviewers asked the parent or guardian questions to identify all children in the household with special healthcare needs. It was administered three times between 2001 and 2010. In the 2009-2010 survey, CHDs were a specific condition prompt. Topics covered include child's health and functional status, insurance coverage, access to healthcare, care coordination, and impact of health conditions on the child and their family.[41] The survey is being integrated into the National Survey of Children's Health, but will still provide the same in-depth look at the lives of children with special health care needs. Survey strengths included that it was population-based and provided publically available comparison datasets. It described the population of CSHCN and provided a snapshot of the impact of special healthcare needs. However, CHDs and treatment are not confirmed by a medical record source.

Vital records

The U.S. vital records system is a federal-state partnership in which state vital records agencies receive federal funds for providing statistical data concerning vital events (live birth, death, and fetal death). Birth and death certificates enumerate all live births and deaths occurring in the United States, providing a comprehensive population-based cohort. Thus, vital records are important in CHD public health studies. Although all states have vital records, data content varies slightly by state. The National Center for Health Statistics (NCHS) has promulgated national standard certificates that define the content and data elements.[42, 43] Researchers should contact the department of health in the particular states of interest to obtain information on available state-specific vital records databases.

Birth and death certificates contain protected personal identifiable information. However, NCHS has national, de-identified, publicly available data files (e.g., birth, death, and period-linked birth-infant death data)[44] useful for public health studies. For example, causes of death information from death certificates was used to described annual CHD mortality in the United States by age, race, and sex.[45] Period-linked birth-death data were used to identify racial differences in infant mortality due to birth defects such as CHDs.[46] The NCHS also maintains the National Death Index (NDI), a restricted-access, centralized database of all state death records.

Although vital records data are useful, there are some limitations to consider. The quality of birth defects reporting on birth and fetal death certificates is generally poor, and thus may influence the quality of a particular study.[47-49] Researchers have identified limitations in ability to identify all decedents with a specific illness or health condition.[50, 51] The coding on birth or death records, or the checkboxes used on many birth/fetal death certificates, may not provide accurate or sufficient diagnostic details for some studies. Furthermore, birth and death certificates may use different coding systems. Death certificates have been coding underlying cause of death using ICD-10 since 1999, well ahead of clinical utilization of ICD-10-CM for billing purposes, which became official as of October 1, 2015. Resource utilization and cost/charge data, are not presently reported in these documents. Finally, due to the personal identifying information, individual-level vital records are not easily accessible to general researchers and often must be linked at the health department or via the NDI.

COMBINING DATABASES ACROSS CATEGORIES

Combining databases can maximize strengths and minimize limitations of individual databases to address issues that may not be possible using a single database (Table 5). For example, by linking data from a clinical database (STS-CHSD) with an administrative database (Pediatric Health Information System (PHIS)), multiple studies on healthcare utilization with robust clinical data have been conducted.[52-54] Leveraging existing databases through linkage is also important to understand long-term and longitudinal outcomes for persons with CHDs. One example is the linkage of the Pediatric Cardiac Care Consortium (PCCC) with national registries. The PCCC contains data on patients who have undergone CHD interventions at 47 U.S. centers between 1982 and 2011, with direct identifiers available for patients enrolled up to April 2003.[55] The availability of direct identifiers allowed linkage of PCCC data with the NDI and the United Network for Shared Organs (UNOS), thereby providing significant information regarding the long term outcomes after palliative or corrective procedures.[56] These linkages may address some of the individual database weaknesses regarding longer-term and longitudinal follow-up. Experts across disciplines agree that there needs to be a better mechanism for longitudinal follow-up of persons with CHDs across the lifespan. Longitudinal data can provide unique outcomes information.[1, 57] Restricted-access data files, such as NDI and the corresponding state-level records, may also be useful for other record-based linkage studies of persons with CHDs. Birth defects surveillance data have been linked to vital records to examine CHD prevalence[13] and survival,[15, 58] and to longitudinal school records to investigate receipt of special education services among children with CHDs.[39] Such population-based estimates are attainable only through linkage of multiple databases.

Throughout this paper we have noted unique databases which span two database categories. However, databases from different categories have also been combined to form new stand-alone databases. One example of a database that spans two categories (i.e., birth defect surveillance and surveys) is the National Birth Defects Prevention Study (NBDPS). The NBDPS is a multi-site collaborative case-control study to evaluate potential genetic and environmental risk factors for major congenital malformations, including CHDs.[59-62] Cases of CHDs are identified from birth defect surveillance data, and structured telephone interviews are conducted with mothers of case and controls. Investigations using NBDPS data have contributed to understanding CHDs, including occurrence risk associated with maternal smoking,[63] obesity,[64] medication use,[65, 66] and descriptive epidemiologic studies of select CHDs.[67, 68] The strength of studies such as the NBDPS is that they are large, population-based, multi-center studies with standardized interview protocol, medical record review and classification of CHDs. However, limitations exist, including potentially inaccurate or biased recall of exposures of interest due to self-report.

CDC recognized the possibilities for research and surveillance through linking data across various sources. In 2012, CDC awarded grants to the New York State Department of Health, Emory University in Atlanta, Georgia, and the Massachusetts Department of Public Health for a pilot study to develop population-based surveillance of adolescents and adults with CHDs. The grantees combined data within their states from a variety of data sources including birth defects surveillance data, Medicaid data, hospital discharge data, vital records, provider reports, and clinic billing data.[69] As results are being analyzed from this pilot, a new collaborative with five sites is expanding on this work.

While examples of specific database combinations exist, a coordinated effort to use data for answering public health questions concerning CHDs is lacking. The consolidation of heterogeneous datasets raises significant challenges related to confidentiality, governance, nomenclature and coding structure, and information technology capabilities. Even efforts at combining multi-institutional electronic health record (EHR) data on CHD have identified many obstacles.[31] For example, there are inherent complexities of database interaction, such as non-standard variable definitions or database structure. Furthermore, data are from disparate populations and different time points across the life span. Some represent a cross-section of the population, while others include only those patients seen in a specific healthcare setting, or at the time of a specific event (such as surgery or cardiac intervention). Procedural datasets include far more clinical detail than administrative sources. The types of coding schemes used for each database varies, as well as the experience of the database manager or healthcare provider that selects the codes, both of which create inherent heterogeneity in the accuracy and granularity of the congenital diagnosis. Variables for accurate linkage between dataset may not be adequate, though this could be assisted through the use of a global unique identifier as has been endorsed by the National Institutes of Health for other groups (https://ndar.nih.gov/tools_guid_tool.html). Furthermore, issues of HIPAA compliance may be raised, since consent for data use in one database may not carry over to a conglomerate. To help address these challenges, in January 2015, the National Heart, Lung, and Blood Institute (NHLBI) convened a workgroup to develop a vision for an integrated data network for CHD research. The subsequent report summarizes the discussions and identifies critical elements as well as potential barriers for integrating CHD data.[57]

CONCLUSION

There are numerous databases available to address public health knowledge gaps about CHDs across the lifespan. Databases can be grouped into broad categories with particular strengths and limitations. Understanding the relative characteristics of different databases is important for choosing the best data to answer a particular research question, or to identify opportunities to maximize strengths and minimize limitations through database linkages.

Participants in the Congenital Heart Public Health Consortium

Ginnie Abarbanell, Faith Adams, Steven Allen, Sydney Allen, Anand Ambrose, Carl Backer, Andrea Baer, Carissa Baker-Smith, Mona Barmash, Amy Basken, Cassandra Bates, Sarosh Batlivala, Robert Beekman, John Belmont, Joshua Benke, Stuart Berger, Lisa Bergersen, JR Bockerstette, Jeffrey Boris, Lorenzo Botto, Jackie Boucher, Craig Broberg, Dana Brock Hageman, Cheryl Brosig Soto, Kristin Burns, Lenore Cameron, Robert Campbell, Steven Colan, Tiffany Colarusso, Lynn Colegrove, Christina Coleman, Angie Colson, Adolfo Correa, Pamela Costa, Chris Couser, Melissa Crenshaw, Tessa Crume, Rachel Daskalov, Mark Del Monte, Lindsay DeSantis, Kaitlin Doherty, Kenneth Dooley, Charles (Wes) Duke, Pirooz Eghtesady, Saiza Elayda, Alison Ellison, Tim Elsner, Cori Erntz, Michelle Esquivel, Bethany Evans Lloyd Feit, Marcia Feldkamp, William Foley, Elyse Foster, Wayne Franklin, Bridget Freeley, Frank Galioto, Mary George, Michael Gewitz, Katja Gist, Thomas Glenn, Melissa (Jill) Glidewell, Lorraine Gore, Darryl Gray, Johanna Gray, Hannah Green, Scott Grosse, Michelle Gurvitz, Sonia Handa, Melissa Harvey, Emilie Heath, Danielle Hile, John Hokanson, Margaret

(Peggy) Honein, Marius Hubbell, Jeff Hudson, Kelly Huhn, Dawn Ilardi, Jeffrey Jacobs, Marshall Jacobs, Dawn Jacobs, Robert Jaquiss, Kathy Jenkins, Anitha, John, Patrick Johnson, Shakila Johnson, Emily Jones, Antonios Jossif, Jonathan Kaltman, David Kasnic, Alex Kemper, Natalie Kenny, Paul Khairy, Valerie King, Russell Kirby, Donna Knapp, Daisuke Kobayashi, Lazaros Kochilas, Adrienne Kovacs, Asha Krishnaswamy, James Kucik, Karen Kuehl, Alexandra Kuznetsov, Scott Leezer, Jodi Lemacks, Patty Libby, Paul Lipkin, Michele Lloyd-Puryear, Keila Lopez, Nicolas Madsen, Cara Mai, Monica Mann, Ariane Marelli, Bradley Marino, Gerard Martin, G. Paul Matherne, Phillip Mauller, Susan May, Nancy McCabe, Edward McCabe, Michelle McCardle, Ty McCathran, Amy McCathran, Michael McConnell, Kristine McCormick Eric Melsom, William Milionis, Paula Miller, Erika Miller Stephanie Mitchell, Cynthia Moore, Laura Morris, Angela Murray, Kathleen Mussatto, Steven Neish, Sue Nelson, Jane Newburger, Jeremy Nicolarsen, Autumn Niggles, Jacqueline Noonan, Gail Ober, Lori O'Keefe, Matthew Oster, Marc Overcash, Jennifer Page, Matthew Park, Sara Pasquali, Mehul, Patel, Jasmin Patel, Gail Pearson, Cindy Pellegrini, Corrie Pierce, Nelangi Pinto, Kara Polen, Jose Quinones, Carol Raimondi, Pat Richter, Michelle Rintamaki, Elisa Robles, Geoffrey Rosenthal, Grahame Rush, Laura Russell, Annamarie Saarinen, Craig Sable, Joel Saltz, Terri Schaefer, Kathryn Schubert, Vida Schwartz, Stuart Shapira, Kathleen Sheehan, Brenda Silverman, Regina Simeone, Juanita Smith, Kristina Smith, Kimberly Smith, Marci Sontag, Shubhika Srivastava, Corrie Stassen, Corey Stiver, Kathryn Taubert, Judy Thibadeau, John Thomas, Dena Thomas, Vivian Thorne, Linda Tiernan, Susan Timmins, Colby Tiner, Natalie Torentinos, Glenn Tringali, James Tweddell, Lisa Vasquez, Amy Verstappen, Janice Ware, Caron Watkins, Catherine Webb, Ellen Weiss, Marina Weiss, Gil Wernovsky, Gretchen Whitehurst, Herbert Whitley, Jennifer Witten, Austin Wong, Thalia Wood, Matthew Wright, Robert Wynbrant, Bistra Zheleva

Disclosures: None

REFERENCES

1. Oster ME, Riehle-Colarusso T, Simeone RM, Gurvitz M, Kaltman JR, McConnell M, Rosenthal GL and Honein MA. Public health science agenda for congenital heart defects: report from a Centers for Disease Control and Prevention experts meeting. *J Am Heart Assoc*. 2013;2:e000256.

2. The Congenital Heart Public Health Consortium. 2011.

3. Riley GF. Administrative and claims records as sources of health care cost data. *Med Care*. 2009;47:S51-5.

4. Agency for Health Research and Quality, HCUP Databases - State Inpatient Databases Overview. 2015.

5. Boulet LS, Grosse SD, Riehle-Colarusso T and Correa A. Health care costs of congenital heart defects. In: D. F. Wyszynski, A. Correa and T. P. Graham, eds. *Congenital Heart Defects From origin to Treatment*: Oxford University Press; 2010: 493-501.

6. Okumura MJ, Campbell AD, Nasr SZ and Davis MM. Inpatient health care use among adult survivors of chronic childhood illnesses in the United States. *Arch Pediatr Adolesc Med*. 2006;160:1054-60.

7. Centers for Disease Control and Prevention. Hospital stays, hospital charges, and in-hospital deaths among infants with selected birth defects – United States, 2003. *MMWR*. 2007;56:25-29.

8. Russo CA and Elixhauser A. Hospitalizations for birth defects, 2004: Statistical brief #24 *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs* Rockville (MD); 2006.

9. Simeone RM, Oster ME, Cassell CH, Armour BS, Gray DT and Honein MA. Pediatric inpatient hospital resource use for congenital heart defects. *Birth Defects Res A Clin Mol Teratol*. 2014;100:934-43.

10. Simeone RM, Oster ME, Hobbs CA, Robbins JM, Thomas Collins R and Honein MA. Factors associated with inpatient hospitalizations among patients aged 1 to 64 years with

congenital heart defects, Arkansas 2006 to 2011. *Birth Defects Res A Clin Mol Teratol*. 2015;103:589-96.

11. Bartoloini E and Paradis R. All Payer Claims Databases: unlocking the potential. *The Network for Excellence in Health Innovation*.

12. Porter J, Love D, Costello A, Peters A and Rudolph B. *All-Payer Claims Database Development Manual: Establishing a Foundation for Health Care Transparency and Informed Decision Making*: APCD Council and West Health Policy Center; 2015.

13. Bjornard K, Riehle-Colarusso T, Gilboa SM and Correa A. Patterns in the prevalence of congenital heart defects, metropolitan Atlanta, 1978 to 2005. *Birth Defects Res A Clin Mol Teratol*. 2013;97:87-94.

14. Miller A, Siffel C, Lu C, Riehle-Colarusso T, Frias JL and Correa A. Long-term survival of infants with atrioventricular septal defects. *J Pediatr*. 2010;156:994-1000.

15. Oster ME, Lee KA, Honein MA, Riehle-Colarusso T, Shin M and Correa A. Temporal trends in survival among infants with critical congenital heart defects. *Pediatrics*. 2013;131:e1502-8.

16. Dawson AL, Cassell CH, Riehle-Colarusso T, Grosse SD, Tanner JP, Kirby RS, Watkins SM, Correia JA and Olney RS. Factors associated with late detection of critical congenital heart disease in newborns. *Pediatrics*. 2013;132:e604-11.

17. Kucik JE, Cassell CH, Alverson CJ, Donohue P, Tanner JP, Minkovitz CS, Correia J, Burke T and Kirby RS. Role of health insurance on the survival of infants with congenital heart defects. *Am J Public Health*. 2014;104:e62-70.

18. Peterson C, Dawson A, Grosse SD, Riehle-Colarusso T, Olney RS, Tanner JP, Kirby RS, Correia JA, Watkins SM and Cassell CH. Hospitalizations, costs, and mortality among infants with critical congenital heart disease: how important is timely detection? *Birth Defects Res A Clin Mol Teratol*. 2013;97:664-72.

19. Guidelines for Conducting Birth Defects Surveillance - National Birth Defects Prevention Network. 2004.

20. Anderka M, Mai CT, Romitti PA, Copeland G, Isenburg J, Feldkamp ML, Krikov S, Rickard R, Olney RS, Canfield MA, Stanton C, Mosley B and Kirby RS. Development and implementation of the first national data quality standards for population-based birth defects surveillance programs in the United States. *BMC Public Health*. 2015;15:925.

21. Correa A, Cragan JD, Kucik JE, Alverson CJ, Gilboa SM, Balakrishnan R, Strickland MJ, Duke CW, O'Leary LA, Riehle-Colarusso T, Siffel C, Gambrell D, Thompson D, Atkinson M and Chitra J. Reporting birth defects surveillance data 1968-2003. *Birth Defects Res A Clin Mol Teratol*. 2007;79:65-186.

22. Riehle-Colarusso T, Strickland MJ, Reller MD, Mahle WT, Botto LD, Siffel C, Atkinson M and Correa A. Improving the quality of surveillance data on congenital heart defects in the metropolitan Atlanta congenital defects program. *Birth Defects Res A Clin Mol Teratol*. 2007;79:743-53.

23. Reller MD, Strickland MJ, Riehle-Colarusso T, Mahle WT and Correa A. Prevalence of congenital heart defects in metropolitan Atlanta, 1998-2005. *J Pediatr*. 2008;153:807-13.

24. Strickland MJ, Klein M, Correa A, Reller MD, Mahle WT, Riehle-Colarusso TJ, Botto LD, Flanders WD, Mulholland JA, Siffel C, Marcus M and Tolbert PE. Ambient air pollution and cardiovascular malformations in Atlanta, Georgia, 1986-2003. *Am J Epidemiol*. 2009;169:1004-14.

25. Strickland MJ, Riehle-Colarusso TJ, Jacobs JP, Reller MD, Mahle WT, Botto LD, Tolbert PE, Jacobs ML, Lacour-Gayet FG, Tchervenkov CI, Mavroudis C and Correa A. The importance of nomenclature for congenital cardiac disease: implications for research and evaluation. *Cardiol Young*. 2008;18 Suppl 2:92-100.

26. Mai CT, Riehle-Colarusso T, O'Halloran A, Cragan JD, Olney RS, Lin A, Feldkamp M, Botto LD, Rickard R, Anderka M, Ethen M, Stanton C, Ehrhardt J, Canfield M and National Birth Defects Prevention N. Selected birth defects data from population-based birth defects

surveillance programs in the United States, 2005-2009: Featuring critical congenital heart defects targeted for pulse oximetry screening. *Birth Defects Res A Clin Mol Teratol*. 2012;94:970-83.

27. Olney RS and Botto LD. Newborn screening for critical congenital heart disease: essential public health roles for birth defects monitoring programs. *Birth Defects Res A Clin Mol Teratol*. 2012;94:965-9.

28. Canfield MA, Mai CT, Wang Y, O'Halloran A, Marengo LK, Olney RS, Borger CL, Rutkowski R, Fornoff J, Irwin N, Copeland G, Flood TJ, Meyer RE, Rickard R, Alverson CJ, Sweatlock J, Kirby RS and National Birth Defects Prevention N. The association between race/ethnicity and major birth defects in the United States, 1999-2007. *Am J Public Health*. 2014;104:e14-23.

29. Meyer RE, Liu G, Gilboa SM, Ethen MK, Aylsworth AS, Powell CM, Flood TJ, Mai CT, Wang Y, Canfield MA and National Birth Defects Prevention N. Survival of children with trisomy 13 and trisomy 18: A multi-state population-based study. *Am J Med Genet A*. 2016;170:825-37.

30. Wang Y, Liu G, Canfield MA, Mai CT, Gilboa SM, Meyer RE, Anderka M, Copeland GE, Kucik JE, Nembhard WN, Kirby RS and National Birth Defects Prevention N. Racial/ethnic differences in survival of United States children with birth defects: a population-based study. *J Pediatr*. 2015;166:819-26 e1-2.

31. Broberg CS, Mitchell J, Rehel S, Grant A, Gianola A, Beninato P, Winter C, Verstappen A, Valente AM, Weiss J, Zaidi A, Earing MG, Cook S, Daniels C, Webb G, Khairy P, Marelli A, Gurvitz MZ and Sahn DJ. Electronic medical record integration with a database for adult congenital heart disease: Early experience and progress in automating multicenter data collection. *Int J Cardiol*. 2015;196:178-82.

32. Jacobs JP. Introduction--databases and the assessment of complications associated with the treatment of patients with congenital cardiac disease. *Cardiol Young*. 2008;18 Suppl 2:1-37.

33. Jacobs JP, Jacobs ML, Mavroudis C, Tchervenkov CI and Pasquali SK. Executive summary: The Society of Thoracic Surgeons Congenital Heart Surgery Database - twenty-second harvest - (Jan 1, 2011 - December 30, 2014). Spring 2015 Harvest.

34. Franklin RC, Jacobs JP, Krogmann ON, Beland MJ, Aiello VD, Colan SD, Elliott MJ, William Gaynor J, Kurosawa H, Maruszewski B, Stellin G, Tchervenkov CI, Walters Iii HL, Weinberg P and Anderson RH. Nomenclature for congenital and paediatric cardiac disease: historical perspectives and The International Pediatric and Congenital Cardiac Code. *Cardiol Young*. 2008;18 Suppl 2:70-80.

35. Jacobs JP, O'Brien SM, Pasquali SK, Gaynor JW, Mayer JE, Jr., Karamlou T, Welke KF, Filardo G, Han JM, Kim S, Quintessenza JA, Pizarro C, Tchervenkov CI, Lacour-Gayet F, Mavroudis C, Backer CL, Austin EH, 3rd, Fraser CD, Tweddell JS, Jonas RA, Edwards FH, Grover FL, Prager RL, Shahian DM and Jacobs ML. The Society of Thoracic Surgeons Congenital Heart Surgery Database mortality risk model: Part 2-clinical application. *Ann Thorac Surg*. 2015;100:1063-70.

36. O'Brien SM, Jacobs JP, Pasquali SK, Gaynor JW, Karamlou T, Welke KF, Filardo G, Han JM, Kim S, Shahian DM and Jacobs ML. The Society of Thoracic Surgeons Congenital Heart Surgery Database mortality risk model: Part 1-statistical methodology. *Ann Thorac Surg*. 2015;100:1054-62.

37. Jacobs JP and Jacobs ML. Transparency and public reporting of pediatric and congenital heart surgery outcomes in North America. *World J Pediatr Congenit Heart Surg*. 2016;7:49-53.

38. Pasquali SK, Jacobs ML, O'Brien SM, He X, Gaynor JW, Gaies MG, Peterson ED, Hirsch-Romano JC, Mayer JE and Jacobs JP. Impact of patient characteristics on hospital-level outcomes assessment in congenital heart surgery. *Ann Thorac Surg*. 2015;100:1071-6; discussion 1077.

39. Riehle-Colarusso T, Autry A, Razzaghi H, Boyle CA, Mahle WT, Van Naarden Braun K and Correa A. Congenital heart defects and receipt of special education services. *Pediatrics*. 2015;136:496-504.

40. United States Census Bureau. American Community Survey Information Guide. 2010.

41. Data Resource Center for Child and Adolescent Health. National Survey of Children with Special Health Care Needs: Child and adolescent health measurement initiative.

42. National Center for Health Statistics. 2003 Revisions of the U.S. standard certificates of live birth and death and the fetal death report.

43. Kirby RS and Salihu HM. Back to the future? A critical commentary on the 2003 U.S. national standard certificate of live birth. *Birth*. 2006;33:238-44.

44. National Center for Health Statistics. Vital statistics data available on-line.

45. Gilboa SM, Salemi JL, Nembhard WN, Fixler DE and Correa A. Mortality resulting from congenital heart disease among children and adults in the United States, 1999 to 2006. *Circulation*. 2010;122:2254-63.

46. Broussard CS, Gilboa SM, Lee KA, Oster M, Petrini JR and Honein MA. Racial/ethnic differences in infant mortality attributable to birth defects by gestational age. *Pediatrics*. 2012;130:e518-27.

47. Watkins ML, Edmonds L, McClearn A, Mullins L, Mulinare J and Khoury M. The surveillance of birth defects: the usefulness of the revised U.S. standard birth certificate. *Am J Public Health*. 1996;86:731-4.

48. Boulet SL, Shin M, Kirby RS, Goodman D and Correa A. Sensitivity of birth certificate reports of birth defects in Atlanta, 1995-2005: effects of maternal, infant, and hospital characteristics. *Public Health Rep*. 2011;126:186-94.

49. Kirby RS. The quality of data reported on birth certificates. *Am J Public Health*. 1997;87:301.

50. Spector LG, Menk JS, Vinocur JM, Oster ME, Harvey BA, St Louis JD, Moller J and Kochilas LK. 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.

51. Brooks EG and Reed KD. Principles and pitfalls: a guide to death certification. *Clin Med Res*. 2015;13:74-82.

52. Pasquali SK, Jacobs JP, Shook GJ, O'Brien SM, Hall M, Jacobs ML, Welke KF, Gaynor JW, Peterson ED, Shah SS and Li JS. Linking clinical registry data with administrative data using indirect identifiers: implementation and validation in the congenital heart surgery population. *Am Heart J*. 2010;160:1099-104.

53. Pasquali SK, Jacobs ML, He X, Shah SS, Peterson ED, Hall M, Gaynor JW, Hill KD, Mayer JE, Jacobs JP and Li JS. Variation in congenital heart surgery costs across hospitals. *Pediatrics*. 2014;133:e553-60.

54. Pasquali SK, Li JS, He X, Jacobs ML, O'Brien SM, Hall M, Jaquiss RD, Welke KF, Peterson ED, Shah SS and Jacobs JP. Comparative analysis of antifibrinolytic medications in pediatric heart surgery. *J Thorac Cardiovasc Surg*. 2012;143:550-7.

55. Vinocur JM, Moller JH and Kochilas LK. Putting the Pediatric Cardiac Care Consortium in context: evaluation of scope and case mix compared with other reported surgical datasets. *Circ Cardiovasc Qual Outcomes*. 2012;5:577-9.

56. Kochilas L, Vinocur JM, St Louis JD, Harvey BA, Moller JH and Spector L. In-hospital vital status after interventions for congenital heart diseases: completeness of case ascertainment using the National Death Index. *J Am Coll Cardiol*. 2016;67:1011.

57. Pasquali SK, Jacobs JP, Farber GK, Bertoch D, Blume ED, Burns KM, Campbell R, Chang AC, Chung WK, Riehle-Colarusso T, Curtis LH, Forrest CB, Gaynor WJ, Gaies MG, Go AS, Henchey P, Martin GR, Pearson G, Pemberton VL, Schwartz SM, Vincent R and Kaltman JR. Report of the National Heart, Lung, and Blood Institute working group: An integrated network for congenital heart disease research. *Circulation*. 2016;133:1410-8.

58. Siffel C, Riehle-Colarusso T, Oster ME and Correa A. Survival of children with hypoplastic left heart syndrome. *Pediatrics*. 2015;136:e864-70.

59. Rasmussen SA, Olney RS, Holmes LB, Lin AE, Keppler-Noreuil KM, Moore CA and National Birth Defects Prevention S. Guidelines for case classification for the National Birth Defects Prevention Study. *Birth Defects Res A Clin Mol Teratol*. 2003;67:193-201.

60. Yoon PW, Rasmussen SA, Lynberg MC, Moore CA, Anderka M, Carmichael SL, Costa P, Druschel C, Hobbs CA, Romitti PA, Langlois PH and Edmonds LD. The National Birth Defects Prevention Study. *Public Health Rep*. 2001;116 Suppl 1:32-40.

61. Botto LD, Lin AE, Riehle-Colarusso T, Malik S, Correa A and National Birth Defects Prevention S. Seeking causes: classifying and evaluating congenital heart defects in etiologic studies. *Birth Defects Res A Clin Mol Teratol*. 2007;79:714-27.

62. Reefhuis J, Gilboa SM, Anderka M, Browne ML, Feldkamp ML, Hobbs CA, Jenkins MM, Langlois PH, Newsome KB, Olshan AF, Romitti PA, Shapira SK, Shaw GM, Tinker SC, Honein MA and National Birth Defects Prevention S. The National Birth Defects Prevention Study: A review of the methods. *Birth Defects Res A Clin Mol Teratol*. 2015;103:656-69.

63. Malik S, Cleves MA, Honein MA, Romitti PA, Botto LD, Yang S, Hobbs CA and National Birth Defects Prevention S. Maternal smoking and congenital heart defects. *Pediatrics*. 2008;121:e810-6.

64. Gilboa SM, Correa A, Botto LD, Rasmussen SA, Waller DK, Hobbs CA, Cleves MA, Riehle-Colarusso TJ and National Birth Defects Prevention S. Association between prepregnancy body mass index and congenital heart defects. *Am J Obstet Gynecol*. 2010;202:51 e1-51 e10.

65. Broussard CS, Rasmussen SA, Reefhuis J, Friedman JM, Jann MW, Riehle-Colarusso T, Honein MA and National Birth Defects Prevention S. Maternal treatment with opioid analgesics and risk for birth defects. *Am J Obstet Gynecol*. 2011;204:314 e1-11.

66. Polen KN, Rasmussen SA, Riehle-Colarusso T, Reefhuis J and National Birth Defects Prevention S. Association between reported venlafaxine use in early pregnancy and birth defects, national birth defects prevention study, 1997-2007. *Birth Defects Res A Clin Mol Teratol*. 2013;97:28-35.

67. Hartman RJ, Riehle-Colarusso T, Lin A, Frias JL, Patel SS, Duwe K, Correa A, Rasmussen SA and National Birth Defects Prevention S. Descriptive study of nonsyndromic atrioventricular septal defects in the National Birth Defects Prevention Study, 1997-2005. *Am J Med Genet A*. 2011;155A:555-64.

68. Lin AE, Krikov S, Riehle-Colarusso T, Frias JL, Belmont J, Anderka M, Geva T, Getz KD, Botto LD and National Birth Defects Prevention S. Laterality defects in the National Birth Defects Prevention Study (1998-2007): birth prevalence and descriptive epidemiology. *Am J Med Genet A*. 2014;164A:2581-91.

69. Centers for Disease Control and Prevention. Population-based surveillance of congenital heart defects among adolescents and adults.

Table 1. Administrative healthcare database examples in the United States and Canada from 1990 onward for potential use in congenital heart defects (CHDs) public health investigations. Cardiac-specific databases are marked with an asterisk (*).

Name	Brief Description	Sponsoring Organization	Data Years	URL (accessed as of June 1, 2016)
ADMINISTRATIVE – Heathcare Cost and Utilization Project (HCUP) Databases				
Kids' Inpatient Database (KID)	Weighted sample of the SID data (see below) used to identify, track, & analyze national trends in pediatric inpatient healthcare. Sampling weights help provide national estimates.	Agency for Healthcare Research and Quality (AHRQ)	Every 3 years; 1997-present	http://www.hcup-us.ahrq.gov/kidoverview.jsp
Nationwide Emergency Department Sample (NEDS)	Sampled from the SID and SEDD data (see below), is the largest all-payer emergency department (ED) database in the U.S. Used to create estimates of ED care.	Agency for Healthcare Research and Quality (AHRQ)	2006-2013	http://www.hcup-us.ahrq.gov/nedsoverview.jsp
Nationwide/National Inpatient Sample (NIS)	Weighted sample of discharges from U.S. community hospitals, which is the largest publicly available all-payer inpatient healthcare database in the US. Sampling weights help provide national estimates.	Agency for Healthcare Research and Quality (AHRQ)	1988-present	http://www.hcup-us.ahrq.gov/nisoverview.jsp
Nationwide Readmission Database (NRD)	Sampled from the SID data (see below), used to create estimates of national readmission rates for all payers and the uninsured.	Agency for Healthcare Research and Quality (AHRQ)	2013	http://www.hcup-us.ahrq.gov/nrdoverview.jsp
State Ambulatory Surgical and Service Databases (SASD)	Encounter data for ambulatory surgery & other outpatient services from hospital-owned facilities. Capture of hospital-based outpatient diagnostic and/or pediatric cardiac catheterization is variable, as is data content & years. Some states have non-hospital outpatient data.	Agency for Healthcare Research and Quality (AHRQ)	1997-present	http://www.hcup-us.ahrq.gov/sasdoverview.jsp
State Emergency	Discharge data on all ED visits in a given state that do not	Agency for Healthcare	1999-	http://www.hcup-us.ahrq.gov/seddoverview.jsp

Department Databases (SEDD)	result in an admission. ED visits resulting in admissions are captured in the SID	Research and Quality (AHRQ)	present	
State Inpatient Databases (SID)	Inpatient discharge data from participating states used to identify, track, & analyze state trends in healthcare utilization, access, charges, quality, and outcomes.	Agency for Healthcare Research and Quality (AHRQ)	1990-present	http://www.hcup-us.ahrq.gov/sidoverview.jsp
ADMINISTRATIVE - Others				
Healthcare Cost Institute (HCCI) Database	Collection of claims data by non-partisan, non-profit organization on over 50 million people with employer-sponsored insurance. Annual reports published and data available to researchers to better understand determinants of US health care costs and utilization	Healthcare Cost Institute (HCCI)	2007-present	http://www.healthcostinstitute.org/
MarketScan® Research Databases	Database system linking healthcare usage through commercial insurance, Medicaid and Medicare to analyze a variety of outcomes.	Truven Health Analytics	1995-present	http://truvenhealth.com/your-healthcare-focus/analytic-research/marketscan-research-databases
Medicaid Analytic Extracts (MAX)	Contain enrollment information and paid claims at the state level for all Medicaid beneficiaries for inpatient and outpatient care, prescription medications.	Centers for Medicare and Medicaid Services (CMS)	1999-2012	http://www.resdac.org/cms-data/file-family/Medicaid-Analytic-Extracts-MAX
National Association of Children's Hospitals (NACH) Case Mix Comparative Data Program	Database of 95 children's hospitals in the US with data to analyze inpatient populations, target quality improvement, enhances hospital utilization, & support advocacy on behalf of children's hospitals.	National Children's Hospital Association	2000-2012	http://www.childrenshospitals.net/Am/Template.cfm?Section=Home3
Pediatric Health Information System (PHIS)	Database of clinical and financial data from 49 tertiary care pediatric hospitals in the US affiliated with the Children's Hospital Association. Data can be linked across encounters within the same hospital.	Children's Hospital Association	1992-present	http://www.childrenshospitals.org/

* Quebec Congenital Heart Disease Database	Database from 3 province-wide administrative databases, capturing demographics, diagnoses, procedures, and health services used throughout a patient's life.	McGill Adult Unit Congenital Heart Disease Excellence	1983-present	none

Table 2. Birth defects surveillance database examples in the United States and Canada from 1990 onward for potential use in congenital heart defects (CHDs) public health investigations.

Name	Brief Description	Sponsoring Organization	Data Years	URL (accessed as of June 1, 2016)
Florida Birth Defects Registry (FBDR)	Statewide population-based birth defects surveillance program for live born infants ≤ age 1 year. Multiple data sources & linkages, including hospital & ambulatory discharge data, Children's Medical Services Florida, vital records, other administrative and clinical data.	Florida Department of Health	1999-present	http://fbdr.org/
Metropolitan Atlanta Congenital Defects Program (MACDP)	Population-based birth defects surveillance program for live born & stillborn infants, fetuses, and children diagnosed up to 6 years of age born to residents of metropolitan Atlanta, Georgia. Multiple clinical data sources with linkage to vital records. Active case finding, review and classification of CHDs.	CDC[1] – National Center on Birth Defects and Developmental Disabilities	1967 - present	http://www.cdc.gov/ncbddd/birthdefects/macdp.html
National Birth Defects Prevention Network (NBDPN) Data Repository	Data registry from 12 birth defect surveillance systems (including FBDR and MACDP) collaborating on birth defects surveillance, research, and prevention projects.	National Birth Defects Prevention Network	1998-2007	http://www.nbdpn.org/

Table 3. Clinical database examples in the United States and Canada from 1990 onward for potential use in congenital heart defects (CHDs) public health investigations. Cardiac-specific databases are marked with an asterisk (*).

Name	Brief Description	Sponsoring Organization	Data Years	URL (accessed as of June 1, 2016)
* Congenital Cardiac Catheterization Outcomes Project (C3PO)	Database of patient and procedural characteristics on catheterization procedures performed for congenital & acquired heart disease in infants, children, and adults at 15 pediatric heart centers.	Boston Children's Hospital Cardiovascular Program	2007-present	https://c3po-qi.chboston.org
* Congenital Cardiac Interventional Study Consortium (CCISC)	Registry for demographic and procedural information on patients undergoing diagnostic & interventional cardiac catheterizations for CHD.	Children's Hospital of Michigan Foundation	2005-2010	http://www.chmfoundation.org/congenital-cardiac-interventional-study-consortium-case/
* Congenital Evaluation, Reporting, and Tracking Endeavor (CONGENERATE)	Database for providers of adults with CHD for multicenter collaboration, research, & quality metric initiatives.	McGill University, University of Sherbrooke, University of Montreal	2010-present	http://www.congenerate.org/
* Congenital Heart Surgeons Society (CHSS) Database	Database for multi-institutional clinical studies evaluating surgical interventions for CHD. Goals: increase, correlate, & disseminate knowledge of physiology, pathology and therapy	Congenital Heart Surgeons' Society (CHSS)	1985-present	http://www.chssdc.org/
* IMproving Pediatric and Adult Congenital Treatments (IMPACT™) Registry	Registry of demographics, management & outcomes of pediatric and adult patients with CHD undergoing diagnostic & intervention cardiac catheterizations and electrophysiology procedures at 55 sites. Data for performance measurement, benchmarking, and quality improvement initiatives.	American College of Cardiology/ National Cardiovascular Data Registry	2010-present	https://www.ncdr.com/webncdr/impact/home
* Mid-Atlantic Group of Interventional Cardiology	Registry of outcomes for specific cardiac interventional catheterizations for CHDs and pulmonary hypertension at 14	Johns Hopkins University	2003-2010	http://www.magicgroup.org/html/news.html

(MAGIC)	sites.			
* National Pediatric Cardiology Quality Improvement Collaborative	Providers & family network which collects data, conducts research, & uses quality improvement science to improve outcomes. Multi-center database to identify care variations & best practices, and test hypotheses.	Joint Council on Congenital Heart Disease	2006-present	https://jcchdqi.org/
Organ Procurement Transplant Network Database	Database containing secure data on all wait lists, organ donation & transplant events in the U.S. Database can be queried online and reports available.	United Network for Organ Sharing (UNOS)	1987-present	http://optn.transplant.hrsa.gov/
* Pediatric Cardiac Care Consortium (PCCC)	Registry of cardiac catheterizations, surgeries, & autopsies for infants, children, and adults with congenital or acquired heart disease from 57 pediatric cardiac centers. Includes outcomes & longitudinal patient tracking	University of Minnesota	1982-2011	http:// www.pcccweb.com
* Pediatric Cardiac Critical Care Consortium (PC4)	Consortium of pediatric cardiac critical care, cardiac surgery, & cardiology which collects data on outcomes & practice, provides performance feedback, and promotes improvement based on empirical analysis and collaborative learning.	National Institutes of Health / University of Michigan / Participating Sites	2009-present	http://pc4quality.org/
* Pediatric Heart Network (PHN)	Collaboration of clinical sites & a data coordinating center that conducts research to improve outcomes and quality of life of children with congenital and acquired heart disease. Centers follow study protocol to collect identical data and treat patients in similar ways	National Heart, Lung, & Blood Institute	2001-present	http://www.pediatricheartnetwork.com/
* Pediatric Heart Transplant Study (PHTS) Database	International, prospective, event driven database for research in the field of pediatric heart transplantation. PHTS advances the science & treatment of children during listing for and following heart transplantation	University of Alabama Birmingham School of Medicine	1993-present	http://www.uab.edu/medicine/phts/

* Society of Thoracic Surgeons Congenital Heart Surgery Database (STS-CHSD)	Database for quality improvement, patient safety, and research which contains data on >95% of pediatric cardiac operations in the US. Represents120 United States pediatric cardiac surgery hospitals & 3 in Canada	Society of Thoracic Surgeons	1994-present	http://www.sts.org/sts-national-database/database-managers/congenital-heart-surgery-database
Virtual Pediatric Intensive Care Unit Systems, LLC (VPS)	Collaboration of 115 hospitals to improve critical care quality& outcomes through actionable reports, data management, & research	none	1998-present	http://www.myvps.org/about-vps.html
* Western Canadian Children's Heart Network Database (WCCHN)	Database containing data on all diagnosed pediatric heart disease and CHD cases for five Canadian sites, and adult CHD cases in one site.	Western Canadian Children's Heart Network	2006-present	http://www.westernchildrensheartnetwork.ca/

Table 4. Survey database examples in the United States and Canada from 1990 onward for potential use in congenital heart defects (CHDs) public health investigations.

Name	Brief Description	Sponsoring Organization	Data Years	URL (accessed as of June 1, 2016)
American Community Survey (ACS)	Part of the Decennial Census Program, it is a nationwide continuous survey sent to a small percentage of U.S. households to gather demographic, housing, social, and economic data and provide yearly reports.	United States Decennial Census Program – Census Bureau	2005-present	http://www.census.gov/programs-surveys/acs/
Decennial Census	Survey of all U.S. households done every ten years, consisting of short and long forms. As of 2010, only the short-form is done – the long-form replaced by the ACS. Data is used for numerous purposes.	United States Census Bureau	1790-present	https://www.census.gov/

Medical Expenditure Panel Survey (MEPS)	Survey of households to estimate use of health services, cost, payment & availability. Surveys have 3 components: core household, insurance/employer, and the medical provider	Agency for Healthcare Research and Quality (AHRQ)	1996-Present	http://meps.ahrq.gov/mepsweb/
National Health Interview Survey (NHIS)	Survey of households to estimate the amount, distribution, & effects of illness & disability in the US across demographics and socioeconomic status. Updated questions on select topics. Main source of health information of the US population.	CDC[1] -National Center for Health Statistics	1957-present.	http://www.cdc.gov/nchs/nhis.htm
National Survey on Children with Special Health Care Needs (NS-CSHCN)	Random sample survey of households in all states to assess prevalence & impact of special healthcare needs among children in the US. Survey has core & special topic areas such as CHDs	CDC[1] -National Center for Health Statistics – Maternal Child Health Bureau	2000, 2005, 2009	http://www.cdc.gov/nchs/slaits/cshcn.htm

Table 5. Combined database examples in the United States and Canada from 1990 onward for potential use in congenital heart defects (CHDs) public health investigations.

Name	Brief Description	Sponsoring Organization	Data Years	URL (accessed as of June 1, 2016)
ADMINISTRATIVE AND CLINICAL				
Healthcare Systems Research Network (HCSRN)	Collaboration of 18 integrated healthcare delivery systems implementing research findings in clinical practice. Working over a broad scope of indicators, they aim to develop an extensive and usable database.	none	2006-present	http://www.hcsrn.org/en/
Pediatric Health	Database augmenting the existing PHIS database by linking	Agency for Healthcare	2009-	http://www.prisnetwork.org/research/phis_plus.ht

Information System Plus (PHIS+)	electronic health records of laboratory and radiology reports from 6 of the 49 Children's Hospital Association hospitals to conduct clinical comparative effectiveness research projects.	Research and Quality (AHRQ) and Children's Hospital Association	2012	ml
ADMINISTRATIVE AND SURVEY				
National Ambulatory Medical Care Survey (NAMCS)	National sample survey of non-federal office-based physicians to provide data on ambulatory medical care services in the U.S.	CDC[1] -National Center for Health Statistics	1973-present	http://www.cdc.gov/nchs/ahcd/about_ahcd.htm
National Hospital Ambulatory Medical Care Survey (NHAMCS)	National sample survey of hospital emergency, outpatient, hospital-based & non-hospital ambulatory surgery centers. Provides data on care at hospital-based ambulatory services and ambulatory surgical centers.	CDC[1] -National Center for Health Statistics	1992-present	http://www.cdc.gov/nchs/ahcd/about_ahcd.htm
National Hospital Discharge Survey (NHDS)	Probability survey of inpatients discharged from non-Federal short-stay U.S. hospitals. Provides national estimates of hospital inpatient services.	CDC[1] -National Center for Health Statistics	1965-2010	http://www.cdc.gov/nchs/nhds.htm
National Hospital Care Survey (NHCS)	Survey combining data from NHAMCS, NHDS, and drug abuse network.	CDC[1] -National Center for Health Statistics	2011-present	http://www.cdc.gov/nchs/nhcs.htm
BIRTH DEFECT SURVEILLANCE AND SURVEY				
Birth Defects Study To Evaluate Pregnancy exposureS (BD-STEPS)	Multi-site population-based, case-control study of 17 birth defects, building upon findings from the NBDPS.	Centers for Birth Defects Research and Prevention	2014-present	http://bdsteps.org/
National Birth Defects Prevention Study (NBDPS)	Multi-site population-based, case-control study of 30 birth defects. Includes maternal interview & cheek cell specimens from family members. Excludes syndromes & chromosomal abnormalities.	Centers for Birth Defects Research and Prevention	1997-2011	http://www.nbdps.org/

Pregnancy Health Interview Study (Birth Defects Study)	Multi-site case-control study of birth defects & newborn health focuses on environmental exposures (primarily medications) in pregnancy. Includes maternal interview, medical record release. Genetic specimens 1992-2008.	Slone Epidemiology Center at Boston University	1979-present	http://www.bu.edu/slone/research/studies/phis/