A Systematic Review of the Literature on Birth Defects Surveillance Methods: Where are the Gaps?

Nicole P. Lindsey, Victoria K. Markwalter, and Laura J. Pabst

Birth Defects Monitoring and Research Branch

Division of Birth Defects and Infant Disorders

National Center on Birth Defects and Developmental Disabilities

Centers for Disease Control and Prevention

August 2026

Introduction

Historically, birth defects surveillance programs implemented one of two approaches to surveillance: active or passive case ascertainment. The active case ascertainment approach encompassed an intensive level of case identification that involved staff visiting healthcare facilities and having access to medical information such as logbooks and medical records to identify cases. Staff would then abstract relevant information from maternal and infant medical records onto a reporting form. Programs using this surveillance method usually had very complete ascertainment and each diagnosis was confirmed or verified. Although active case ascertainment required considerable resources and personnel, such efforts resulted in complete and accurate capture of case information.

In a passive case ascertainment approach the surveillance program relied upon data sources to report cases as they occurred. Traditionally, case reports were not confirmed by staff in a passive surveillance program but were accepted as reported. Information on the prevalence of birth defects reported by active and passive surveillance systems varied considerably due to differences in methods of case ascertainment and verification, limiting the utility of collected data for calculating national estimates and identifying significant differences or trends.

Birth defects surveillance methods have evolved substantially in the past two decades as electronic health records and other technological advancements have become widely available and implemented. As methods have evolved, the distinction between active and passive approaches has become less distinct, with traditionally passive programs incorporating electronic data sources for case identification and verification and active programs relying more on automated electronic case notifications, such as hospital discharge indices. Despite the changes, there are also historical methods still in practice in many programs. As more data sources, healthcare coding systems, and data transfer options come into use, it is critical that these potential sources and methods are evaluated to know which will improve efficiency, completeness, and accuracy of birth defects surveillance.

This systematic review examines the available literature around birth defects case identification (finding infants that may have defects under surveillance), investigation (collecting data needed for surveillance purposes), and verification (review of collected data to confirm defect diagnoses) processes. Questions to be addressed by this literature review were developed considering the current NBDPN Birth Defects Surveillance Guidelines[1] and the NBDPN Performance Summary, which is a survey completed by state, territorial, or regional birth defects programs biennially about

surveillance methods being used, including the sources used to identify potential cases and the criteria used to require or trigger reporting from a data source to the program. Additional methodological information collected in the survey included the defects under surveillance, residency criteria for inclusion, pregnancy outcomes under surveillance (e.g., live births, fetal deaths), maximum age of diagnosis for identification, modes of case notification (e.g., individual reports, multiple case files), and medical record abstraction and clinical review practices. Where possible, we have aligned the findings of the literature review with the organization of the Performance Summary survey.

The following were identified as important questions for developing standardized methods for birth defects surveillance:

1. What data sources should be used to identify birth defects cases? What are the sensitivities/unique contributions of various case identification sources?
2. How broad should trigger codes and other reporting criteria be to optimize sensitivity while limiting false positive birth defects case reports?
3. What are the specificities/positive predictive values of case identification sources? Are medical abstraction and clinical review necessary to obtain accurate prevalence estimates?
4. Up to what age do birth defects need to be ascertained to obtain accurate prevalence estimates?
5. What pregnancy outcomes need to be included to obtain accurate prevalence estimates?

A synthesis of the current evidence base for BD surveillance methods is essential at this time, as the National Birth Defects Prevention Network (NBDPN) is undergoing a major update to the Surveillance Guidelines used across programs to guide and standardize data collections (NBDPN 2004). This systematic literature review supports efforts to determine where there is adequate evidence to inform NBDPN standards and guidelines and identify critical gaps in the current knowledge base to highlight areas needing research can guide future analyses and evaluations.

Data Source and Approach

A literature search was conducted in MEDLINE (via PubMed) for articles published between Jan 1, 2010, and April 23, 2026. Limiting the search to studies published during this time frame was designed to capture papers describing current surveillance practices. The search was also limited to articles published in English and to surveillance programs in US states and territories to ensure the results would be applicable to US programs. Only articles describing primary data were included in the analysis. Other articles (case reports, reviews, meta-analyses, editorials, recommendations, and guidelines) identified through the search were reviewed to identify additional references to studies that might meet the inclusion criteria.

The initial MEDLINE search returned 4,474 articles. All article titles identified through the search were reviewed to identify those needing abstract review and subsequent full paper review per the criteria outlined above. A total of 4,395 articles were excluded from just the title review, for being review/commentary pieces, assessments of international surveillance programs, or of an unrelated topic. Abstracts were reviewed for each of the remaining 79 articles, and an additional 37 articles were excluded for similar reasons as above. Forty-two articles were reviewed in full and 21 were excluded because they did not address the specific questions of the project, were older studies replicated in more recent included articles, or did not address surveillance of at least one of the NBDPN defects.

To identify relevant work that might have been unpublished, a call for abstracts, papers, and presentations was sent to the full NBDPN membership in the fall of 2024, asking for submission of those for inclusion in the review. This call for unpublished work did not identify any additional studies for inclusion. However, we reviewed abstracts accepted for poster, oral, or breakout session presentation at the 2025 NBDPN Conference to identify relevant new work not yet published and did identify two additional studies for inclusion. One of these two was subsequently published.

In sum, 23 studies are included in the evidence summarized in this paper. Several of the identified studies come from just two state programs (Florida, 8; New York, 3). Other programs represented included Alaska, Arizona, metro-Atlanta, Massachusetts, Minnesota, New Jersey, North Carolina, Tennessee, and Texas. The included studies use different inclusion criteria (e.g., pregnancy outcomes), different methods for identifying, abstracting and reviewing cases, and different defects and groupings of defects. This context is important when considering results.

Summary of Evidence

Question 1: What data sources should be used to identify birth defects cases? What are the sensitivities/unique contributions of case identification sources?

We identified 11 studies that assessed the contributions of one or more potential case identification sources assessed in the 2025 NBDPN Performance Summary (**Box 1)**. The identified studies did not separate delivery and pediatric/tertiary care hospitals in their analysis, so these are presented together, as hospital discharge databases. Articles assessed the following additional case identification sources: birth certificates, death and fetal death certificates, cytogenetic laboratory results, emergency department discharge data, and pulse oximetry screening program data. The remainder of the sources were not assessed by any identified studies. Two studies evaluated the sensitivity of a passive surveillance system network to capture cases compared to those captured by an active program.

Box 1. Case Identification Sources: 2025 NBDPN Performance Summary

Vital records (birth, death, fetal death, elective termination certificates)
Delivery hospitals
Pediatric and tertiary care hospitals
Other healthcare providers
Midwifery or birthing facilities
Third party payors (e.g., Medicaid, HMOs, Indian Health Service)
Maternal fetal medicine specialist
Specialty facilities (e.g., prenatal diagnostics, cytogenetics laboratories, genetic counseling services)
State-based registries (e.g., children with special needs, newborn screening, developmental disabilities)

Hospital inpatient discharge data

Two articles compared the sensitivity of hospital inpatient discharge data to cases captured by a passive surveillance system with multiple inputs (**Table 1**) (Wang 2010, Rutkowski 2015). Overall, hospital inpatient data had high (~94%) sensitivity compared to passive surveillance systems with multiple inputs. However, in both studies, hospital discharge data had lower sensitivity for some defect classes, such as ear/eye defects, which may be less likely to be detected at birth. This limited evidence suggests that hospital discharge data alone may be sufficient to identify most defects, particularly if the data are primarily collected to monitor trends. One additional paper evaluated the sensitivity of multiple data sources to capture critical congenital heart defects (CCHD) cases confirmed by an active surveillance program (Atkinson 2025). Hospital discharge records were the most sensitive with (88%) of confirmed CCHD cases captured (data not shown).

Table 1. Sensitivity of hospital inpatient discharge data compared to passive surveillance

	Wang et al. New York, 2010	Rutkowski et al. Florida, 2015*
Overall	**94%**	**94-96%**
Cardiovascular	96%	92-99%
Chromosomal	96%	89-97%
Central nervous system	94%	84-94%
Eye and ear	86%	73-92%
Gastrointestinal	95%	88-99%
Genitourinary	91%	91-92%
Musculoskeletal	95%	93-96%
Orofacial	97%	94-98%

**Rutkowski results presented as range because of differences in data sources included in the comparator (passive surveillance system) by year*

Birth certificates

Three articles compared the sensitivity of birth certificates to a gold standard of an active surveillance program (**Table 2**) (Boulet 2011, Banerjee 2012, Salemi 2017). Despite variations in geography and program methods, sensitivities among studies were consistently low, although they were higher for anencephaly (range 55-80%) than for other defects. One additional article evaluated the sensitivity of multiple data sources to capture CCHD cases confirmed by an active surveillance program; birth certificates had a sensitivity of only 7% (Atkinson 2025) (data not shown).

Table 2. Sensitivity of birth certificates compared to overall active surveillance program

	Boulet et al. Metro Atlanta, 2011	Banerjee et al. Minnesota, 2012	Salemi et al. Florida, 2017
All included defects	**23%**	**28%**	**19%**
By Defect			
Anencephaly	69%	80%	55%
Omphalocele/gastroschisis*		59%	54%
Hydrocephalus		50%	
Spina bifida	35%	33%	28%
Cleft lip/palate	31%	32%	36%
Limb reduction deficiency			9%
Clubfoot	17%	--	
Trisomy 21	18%	22%	25%
Diaphragmatic hernia		16%	
Renal agenesis		13%	
Hypospadias			5%
Microcephaly		7%	
Rectal atresia	7%	5%	
Tracheo-esophageal fistula		0%	
By Class			
Central nervous system		36%	
Musculoskeletal		33%	
Orofacial		30%	
Chromosomal		28%	
Genitourinary		8%	
Cardiovascular		5%	

	Boulet et al. Metro Atlanta, 2011	Banerjee et al. Minnesota, 2012	Salemi et al. Florida, 2017
Gastrointestinal		2%	

** This combined variable is not very useful given the vast epidemiological differences between these two defects*

Infant death certificates

Three studies were identified that evaluated the sensitivity of infant death certificates for case identification. One paper evaluated the sensitivity of multiple data sources to capture CCHD cases confirmed by an active surveillance program; infant death certificates had a relatively high sensitivity of 73% (Atkinson 2025).

A Florida study reported that the passive surveillance system captured 73% of infants reported as having died from a birth defect but case capture varied significantly by defect (Tanner 2010). If the infant death certificates were added as a case identification source, the overall prevalence of birth defects would only have increased by 0.6% but substantially increased the prevalence of some defects (e.g., anencephaly by 80%, Trisomy 18 by 25%, Trisomy 13 by 17%, hypoplastic left heart syndrome by 11% and renal agenesis by 10%). A major limitation of this study is the absence of any validation to confirm defects reported on the death certificates; therefore, the false positive rate is unknown.

A later Florida study similarly reported that death certificates uniquely identified <1% of birth defects overall, however they contributed significantly to the capture of defects commonly associated with early mortality, including anencephaly (64%), Trisomy 13 (48%), and interrupted aortic arch (10%) (Rutkowski 2015). Only for Trisomy 13 cases did the addition of death certificates result in a significant increase in prevalence rate. This study was also limited by a lack of any validation to confirm defects reported on death certificates.

Fetal death certificates

One unpublished abstract evaluated the sensitivity of multiple data sources to capture CCHD cases confirmed by an active surveillance program; fetal death certificates had a sensitivity of about 30%. This paper was subsequently published (Atkinson 2025). The unique contribution of this data source was not assessed.

Emergency department data

One article evaluated the contribution of emergency department data in relation to the overall passive surveillance system (Rutkowski 2015). The emergency department data uniquely identified 0.7%

cases overall and unique contribution was not above 5% for any specific defect, suggesting that this case identification source would not be the most efficient use of time for case finding when resources are limited.

Cytogenetics laboratories

One article evaluated the contribution of cytogenetic testing data to their existing passive surveillance system (Tao 2013). Overall, 55% of children with cytogenetic testing reported in the electronic clinical laboratory reporting system had been already identified by the passive surveillance system. Of the 45% not captured by the surveillance system, about half had an abnormal test result and 38% were determined to have a reportable defect following review. The addition of this data source increased reporting of all chromosomal abnormalities by 8%. The addition of this data source increased reporting of Down syndrome, specifically, by 5%. Other NBDPN chromosomal defects were not specifically mentioned in the article. Among cases with laboratory testing data who had already been identified by the passive system, 26% had new diagnoses added or had their diagnosis modified. The percentage of anomalies that would have been missed without the cytogenetic testing data ranged from 1.3% (other specified anomalies and syndromes) to 26.9% (autosomal deletion syndromes). Given these results, cytogenetics laboratory data might be considered for inclusion as a recommended case identification source, depending on the chromosomal defects under surveillance.

Pulse oximetry screening program

One unpublished abstract evaluated the sensitivity of multiple data sources to capture CCHD cases confirmed by an active surveillance program; pulse oximetry screening (POxS) records had a sensitivity of about 60% (Atkinson 2025). The unique contribution of this data source was not assessed. One other article reported on the utility of a POxS program to support birth defects surveillance for CCHD (Garg 2013). Although they did not calculate sensitivity nor the unique contributions to case identification, the authors reported that about 60% of reported failed POxS had no clinical indicators for CCHD, and diagnostic evaluation was solely attributable to failed screen. About 10% of those without clinical indicators ultimately had CCHDs that might have resulted in serious morbidity if diagnosis had been delayed. From this study, it is not clear if these cases would have been identified through hospital discharge or other data and therefore the utility of this diagnostic tool as a case identification source is unclear.

Electronic case reporting

Electronic case reporting (eCR), the automated, secure, real-time exchange of public health disease information between reporting facilities/providers and public health agencies, is a relatively new potential data source for birth defects surveillance programs. Few birth defects surveillance programs have piloted or implemented eCR for birth defects as of the writing of this paper. The Texas Birth Defects Epidemiology and Surveillance Branch is one of the first to implement eCR for birth defects. Allred et al. described an evaluation conducted in 2024 to determine the feasibility of using eCR for case identification across eight birth defects (anencephaly, cleft lip alone, cleft palate alone, cleft lip with cleft palate, trisomy 21, gastroschisis, limb reduction, and spina bifida) (Allred 2025). The authors reported high concordance for several data elements that are important for case identification, including sex, date of birth, primary birth defect, and contact information. Though information in the electronic initial case report (eICR) is taken directly from the electronic health record (EHR), the eICR did not always contain all the relevant birth defect diagnostic information available in the EHR. However, this evaluation reported a high degree of concordance for the primary birth defect (overarching category or 4-digit BPA code), indicating that the eICR data are suitable for case notification and for timely identification of cases that are potentially eligible for referral to early intervention services.

Sensitivity of passive surveillance compared to active surveillance

Salemi et al. evaluated the sensitivity of a passive surveillance system to capture eight major defects including gastroschisis, orofacial clefts, trisomies 13,18, 21, anencephaly, spina bifida, and encephalocele in selected counties in Florida that were also covered by an enhanced surveillance program (Salemi 2012). They reported an overall sensitivity (not defect specific) of 89% compared to active, enhanced surveillance (which included medical record abstraction and clinical review). There was significant variability in sensitivity by specific defect, from only 46% for anencephaly to 89% for Trisomy 21. The authors concluded that for rapidly fatal conditions, fewer medical records for the infants may contribute to the poor sensitivity.

A similar study was conducted for 16 defects in selected counties in New York (Howley 2024). The passive surveillance system captured more than 90% of cases for eight included defects: CCHDs (100%, as a group, not for specific CCHDs), spina bifida (96%), gastroschisis (95%), and esophageal atresia (93%). Sensitivity did vary by specific CCHD defects from 98% (tetralogy of Fallot) to 61% (tricuspid atresia). Sensitivities were lowest for anophthalmia/microphthalmia (81%) and anotia/microtia (77%). The authors reported that the missing specific birth defect diagnoses were

primarily due to hospital coding errors; 76% of those missed by passive surveillance were reported as having another defect.

Summary of Evidence for Question 1:

In summary, there is incomplete evidence in the published literature to make specific recommendations about best practices for identification of birth defects cases. The available evidence suggests that hospital inpatient discharge data capture >90% of most defects. Defect classes that might be under-represented include those less likely to be detected at birth and perhaps those most likely to result in death shortly after delivery. Future studies are needed to evaluate which data sources are best suited to fill those gaps in case identification, but it seems possible that hospital discharge data accompanied by just a few additional case identification sources targeted to identify particular defects or classes of defects can be used to identify the vast majority of cases.

Question 2: How broad should trigger codes and other reporting criteria be to optimize sensitivity while limiting false positive birth defects case reports?

No publications were identified that evaluated use of specific diagnosis or procedure codes or other triggers such as text descriptors as outlined in the Performance Summary survey (**Box 2**).

Box 2. Case Identification Criteria: 2025 NBDPN Performance Summary

Specific indication of birth defects
ICD-9-CM code 740-759 or ICD-10-CM code Q00-Q99
Prenatally diagnosed or suspected defects
Birth certificates with indication of defect
Death or fetal death certificates with indication of defect
Charts with text that describes a defect specifically (e.g., Down syndrome)

No specific indication of birth defects
ICD-9-CM code outside 740-759 or ICD-10-CM code outside Q00-Q99
Charts with selected procedure codes
Infants with low APGAR scores
Infants with low birth weight or gestational age at birth
Infants in NICU or other special care nursery
Charts with text that describe a condition that may suggest a defect (e.g., abnormal facies)
Neonatal deaths, fetal deaths/stillborn infants

One article assessed the predictive value of an O35 code (indicates maternal care for known or suspected fetal abnormality and damage) on fetal death certificates, finding that the positive

predictive value (PPV) was 69% and the negative predictive value (NPV) was 96% for detecting cases with a birth defect (Tessarolo 2021).

One article assessed the impact of expanding the number of diagnoses included in data pulled from inpatient discharge databases on counts and rate of birth defects from the first ten codes to a set of 31 codes (Salemi 2018). When restricted to first ten codes, 3% of birth defect cases overall would have been missed. Less than 1% of cases among anencephaly and multiple congenital heart defects would have been missed. However, higher percentages of cases among six of the defects assessed would have been missed, including anophthalmia/microphthalmia (22%), anotia/microtia (22%), congenital cataract (20%), aniridia (14%), choanal atresia (14%), and limb deficiencies (14%). For 11 defects, the estimated prevalence using ten codes was statistically different from that estimated using 31 codes. The impact of the code expansion was not significant for severe, lethal, and easily recognized defects.

Two studies evaluated the potential effect of reviewing medical records for all fetal deaths (not just those with anomaly codes) on case capture. An Arizona study found that if all fetal deaths were not reviewed, the overall impact on birth defect prevalence was small (~4%), but about 60% of fetal deaths associated with a birth defect would have been missed (Tessarolo 2021). This method did result in substantial additional work; instead of reviewing only 51 charts with O35 codes, 1,440 charts required review. The authors reported that for well-trained abstractors, the chart review process was quick, taking only five to ten minutes per case. A similar assessment of all fetal death reports regardless of birth defect indication was conducted in Minnesota (Thom and Cho 2025). Of the fetal death reports, 20% were found to have a birth defect. Of those, 85% had a defect indicated on the death report, but 15% did not. False negative and false positive rates were 15% and 7%, respectively. No publications were identified that evaluated the other non-specific criteria, such as all babies admitted to NICU, low APGAR, neonatal deaths, or with low gestational age or birth weight.

Summary of Evidence for Question 2:

In summary, the evidence to guide recommendations for trigger codes or reporting criteria for birth defects surveillance is very limited. Further research is needed to evaluate these elements.

Question 3: What are the specificities/positive predictive values of case identification sources? Are medical abstraction and clinical review necessary to obtain accurate prevalence estimates?

Four articles were identified that assessed PPV of one or more case identification sources (**Box 1**). The identified articles assessed PPV of hospital discharge data, birth certificate data, and of overall passive surveillance systems with multiple inputs compared to a gold standard of cases confirmed by medical record abstraction and clinical review. The defects included varied by study. For most case identification sources, there were no articles identified that assessed PPVs.

Hospital discharge data

One article calculated the PPV of hospital discharge data for CCHDs following medical abstraction and pediatric cardiologist review (Barnett 2024). The proportion verified varied from 19% (common truncus) to 84% (TAPVC). The article did not quantify how many were miscoded or were determined not to have a CCHD or birth defect at all. Additionally, the analysis did not address CCHDs that were missed (false negatives). The authors concluded that for CCHDs, review of passive reports is necessary because false positive rates did significantly impact prevalence estimates.

Birth certificates

Three articles assessed the positive predictive value of birth certificates for identifying birth defects (**Table 3**) (Boulet 2011, Banerjee 2012, Salemi 2016). Cases identified by a birth certificate were reviewed by trained abstractors and/or clinicians; false positives were suspected defects that were not confirmed following these verification steps. While PPVs for birth defect indication on birth certificates are generally much better than their sensitivity, there were some notably low findings such as a 49% PPV for anencephaly reported by Salemi et al. The authors noted that anencephaly is the first checkbox on the Florida birth certificate, and it is plausible that the box may be mistakenly confused for the box indicating no defect. Banerjee et al. reported <40% PPV for 5 defects (hydrocephalus, renal agenesis, microcephaly, rectal atresia, and tracheo-esophageal fistula) but did not discuss potential reasons for these findings. Birth certificates do not appear to be a reliable ascertainment source without a case verification system in place for many defects.

Table 3. Positive predictive value of defect indication for select defects on birth certificates following medical record abstraction and/or clinical review

	Boulet et al. Metro Atlanta, 2011	Banerjee et al. Minnesota, 2012	Salemi et al. Florida, 2016
All included defects	96%		87%
By Defect			
Anencephaly	91%	80%	49%
Omphalocele/gastroschisis*		90%	94%
Hydrocephalus		39%	
Spina bifida	96%	100%	74%
Cleft lip/palate	97%	90%	97%
Limb reduction deficiency			55%
Clubfoot	87%	--	
Trisomy 21	97%	92%	87%
Diaphragmatic hernia		50%	
Renal agenesis		38%	
Hypospadias			77%
Microcephaly		33%	
Rectal atresia	100%	25%	
Tracheo-esophageal fistula		0%	
By Class			
Central nervous system		47%	
Musculoskeletal		16%	
Orofacial		90%	
Chromosomal		63%	
Genitourinary		23%	
Cardiovascular		36%	
Gastrointestinal		20%	

* This combined variable is not very useful given the vast epidemiological differences between these two defects

Two studies looked at the predictive value of birth certificates among cases uniquely identified from that source. In Florida, although birth certificates identified 5% of birth defects cases that would have been otherwise missed, over half of those were false positive (not confirmed defects following investigation) (Salemi 2017). In Alaska, birth certificates uniquely identified 12% of potential neural tube defects, but none of those cases were confirmed following investigation (Schoellhorn and Collins 2010). This limited evidence suggests that among cases uniquely identified by birth certificates, the predictive value is poor.

We also reviewed the literature for studies that evaluated if medical abstraction and clinical review were necessary to obtain accurate prevalence estimates. Two studies evaluated the predictive value of cases identified by passive surveillance systems compared to those verified through chart abstraction and clinical review and are summarized below. Another two articles were identified that evaluated accuracy of algorithms to verify cases compared to medical abstraction and review; these results are reviewed in this section.

Positive predictive value of passive surveillance compared to active surveillance

Salemi et al. quantified the agreement between active (with medical record abstraction and clinical review) and passive surveillance systems in Florida counties where both were implemented (Salemi 2012). There was significant variation in accuracy across defects, with PPVs ranging from 96% for gastroschisis to 69% for limb reduction deformities. When considered together, all orofacial clefts had a high PPV (96%) but when types of clefts were considered separately, PPVs were markedly lower (82% to 88%). Similarly, when grouped, diagnoses were relatively accurate for CCHDs (90%) but for specific defects, like hypoplastic left heart syndrome, the PPV was much lower (75%). This suggests that the passive system captures the defect class quite well but often fails to identify the specific defect accurately.

A more recent study in New York reported that most infants identified by active surveillance were captured by passive surveillance (98%); however, the ability of passive surveillance to correctly identify specific defects varied (Howley 2024). Here, the PPV of passive surveillance was >90% for gastroschisis and cleft lip but was <70% for spina bifida, diaphragmatic hernia, truncus arteriosus, tricuspid atresia, hypoplastic left heart syndrome, coarctation of aorta, and pulmonary atresia.

Use of electronic case reporting (eCR) to complete case investigation

An evaluation project on the utility of eCR in birth defects surveillance in Texas demonstrated that the data contained in electronic initial case reports (eICR) are largely not equivalent to that found in the full electronic health record, primarily due to many missing fields (Allred 2025). Variables such as maternal characteristics (e.g., maternal age, prenatal care data), infant clinical characteristics (e.g., gestational age, birth weight), and key postnatal procedural information were more consistently available and detailed in the electronic health record than in the eICR.

Use of algorithms to verify cases

Three articles evaluated the use of algorithms to verify cases to limit staff resources needed for medical record abstraction and review by clinicians. In Florida, the authors compared the PPV of their standard surveillance case definition (suspected cases confirmed by medical record review) to the PPV of more restrictive algorithms (Salemi 2018). Both were compared to a gold standard of medical record abstraction and clinical review. The overall case definition had a PPV of 94%, but PPVs varied by defect from 61% for reduction deformities of lower limb to 96% for gastroschisis. More restrictive algorithms requiring multiple encounters from a single source or encounters from more than one source did increase the overall PPV to 98% or 99% but greatly decreased program's inclusion of

potential cases (range of loss of 58-82%) if used to eliminate cases from inclusion. This decrease in sensitivity would be variable by defect as not all defects result in multiple medical encounters.

A limited study in Alaska found that neural tube defect cases were significantly more likely to be confirmed following medical record review if they were reported from ≥2 sources (PPV 69%) than from just a single source (PPV 23%) (Schoellhorn and Collins 2010).

A recent study in Arizona found that the ICD-10 code for gastroschisis in the hospital discharge database had a PPV of 69% (Contreras 2026). If, however, the ICD-10 code was accompanied by at least one procedure code indicating gastroschisis repair, the PPV increased to 100% (all cases meeting these criteria were confirmed as true cases following medical record review). The findings suggest that the less restrictive method (diagnosis code alone) continue to be used for case identification as 22% of cases confirmed following medical record review only had diagnosis codes included in the discharge data (no procedure codes).

Summary of Evidence for Question 3

The evidence suggests that hospital discharge databases largely reflect accurate birth defects diagnoses. However, many defects do require medical record abstraction and clinical review to ensure that the defects assigned to a particular case are accurate. Further studies are needed to make recommendations on which defects may not require investigation and validation, but the current evidence suggests that gastroschisis and cleft lip might be included in that grouping. Birth certificates appear to capture defects accurately for some defects but not for others; case investigation is needed as a common practice following case identification from a birth certificate. There is some evidence that using algorithms to verify cases increases specificity, but the resulting loss of true cases indicates concerns in it being used exclusively to eliminate cases.

Question 4: Up to what age do birth defects need to be ascertained to obtain accurate prevalence estimates?

One article compared different age-of-diagnosis cut points to an existing standard of inclusion up to age 1 year for 44 defects across several defect classes (Salemi 2015). For this passive surveillance system, restricting case ascertainment to nine months of age resulted in loss of 1.4% of cases; restricting to six months resulted in a loss of 3.2%. However, if only the birth hospitalization were used for case ascertainment, 25% of cases would have been missed. Restricting the age of diagnosis to

nine months had a significant impact on prevalence rate for only one of the 44 defects included in the analysis, hypospadias. Restricting the age of diagnosis to six months significantly decreased prevalence rates for four defects (hydrocephalus without spina bifida, hypospadias, pulmonary valve atresia and stenosis, and single ventricle).

One unpublished abstract was reviewed which assessed what percentage of true cases of 26 defects were added when their age of diagnosis was expanded beyond one year of age (Hodson 2025). Overall, 96% of children with a true birth defect diagnosis were detected during their first year of life. Only 3% of additional cases were identified from age 1-2 years, and an additional 1% of cases were identified from age 3-5. Cases identified within the first year of life ranged by diagnosis from 84% to 100%. For 11 of 26 diagnoses, 100% of true cases were identified in the first year of life. Cleft palate had the lowest percentage of cases detected before 1 year (84%). Prevalence estimates did not change significantly for any of the 26 conditions by including cases detected after 1 year of age. Additionally, diagnoses first assigned to a case beyond 1 year of age were much more likely to be inaccurate after case verification. Of the diagnoses first detected during the child's first year of life, 10% were found to be inaccurate, while 25% of those detected after the first year of life were inaccurate.

Summary of Evidence for Question 4:

There is limited evidence to guide any recommendations about the age of case ascertainment for birth defects surveillance. The available information suggests that restricting case ascertainment to 9 months from 1 year would not have substantial impacts on prevalence estimates. There appears to be limited value in continuing to capture cases diagnosed after one year of birth, given low numbers of additional cases and higher false positive rates.

Question 5: What pregnancy outcomes need to be included to obtain accurate prevalence estimates?

There were limited published articles assessing the impact of capturing non-live birth outcomes on birth defects prevalence. Mai et al. presented national birth defects counts and prevalence by pregnancy outcome for two central nervous system defects and three chromosomal defects (Mai 2019). Non-live birth cases contributed to 67% of the anencephaly cases, 56% for trisomy 18 cases, and 45% for trisomy 13 cases, but only 16% for spina bifida and 13% for trisomy 21. For all five defects, the estimated prevalence was significantly different when non-live births were included.

Another study reported that overall, 6% of cases within their surveillance system are fetal deaths but there was no information about the importance of this data source for particular defects or classes of defect (Tessarolo 2021). One article was identified that assessed the inclusion of other pregnancy losses (e.g., elective terminations at any gestational age and miscarriages <20 weeks) in addition to live births and fetal deaths after 20 weeks (Fothergill 2024). Including these additional outcomes resulted in 30% increase in reported cases and significantly higher prevalence rates overall and for neural tube defects and trisomies 13,18, and 21. However, the authors reported that collecting these data required substantial programmatic effort (e.g., review of free text from prenatal sources, collection of demographic data from other than vital records linkage).

Summary of Evidence for Question 5:

The limited evidence available suggests that data sources that capture non-live births (e.g., fetal death certificates, prenatal diagnostic clinics, etc.) are important case identification sources for accurate surveillance of particular conditions, such as neural tube defects and chromosomal abnormalities. However, in the current state, conducting surveillance for these additional pregnancy outcomes can be logistically difficult, resource intensive, and restricted based on jurisdictional policy or regulation. Expanding routine surveillance outside of live births is likely to remain an enhanced activity not attainable for all programs for all defects. More evidence is needed to determine which defects require case identification for all birth outcomes.

Conclusions

Determining which data sources and criteria to use for birth defects case identification must balance the desire to identify every case while limiting the capture of too many false positive cases that would overburden surveillance program staff. Similarly, determining which cases require enhanced verification procedures must balance the accuracy of defect classification with available program resources. This project was undertaken, in large part, to determine where there is adequate evidence for NBDPN to develop standards and guidance on best practices and to identify the critical gaps in current knowledge base to highlight areas needing research. NBDPN is developing standards around three areas:

1. Which sources (e.g., hospitals, prenatal diagnostic facilities, vital records, specialty clinics, etc.) should be used for identification of birth defects cases?
2. What criteria (ICD-10-CM codes, text descriptions, procedure codes, etc.) should be used to trigger the report of suspected birth defects cases from included sources?

3. Which cases (all, core defects, particular defect classes, etc.) require medical record abstraction and clinical review?

Our review of the literature revealed that there are limited data to guide the development of recommendations for best practices of birth defects surveillance. Much of the published literature is based upon surveillance data from more than ten years ago and may no longer be applicable, particularly as access to electronic health records and electronic case reporting increases within the community. Published data is also derived from only a few programs, limiting the generalizability of the findings. NBDPN may consider supporting research to evaluate birth defects surveillance methodology to allow for the development of evidence-based national guidelines.

References

Atkinson HP, Bergman K, Forestieri N. Evaluation of Data Sources for Critical Congenital Heart Defect Case Ascertainment in North Carolina. *Birth Defects Res*. 2025 Dec;117(12):e70005.

Banerjee E, Zabel E, Alexander B. Comparison of defects identified through Minnesota's Birth Defects Information System and Vital Records, 2006-2008. *J Registry Manag*. 2012;39(1):8-12.

Barnett C, Christiansen J, Mills M, et al. Measuring misclassification and sample bias in passive surveillance systems: Improving prevalence estimates of critical congenital heart defects in state-based passive surveillance systems. *Birth Defects Res*. 2024 Aug;116(8):e2386.

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 Mar-Apr;126(2):186-94.

Contreras D, Miura-Akagi B, Gibbons M, Ramirez GM. Gastroschisis: Is the Combination of Diagnosis and Procedure Code Sufficient for Case-Confirmation? *Birth Defects Res*. 2026 Feb;118(2):e70018.

Fothergill A, Lieberman RF, Nestoridi E, et al. Expanding the Massachusetts Birth Defects Monitoring Program to include additional pregnancy outcomes: Programmatic efforts and impacts on case ascertainment, 2012-2020. *Birth Defects Res*. 2024 Mar;116(3):e2323.

Garg LF, Braun KVN, Knapp MM, et al. Results from the New Jersey statewide critical congenital heart defects screening program. *Pediatrics*. 2013 Aug;132(2):e314-23.

Hodson E, Lolley K, Waldrop J. Impact of Expanding Surveillance beyond One Year of Age for 26 Major Birth Defects. 2025 National Birth Defects Prevention Network Annual Conference. Oral Presentation. Denver, Colorado. July 2025.

Howley MM, Williford E, St. Louis AM, et al. A Comparison of Active and Passive Surveillance Strategies for Selected Birth Defects in New York. *Birth Defects Res*. 2024 Sep;116(9):e2399.

Mai CT, Isenburg JL, Canfield MA, et al. National population-based estimates for major birth defects, 2010-2014. *Birth Defects Res*. 2019 Nov 1;111(18):1420-1435.

National Birth Defects Prevention Network. Guidelines for Conducting Birth Defects Surveillance. Sever, LE, ed. Atlanta, GA: National Birth Defects Prevention Network, Inc., June 2004.

Rutkowski RE, Salemi JL, Tanner JP, et al. Improving the Completeness of Ascertainment in Florida's Birth Defects Surveillance Program: The Impact of Adding Infant Death and Emergency Department Data. *J Registry Manag*. 2015 Fall;42(3):91-102.

Salemi JL, Tanner JP, Kennedy S, et al. A comparison of two surveillance strategies for selected birth defects in Florida. *Public Health Rep*. 2012 Jul-Aug;127(4):391-400.

Salemi JL, Tanner JP, Anjohrin SB, et al. Evaluating Difficult Decisions in Public Health Surveillance: Striking the Right Balance between Timeliness and Completeness. *J Registry Manag*. 2015 Summer;42(2):48-61.

Salemi JL, Tanner JP, Sampat D, et al. The Accuracy of Hospital Discharge Diagnosis Codes for Major Birth Defects: Evaluation of a Statewide Registry with Passive Case Ascertainment. *J Public Health Manag Pract*. 2016 May-Jun;22(3):E9-E19.

Salemi JL, Tanner JP, Sampat D, et al. Evaluation of the Sensitivity and Accuracy of Birth Defects Indicators on the 2003 Revision of the U.S. Birth Certificate: Has Data Quality Improved? *Paediatr Perinat Epidemiol*. 2017 Jan;31(1):67-75.

Salemi JL, Rutkowski RE, Tanner JP, et al. Evaluating the impact of expanding the number of diagnosis codes reported in inpatient discharge databases on the counts and rates of birth defects. *J Am Med Inform Assoc*. 2018 Nov 1;25(11):1524-1533.

Salemi JL, Rutkowski RE, Tanner JP, et al. Identifying Algorithms to Improve the Accuracy of Unverified Diagnosis Codes for Birth Defects. *Public Health Rep*. 2018 May/Jun;133(3):303-310.

Schoellhorn KJ and Collins S. Evaluation of passive surveillance for NTD-affected births in Alaska: a case verification study, birth years 1996-2007. *J Registry Manag*. 2010 Spring;37(1):4-9.

Tanner JP, Salemi JL, Hauser KW, et al. Birth defects surveillance in Florida: infant death certificates as a case ascertainment source. *Birth Defects Res A Clin Mol Teratol*. 2010 Dec;88(12):1017-22.

Tao Z, Wang Y, DiCesare DK, et al. Electronic clinical laboratory reports as a source for ascertaining and confirming chromosomal anomalies reported to the New York State Congenital Malformations Registry. *J Public Health Manag Pract*. 2013 Mar-Apr;19(2):E17-24.

Tessarolo A, Ramirez GM, Contreras D. Assessing the value of reviewing all fetal death records for birth defects surveillance, compared to only those with maternal anomaly codes. *Birth Defects Res*. 2021 Jan 15;113(2):123-127

Thom R and Cho SJ. Birth Defect Case Identification in Stillbirths: Is Using both Fetal Death Reports and Facility-reported ICD-10 Discharge Codes Worthwhile? Breakout Session. 2025 National Birth Defects Prevention Network Annual Conference. Oral Presentation. Denver, Colorado. July 2025.

Wang Y, Cross PK, and Druschel CM. Hospital discharge data: can it serve as the sole source of case ascertainment for population-based birth defects surveillance programs? *J Public Health Manag Pract*. 2010 May-Jun;16(3):245-51.