Improving Case Ascertainment of a Population-Based Birth Defects Registry in New York State Using Hospital Discharge Data

Ying Wang,[1,2]* Monica Sharpe-Stimac,[1] Philip K. Cross,[1] Charlotte M. Druschel,[1,2] and Syni-An Hwang[1,2]

[1]Congenital Malformations Registry, Center for Environmental Health, New York State Department of Health, Troy, NY
[2]School of Public Health, New York State University at Albany, Rensselaer, NY

BACKGROUND: The assessment of the data quality of population-based registration systems is essential to understanding the reliability and usefulness of disease surveillance and research findings resulting from the use of registry data. Since the New York State Congenital Malformations Registry (CMR) uses passive case ascertainment, the completeness of the registry data is an important aspect of the quality of information. This paper presents the results of hospital audits, which were conducted to capture the unreported cases using hospital discharge files, and evaluates the effectiveness of the audits. **METHODS:** Children age 2 years or younger and diagnosed with reportable birth defects for the birth years 1998-2000 were selected from hospital discharge files of all reporting hospitals in the New York Statewide Planning and Research Cooperative System (SPARCS) and matched to the CMR database for the same birth year period. The unmatched reports from the SPARCS hospital discharge files that the CMR possibly missed were sent to hospitals, requesting submission of the missed reports. Two audits on all reporting hospitals in New York State were conducted: 1) 1998 and 1999 birth cohorts audited from June 2000 to March 2002, and 2) 2000 birth cohort audited from November 2001 to November 2002. **RESULTS:** Hospital audits using SPARCS hospital discharge data identified 5,460 reports that the CMR missed for the selected 66 hospitals analyzed. About 86% of these reports had reportable conditions and were added to the CMR, which comprised 21.4% of all reports from the 66 hospitals for the birth years 1998-2000. The number of reports that would have been missed without audits decreased from the 1998 and 1999 birth cohort (25.1%) to the 2000 birth cohort (13.9%). Low reporting rates and, thus, a high percent of added reports, were found for hospitals with a relatively small number of annual reports and for some specific birth defects such as chromosomal anomalies, anencephalus and congenital anomalies of the urinary system. **CONCLUSION:** The current study demonstrates that using hospital discharge data to improve case ascertainment is a valuable and effective method of enhancing birth defect surveillance, particularly for those hospitals with low reporting rates.

Key words: birth defects, congenital malformations registry, hospital discharge data, case ascertainment, audit

INTRODUCTION

Birth defect registries have served as important sources of information when used in surveillance and epidemiological research to investigate the effect of prevention programs (Kallen and Olausson, 2002; Berry et al., 1999) and to identify possible genetic and environmental risk factors for birth defects (Reefhuis et al., 2002; Wu et al., 2004; Manson and Carr, 2003). Assessing the data quality (completeness and accuracy) of population-based registration systems is, therefore, essential to understanding the reliability and usefulness of disease surveillance and research findings resulting from the use of birth defects registry data. For the past decade, studies have been conducted to evaluate the accuracy and completeness of birth defects data by researchers and registry staff (Cronk et al., 2003; Larsen et al, 2003; Wang et al, 2001; Berghold et al., 2001; Wen et al, 2000; Bower et al., 2000; Honein and Paulozzi, 1999; Czeizel, 1997; Schulman and Hahn, 1993).

The Congenital Malformations Registry (CMR) of the New York State Department of Health (NYSDOH) began operations in late 1982. It is one of the largest statewide, population-based birth defects registries in the nation. The CMR receives case reports from hospitals on children two years of age or younger, who were born or reside in New York State and were diagnosed with reportable birth defects. Since the CMR uses passive case ascertainment, ensuring completeness of the registry is an important aspect of the quality of information. Efforts have been made by CMR staff to improve and evaluate the completeness of the registry using various monitoring systems. When matching CMR data to New York State Vital Records files, it was estimated that the number of unreported birth defects cases found in the birth certificate files comprised only 2.1% of all registry cases (Olsen et al., 1996). By using spatial analyses of disease rates, CMR staff were able to identify the areas with lower than expected rates due to hospital underreporting (Forand et al., 2002).

In an attempt to capture unreported cases and improve the completeness of reporting, the CMR started using a new monitoring system in 1995 to audit all reporting hospitals with obstetric or pediatric departments. This system uses hospital discharge data from

*Correspondence to: Ying Wang, PhD, Congenital Malformations Registry, Center for Environmental Health, New York State Department of Health, Flanigan Square, 547 River Street, Troy, NY 12180-2216. E-mail: wxy01@health.state.ny.us
Published online in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/bdra.20208

the New York Statewide Planning and Research Cooperative System (SPARCS) for children born with major birth defects. This paper presents the results of hospital audits using hospital discharge files for the birth years 1998 to 2000 and evaluates the effectiveness of the audits.

METHODS

Data sources

CMR Database: Hospitals and physicians are required to report children born with major birth defects and the narratives of the defects to the CMR. They are not required to report children with only minor malformations. Three options were available for reporting to the Registry: 1) hospitals and physicians could send reports manually on forms provided by the CMR; 2) hospitals could report cases electronically as attachments to their enhanced SPARCS discharge data including specific narratives; and 3) since 2001, hospitals have had the option of reporting via the Internet, using the Health Provider Network (HPN), developed by the NYSDOH as a secure system for electronically collecting and distributing health related data. Twenty-one hospitals had been reporting cases through the enhanced SPARCS system since 1998. The reports comprised 9% of all reports in the CMR for the birth years 1998 to 2000. It was estimated that over 95% of case reports come from hospitals. Annually, the CMR receives birth defects reports for more than 10,000 children of New York State residents, which comprise 4% of live births.

Statewide Planning and Research Cooperative System: SPARCS, implemented by the NYSDOH in 1979, is a comprehensive, integrated information system available to assist hospitals and organizations in the health care industry with health care resource planning, financial analysis, decision making, and surveillance of New York State hospital and ambulatory surgery services and costs. SPARCS receives, processes, stores, and analyzes the inpatient hospitalization data from all facilities in New York State and ambulatory surgery data from hospital-based ambulatory surgery services and all other facilities providing ambulatory surgery services. Each health care provider submits its SPARCS data, as mandated, in the uniform, computer-readable format described in the Universal Data Set. The diseases in the SPARCS data are coded using the International Classification of Disease (ICD-9) coding system. SPARCS data is estimated to be approximately 99% complete based on previous reporting history of the hospitals (New York State Department of Health, 2001).

Hospital Auditing – Active Case Ascertainment of the Missed Reports

CMR-SPARCS Matching: A data set containing all children age 2 years or younger and diagnosed with major birth defects for the birth year period of 1998 to 2000 was created from SPARCS yearly hospital discharge files. This data file was matched to the CMR records for the same birth year period. A Visual FoxPro application was developed to perform the match using identifying variables such as the hospital's Permanent Facility Identifier, infant's date of birth and medical record number, mother's medical record number, and the infant's patient control number. The matches were conducted without using the variables such as patient's name and address, since the SPARCS hospital database does not contain these variables. The records found in the SPARCS hospital discharge files but not in the CMR database were defined as unmatched/missed reports and were sent to hospitals for audits.

Hospital audits: Children with major birth defects are reportable up to two years of age. An analysis of CMR data for the birth years 1991 to 1995 revealed that about 90% of children reported to the Registry were diagnosed in the first six months of life (CMR unpublished data). To facilitate the audit process, CMR staff begin to audit hospitals 12 to 24 months after the reporting period for each year of birth. The lists of unmatched/missed records from the SPARCS discharge data were sent to the hospitals. A cover letter was sent to explain the rationale for auditing, provide brief instructions for selecting records with reportable conditions and set a deadline for submitting the missed reports. Also included were detailed information on diagnosis and demographic and maternal characteristics. Since the ICD-9 codes used in SPARCS discharge data may contain both reportable and non-reportable anomalies, hospitals need to determine if children should be reported, using the list of reportable anomalies provided by the CMR.

Using SPARCS hospital discharge data, CMR staff conducted two audits on all reporting hospitals in the State of New York for children born between January 1 of 1998 and December 31 of 2000: 1) 1998 and 1999 birth cohort audited from June 2000 to March 2002, and 2) 2000 birth cohort audited from November 2001 to November 2002. The missed reports were submitted by the hospitals and added to the CMR. A report of the same child with one or more new diagnoses was added to the database as a new report.

Table 1. Summary of audit results from 66 selected hospitals. Birth years: 1998-2000

	Birth year cohort			
	1998 & 1999	2000	Total	(%)
Number of reports to the CMR before audit	14,589	7,323	21,912	
Number of reports the CMR missed/audited [a]	4,169	1,291	5,460	(100.0)
Number of new reports submitted to the CMR	3,667	1,020	4,687	(85.8)
Number of reports with non-reportable conditions [b]	502	271	773	(14.2)
Percent added to the CMR from audits [c]	25.1	13.9	21.4	

[a] Reports that were not in the CMR but in the SPARCS
[b] Reports not required when only minor malformations occur without a major anomaly
[c] The number of reports the CMR received due to audits divided by the number of reports before audits (for instance, 4,687/(21,912) = 21.4%)

DATA ANALYSIS

Of all 180 reporting hospitals audited, the 21 hospitals that submit some or all of their reports through the SPARCS reporting system were excluded from the analysis. The audit lists were used to distinguish which reports were added to the CMR as a result of the audit. Since the audit lists were missing for 12 hospitals (16% of the CMR reports, 1998-2000), these hospitals were excluded from data analysis. To reduce the workload of data entry, the 81 hospitals which had less than 100 reports for the three birth years audited (7% of the CMR reports, 1998-2000) were also excluded from the analysis. Finally, data from the remaining 66 hospitals (68% of the CMR reports, 1998-2000) were included in data analysis.

Summary statistics, simple and stratified, were generated using the SAS software package (SAS Institute Inc., Cary, NC). The 95% confidence interval (CI) for a sample proportion was calculated based on a normal distribution model and was used for statistical significance testing. The 95% CI is constructed such that 95% of the intervals will include the true proportion. It reflects how much sampling fluctuation a statistic will show.

RESULTS

Overall, 21,912 reports were received from the 66 hospitals for the study period, 1998-2000 (Table 1). By matching with SPARCS hospital discharge data, 5,460 reports were identified as missing from the CMR. Of these missing reports, 4,687 (85.8%) with reportable anomalies were then submitted to the CMR by the hospitals. This increased the total number of reports by 21.4% for the three years audited. The remaining 773 reports were determined to be non-reportable by either the hospitals or the CMR. As can be seen from Table 1, the number of added reports that would have been missed without audits decreased from the 1998 and 1999 birth cohort (25.1%) to the 2000 birth cohort (13.9%).

Hospitals were grouped into four categories according to the percent of reports to the CMR (Table 2). Hospitals that reported less than 0.3% of all reports from the 66 hospitals had the highest percent of re-

Table 2. Summary of audit results from 66 selected hospitals by the percent cases reported to the CMR. Birth year: 1998-2000

Percent of reports to CMR[a]	Number of Hospitals	Reports to CMR before audit	Reports missed/ audited [b]	Reports Added from audit	% added to CMR by birth year audited		Total percent added to CMR			
					1998 & 1999	2000	%	95% CI [c]		
<= 0.3%	13	1,052	493	407	43.4	30.0	38.7	35.7	-	41.6
0.4 - 0.6%	21	3,223	1,023	796	26.8	20.7	24.7	23.2	-	26.2
0.7 - 1.3%	20	5,919	2,180	1743	36.2	16.7	29.4	28.3	-	30.6
1.4 - 7.2%	12	11,718	2,049	1741	17.7	8.8	14.9	14.2	-	15.5
TOTAL	**66**	**21,912**	**5,460**	**4,687**	**25.1**	**13.9**	**21.4**	**20.8**	-	**21.9**

[a] Hospitals were grouped into four categories according to the percent of reports to the CMR, i.e., the number of reports submitted to the CMR from each hospital divided by total reports received from the 66 hospitals
[b] Reports that were not in the CMR when compared to the SPARCS data
[c] The 95% confidence interval was calculated based on Normal distribution model

ports added (38.7%; 95%CI, 35.7%-41.6%). On the other hand, hospitals that reported greater than 1.3% of all reports from the 66 hospitals had the lowest percent of added reports (14.9%; 95%CI, 14.2%-15.5%). The results (data not shown) from stratified analyses on hospital's geographic location showed that hospitals in the Western New York area had the smallest number of missed reports (5%). Higher percent of added reports (30%) and, thus, lower reporting rates were found for hospitals located in the Finger Lakes region and New York City area.

Stratified analyses on birth defect categories were performed to examine if specific birth defects have any effect on hospital reporting (Table 3). The percent of reports that would have been missed without audits ranges from 2.8% (other congenital anomalies of limbs) to 40.6% (chromosomal anomalies). When comparing reports with a single malformation to those with multiple malformations, the percent of reports that would have been missed without audits was significantly higher when only one malformation existed.

DISCUSSION

Using SPARCS hospital discharge data, the CMR staff identified 5,460 reports that the CMR missed for the selected 66 hospitals analyzed during the birth years 1998-2000. About 86% of these reports had reportable conditions and were added to the CMR as new reports, which comprised 21.4% of all reports. Hospital audits resulted in not only added new cases to the CMR but also improved reporting for subsequent years, probably due to hospitals' positively reacting to the audits. Auditing hospitals by CMR staff

Table 3. Summary of audit results from 66 selected hospitals by birth defect category. Birth years: 1998-2000

ICD-9	Birth defects category	Reports to CMR before audit	Reports from audit	Percent added to CMR %	95% CI [a]
740	Anencephalus and similar anomalies	19	7	36.8	15.2 - 58.5
741	Spina bifida	144	36	25.0	17.9 - 32.1
742	Other congenital anomalies of nervous system	1,198	377	31.5	28.8 - 34.1
743	Congenital anomalies of eye	271	46	17.0	12.5 - 21.4
744	Congenital anomalies of ear, face and neck	208	32	15.4	10.5 - 20.3
745	Bulbus cordis anomalies and anomalies of cardiac septal closure	4,459	1,084	24.3	23.1 - 25.6
746	Other congenital anomalies of heart	883	170	19.3	16.7 - 21.9
747	Other congenital anomalies of circulatory system	1,125	90	8.0	6.4 - 9.6
748	Congenital anomalies of respiratory system	641	78	12.2	9.6 - 14.7
749	Cleft palate and/or cleft lip	675	146	21.6	18.5 - 24.7
750	Other congenital anomalies of upper alimentary tract	969	275	28.4	25.5 - 31.2
751	Other congenital anomalies of digestive system	615	158	25.7	22.2 - 29.1
752	Congenital anomalies of genital organs	3,619	574	15.9	14.7 - 17.1
753	Congenital anomalies of urinary system	1,337	493	36.9	34.3 - 39.5
754	Congenital musculoskeletal deformities	1,828	260	14.2	12.6 - 15.8
755	Other congenital anomalies of limbs	1,555	44	2.8	2.0 - 3.7
756	Other congenital musculoskeletal anomalies	499	158	31.7	27.6 - 35.7
757	Congenital anomalies of the integument	77	5	6.5	1.0 - 12.0
758	Chromosomal anomalies	271	110	40.6	34.7 - 46.4
759	Other and unspecified congenital anomalies	174	35	20.1	14.2 - 26.1
	All other congenital malformations	1,345	509	37.8	35.3 - 40.4
	Reports with only one malformation	15,423	3,457	22.4	21.8 - 23.1
	Reports with more than one malformation	6,489	1,230	19.0	18.0 - 19.9
	TOTAL	**21,912**	**4,687**	**21.4**	**20.8 - 21.9**

[a] The 95% confidence interval was calculated based on Normal distribution model

sent a message to reporting hospitals that both the quality and the quantity of their reports are closely monitored. The effectiveness of using secondary databases to improve the data ascertainment has been shown in recent studies on utilizing hospital discharge data to enhance cancer registries (Middleton et al, 2000; Lang et al, 2003; Penberthy et al, 2003); 12 %- 21% of cancer case reports were added by hospital discharge files relative to the number routinely captured by the registry.

The identified reports for audits consisted of 14% false positives that had non-reportable conditions. False positives in the hospital discharge data had been evaluated in epidemiologic studies (Callif-Daley, et al, 1995; Lanska, et al, 1995; Porta, et al, 2000). Callif-Daley, et al, reported that two-thirds of all false positives were due to the miscoding of correctly diagnosed anomalies, and another quarter were clearly contradicted in notes easily available before the patients were discharged.

Underreporting for some specific birth defect categories such as chromosomal anomalies, anencephalus and congenital anomalies of the urinary system, were found in the current study. Complete hospital reporting of some specific birth defects may be problematic. For instance, Down syndrome, which comprised more than 70% of chromosomal anomalies, is not always observed at birth and hospitals could mistakenly believe that pediatric cases are not reportable to the CMR. Moreover, many of the infants with anencephaly die shortly after birth and hospital staff may be uncertain about whether or not the cases should be reported to the CMR.

It should be noted that 66 hospitals (68% of all reports in CMR) out of 180 hospitals audited were included in the data analysis for evaluating the effectiveness of the audits. As found in the current study, hospitals with a smaller number of reports annually were more likely to underreport their cases compared to hospitals with a relatively larger number of reports annually. Excluding the 81 hospitals with less than 100 reports for birth years 1998-2000 from the analysis may have resulted in lower counts of total reports added to the CMR. On the other hand, the hospitals that report electronically through SPARCS usually have higher reporting rates. These hospitals are also relatively small in size, since the number of births from these hospitals comprised less than 10% of all live births for the birth years of 1998-2000. Thus, the overall effect of excluding some hospitals from the analysis might not make a significant difference in our findings.

Matching CMR reports to hospital discharge data could result in false negatives, i.e., the reports had been submitted to the CMR but were not matched. They were counted as missing. The hospitals would be asked to submit these reports to the CMR and they could be entered into the CMR's database as new reports. However, these duplicate reports are not double counted, an auto-match program checks each added report against existing reports and duplicate reports are removed. It was estimated that the number of false negatives is less than 1% of all added reports from audits.

About 15% of the CMR reports were not matched to SPARCS hospital discharge data. Some of these reports may have been reported to SPARCS or the CMR with errors in one or more of the matching variables. In addition, some CMR reports may be outpatient cases from reporting hospitals that have outpatient clinics and ambulatory surgery services and from physicians' offices. Malformations not diagnosed at birth and treated only in outpatient facilities will not be included in hospital inpatient discharge files. Although reports of all children diagnosed with a reportable birth defect is mandatory for hospitals and physicians, not all hospitals with outpatient clinics and ambulatory surgery services comply. Approaches to increasing the outpatient case reporting from hospitals and physicians should be pursued to ensure the completeness of the CMR. Outpatient data has been recently added to the SPARCS dataset and will be explored in the future.

In conclusion, the current study demonstrates that hospital audits using SPARCS hospital discharge data to improve CMR case ascertainment is an effective approach to enhancing birth defect surveillance, particularly for those hospitals with lower reporting rates. Auditing hospitals using the SPARCS hospital discharge data improved the completeness of the CMR.

LITERATURE CITED

Berry RJ, Li Z, Erickson JD, Li S, Moore CA, Wang H, Mulinare J, Zha P, Wong LY, Gindler J, Hong SX, Correa A. 1999. Prevention of neural tube defects with folic acid in China. China U.S. Collaborative Project for Neural Tube Defect Prevention. N Engl J Med 341(20):1485-90.

Bergold A, Strongnegger WJ, Wernecke KD. 2001. A model and application for estimating completeness of registration. Methods Inf Med. 40(2):122-6.

Bower C, Silva D, Henderson TR, Ryan A, Rudy E. 2000. Ascertainment of birth defects: the effect on completeness of adding a new source of data. J Paediatr Child Health 36(6):574-6.

Callif-Daley FA, Huether CA, Edmonds LD. 1995. Evaluating false positives in two hospital discharge data sets of the

Birth Defects Monitoring Program. Public Health Rep 110(2):154-60.

Cronk CE, Malloy ME, Pelech AN, Miller RE, Meyer SA, Cowell M, McCarver DG. 2003. Completeness of state administrative databases for surveillance of congenital heart disease. Birth Defects Res Part A Clin Mol Teratol 67(9):597-603.

Czeizel AE. 1997. First 25 years of the Hungarian congenital abnormality registry. Teratology 55(5):299-305.

Forand SP, Talbot TO, Druschel C, Cross PK. 2002. Data quality and the spatial analysis of disease rates: congenital malformations in New York State. Health Place 8(3):191-9.

Honein MA, Paulozzi LJ. 1999. Birth defects surveillance: assessing the "gold standard". Am J Public Health 89(8):1238-40.

Kallen BA, Olausson PO. 2002. Use of folic acid and delivery outcome: a prospective registry study. Reprod Toxicol 16(4):327-32.

Lang K, Magi M, Aareleid T. 2003. Study of completeness of registration at the Estonian cancer registry. Eur J Cancer Prev 12(2):153-6.

Lanska MJ, Lanska DJ, Baumann RJ. 1995. A population-based study of neonatal seizures in Fayette County Kentucky: comparison of ascertainment using different health data systems. Neuroepidemiology 14(6):278-85.

Larsen H, Nielsen GL, Bendsen J, Flint C, Olsen J, Sorensen HT. 2003. Predictive value and completeness of the registration of congenital abnormalities in the three Danish population-based registries. Scand J Public Health 31(1):12-6.

Manson JM, Carr MC. 2003. Molecular epidemiology of hypospadias: review of genetic and environmental risk factors. Birth Defects Res Part A Clin Mol Teratol 67(10):825-36.

Middleton RJ, Gavin AT, Reid JS, O'Reilly D. 2000. Accuracy of hospital discharge data for cancer registration and epidemiological research in Northern Ireland. Cancer Causes Control 11(10):899-905.

New York State Department of Health. 2001. Statewide planning and research cooperative system annual report. Bureau of Biometrics, Albany, NY.

Olsen CL, Polan AK, Cross PK. 1996. Case ascertainment for state-based birth defects registries: characteristics of unreported infants ascertained through birth certificates and their impact on registry statistics in New York State. Paediatr Perinat Epidemiol 10(2):161-74.

Penberthy L, McClish D, Pugh A, Smith W, Manning C, Retchin S. 2003. Using hospital discharge files to enhance surveillance. Am J Epidemol 158(1):27-34.

Porta M, Costafreda S, Malats N, Guarner L, Soler M, Gubern JM, Garcia-Olivares E, Andreu M, Salas A, Corominas JM, Alguacil J, Carrato A, Rifa J, Real FX. 2000. Eur J Epidemiol 16(6):533-41.

Reefhuis J, deJong-van den Berg L, Cornel MC. 2002. The use of birth defect registries for etiological research: a review. Community Genet 5:13-32.

Schulman J, Hahn JA. 1993. Quality control of birth defect registry data: a case study. Public Health Rep 108(1):91-8.

Wang FL, Gabos S, Sibbald B, Lowry RB. 2001. Completeness and accuracy of the birth registry data on congenital anomalies in Alberta, Canada. Chronic Dis Can 22(2):57-66.

Wen SW, Rouleau J, Lowry RB, Kinakin B, Anderson-Redick S, Sibbald B, Turner T. 2000. Congenital anomalies ascertained by two record systems run in parallel in the Canadian province of Alberta. Can J Public Health 91(3):193-6.

Wu J, Wang J, Meng B, Chen G, Pang L, Song X, Zhang K, Zhang T, Zheng X. 2004. Exploratory spatial data analysis for the identification of risk factors to birth defects. BMC Public Health 4(1):23.