Published in final edited form as:
Pediatr Blood Cancer. 2014 June ; 61(6): 1129–1131. doi:10.1002/pbc.24930.

Establishing a High-Risk Neuroblastoma Cohort Using the Pediatric Health Information System Database

Ami V. Desai, MD[1,2,3], **Marko Kavcic, MD**[1], **Yuan-Shung Huang, MS**[4], **Nicole Herbst, BS**[5], **Brian T. Fisher, DO, MSCE**[2,4,6,7], **Alix E Seif, MD, MPH**[1,7], **Yimei Li, PhD**[1,2,7], **Sean Hennessy, PharmD, PhD**[2,3,5], **Richard Aplenc, MD, PhD**[#1,2,3,4,7], and **Rochelle Bagatell, MD**[#1,7]

[1]Division of Oncology, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

[2]Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA

[3]Center for Pharmacoepidemiology Research and Training, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA

[4]Center for Pediatric Clinical Effectiveness, The Children's Hospital of Philadelphia, Philadelphia, PA

[5] Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA

[6]Division of Infectious Diseases, The Children's Hospital of Philadelphia, Philadelphia, PA

[7]Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA

[#] These authors contributed equally to this work.

Abstract

International Classification of Diseases, 9th Revision (ICD-9) code(s) for neuroblastoma do not exist, preventing identification of these patients in administrative databases. To overcome this challenge, a 3-step algorithm, using ICD-9 codes, exclusion criteria, and manual review of chemotherapy billing data, was utilized to assemble a high-risk neuroblastoma cohort (n=952) from the Pediatric Health Information System (PHIS) Database and validated at a single institution (sensitivity 89.1%; positive predictive value (PPV) 96.1%). This cohort provides a data source for future comparative effectiveness and clinical epidemiology studies in high-risk neuroblastoma patients.

Keywords

Neuroblastoma; Cohort Validation; Administrative Database

Correspondence: Ami Desai, MD 3501 Civic Center Boulevard Room 4020 Philadelphia, PA 19104 Phone: 267-425-2119 Fax: 267-425-5839 desaia@email.chop.edu.

INTRODUCTION

Much remains to be learned about the impact of current-era treatment regimens for high-risk neuroblastoma (HR-NBL) on the patient experience. Administrative databases such as the Pediatric Health Information System (PHIS) database are well-established resources for addressing clinical questions. [1] Methods for assembly of pediatric leukemia cohorts using PHIS have been established [2, 3]. Leukemia cohort construction relied upon use of disease-specific ICD-9 codes; however neuroblastoma-specific ICD-9 codes do not exist. We sought to establish a HR-NBL cohort in PHIS using an alternative methodology.

MATERIALS AND METHODS

Data Source

The PHIS database includes inpatient data from 43 not-for-profit tertiary pediatric hospitals affiliated with Children's Hospital Association. Data include demographics, dates of service, discharge disposition, payor information, ICD-9 diagnosis and procedure codes, and billing data for medications, laboratory tests, imaging procedures, clinical services, and supplies. Patients are assigned a unique identifier and can be followed from one admission to another. [1, 2]

Study Cohort Assembly

The source population consisted of patients first admitted to a PHIS hospital between January 1, 1999 and February 23, 2011. Patients with potentially relevant ICD-9 codes **(Table I Step 1)** were identified. Patients were excluded if no unique medical record number could be identified, if billing data for chemotherapy used in standard North American HR-NBL induction regimens were absent, or if billing data consistent with an alternative diagnosis were present. Chemotherapy data for up to 365 days from index hospital admission were then reviewed manually.

Five HR-NBL induction regimens were used in PHIS centers during the relevant time period: Children's Cancer Group 3891 [4], Pediatric Oncology Group 9640 [5], Children's Oncology Group (COG) A3973 [6], COG ANBL00P1 [7], and COG ANBL0532 [8]. Pharmacy billing data for each patient were compared to these regimens. Inclusion in the cohort required ≥3 courses of chemotherapy that did not deviate significantly from courses prescribed in a given regimen. Gaps in therapy, minor changes in order of chemotherapy courses, and switching of chemotherapy regimen did not preclude inclusion as long as the requisite 3 cycles were identified. Four authors (MK, NH, AVD, RB) conducted the primary chemotherapy review. To assess inter-rater reliability, chemotherapy records for 10% of patients previously reviewed by others were re-reviewed (RB). The correlation of independent reviews was assessed to measure level of agreement.

Study Cohort Validation

Medical records of patients with *de novo* HR-NBL treated at the Children's Hospital of Philadelphia (CHOP) between January 1, 2004 and February 23, 2011 were reviewed to identify the chemotherapy regimen administered. CHOP patients included in the final PHIS

cohort were matched to patients in the CHOP registry to determine the sensitivity and positive predictive value (PPV) of the process used for cohort assembly.

Negative predictive value (NPV) and specificity were not determined. The PHIS database includes all inpatient pediatric admissions from contributing hospitals and contains a large number of admissions for patients who do not have HR-NBL (true negatives). NPV would be high simply due to the overall size of the database. Similarly, the total number of non-neuroblastoma admissions at CHOP during the 7-year time period was large; therefore, the resultant specificity would be high.[2, 3]

RESULTS

Using the pre-specified group of ICD-9 diagnosis codes, 8117 potential HR-NBL patients were identified in the PHIS database. After applying exclusion criteria, 3390 patients remained. Upon completion of manual chemotherapy review, the cohort included 952 patients (**Table I**). The median age was 3.1 years (range: birth-26.5 years); 78% of patients were between 19 months and 9 years old **(Table II)**. Fifty-five percent were male; 70% were white **(Table II)**.

The sensitivity and PPV of each of the steps in cohort generation were calculated **(Table I)**. Using ICD-9 codes alone, sensitivity was 100% but PPV was only 11.4%. After applying exclusion criteria and completing manual chemotherapy review, PPV increased to 96.1% (95% CI; 86.5%-99.5%) with a modest reduction in sensitivity. Second review of chemotherapy data for 339 patients demonstrated the reproducibility of results of chemotherapy review; concordance with initial review was 99%.

DISCUSSION

Using a national pediatric database, we have assembled a cohort of 952 children with HR-NBL for whom detailed information regarding resource utilization is available. Because neuroblastoma can arise in multiple sites[9], use of a diagnostic coding system based on anatomic site has substantial limitations. In some cases, the diagnosis of neuroblastoma may be made without a primary site biopsy and the ICD-9 code for adrenal neoplasm (194.0) may not be used for patients who have neuroblastoma. Conversely, the ICD-9 code for adrenal neoplasm includes other diagnoses (adrenocortical carcinomas, pheochromocytomas, adrenal adenomas, and ganglioneuromas) and therefore the 194.0 code may be used for patients who do not have neuroblastoma.

Use of multiple ICD-9 diagnosis codes as a first step resulted in a large initial cohort and a low PPV. Manual chemotherapy review and the requirement for 3 cycles of regimen-specific chemotherapy (rather than the single cycle of chemotherapy required for inclusion in leukemia cohorts) were critically important for assembly of the final cohort. Manual review was necessary because the chemotherapy delivered in several widely used regimens (COG A3973, ANBL00P1, ANBL0532) consists of a combination of agents (vincristine/doxorubicin/cyclophosphamide) that can also be used to treat other pediatric malignancies. This requirement may have resulted in exclusion of patients that truly had HR-NBL, including patients with early disease progression or patients with therapy-ending toxicities

that occurred prior to completion of 3 cycles of recognizable neuroblastoma chemotherapy. While the potential for bias due to exclusion of these patients should be recognized, the representative nature of this cohort is reflected in the remarkable similarity between the demographics of the PHIS HR-NBL cohort and the patient population of a recent European trial for HR-NBL[10].

The successful assembly of a large HR-NBL cohort using a validated algorithm provides a data source for clinical epidemiology studies, and the cohort can be leveraged to answer comparative effectiveness questions. However, limitations must be acknowledged. Manual review of pharmacy data was needed to overcome the lack of ICD-9 codes specific for neuroblastoma, but it is possible that some chemotherapy data could be missing, including data regarding chemotherapy administered to outpatients. Chemotherapy delivery and billing practices may vary among PHIS institutions, and assignment of ICD-9 codes and pharmacy data completeness may differ by center. Therefore, patients from some centers could be non-randomly under-represented. Finally, validation of our method was completed at a single institution. It is possible that the sensitivity and PPV would vary if additional centers were included in the validation.

Because advancements in healthcare in recent years have been accompanied by dramatic increases in healthcare expenditures, the need to define the most safe and cost-effective approaches to care has become apparent. This dataset can be used by the neuroblastoma community for multi-center clinical epidemiology and comparative effectiveness studies that have become increasingly important in an evolving healthcare economy. This database can be used to address issues such as antibiotic and other medication use, length of hospital stays, and transplant-associated resource utilization. Data regarding late effects and long-term outcomes are not uniformly captured in this inpatient database at present, but may become available in the future. Development of similar cohorts of patients with high-risk neuroblastoma treated in other countries may permit additional comparisons of outcomes and resource utilization in other healthcare systems.

ACKNOWLEDGEMENTS

This study was supported by NIH/NICHD Pediatric Pharmacoepidemiology Training Program T32-HD-06456702 (A.V.D.), NIH Grant R01CA165277 (R.A.), Slovenian Ministry of Education, Science, Culture, and Sport Grant J3-4220 (M.K.), and the Alex's Lemonade Stand Foundation, Pediatric Oncology Student Training Program (N.H.).

REFERENCES

1. Fisher, BT.; Lindenauer, PK.; Feudtner, C. In-hospital Databases.. In: Strom, B.; Kimmel, SE.; Hennessy, editors. Pharmacoepidemiology. John Wiley & Sons, Ltd; UK: 2012. p. 244-258.
2. Fisher BT, Harris T, Torp K, et al. Establishment of an 11-year Cohort of 8733 Pediatric Patients Hospitalized at United States Free-standing Children's Hospitals With De Novo Acute Lymphoblastic Leukemia from Health Care Administrative Data. Med Care. 2012 doi: 10.1097/MLR.0b013e31824deff9.
3. Kavic M, Fisher BT, Torp K, et al. Brief Report: Assembly of a Cohort of Children Treated for Acute Myeloid Leukemia at Free-Standing Children's Hospitals in the United States Using an Administrative Database. Pediatr Blood Cancer. 2013; 60(3):508–511. [PubMed: 23192853]
4. Matthay KK, Reynolds CP, Seeger RC, et al. Long-Term Results for Children With High-Risk Neuroblastoma Treated on a Randomized Trial of Myeloablative Therapy Followed by 13-cis-

Retinoic Acid: A Children's Oncology Group Study. J Clin Oncol. 2009; 31(5):1007–1013. [PubMed: 19171716]

5. Granger M, Grupp SA, Kletzel M, et al. Feasibility of a Tandem Autologous Peripheral Blood Stem Cell Transplant Regimen for High Risk Neuroblastoma in a Cooperative Group Setting: A Pediatric Oncology Group Study: A Report from the Children's Oncology Group. Pediatr Blood Cancer. 2012; (59):902–907. [PubMed: 22744917]
6. Kreissman SG, Seeger RC, Matthay KK, et al. Purged versus non-purged peripheral blood stem-cell transplantation for high-risk neuroblastoma (COG A3973): a randomised phase 3 trial. Lancet Oncol. 2013 doi:10.1016/S1470-2045(13)70309-7.
7. Seif AE, N.A. Baker DL, et al. A pilot study of tandem high-dose chemotherapy with stem cell rescue as consolidation for high-risk neuroblastoma: Children's Oncology Group study ANBL00P1. Bone Marrow Transplant. 2013; 48:947–952. [PubMed: 23334272]
8. Park JR, Scott JR, Stewart CF, et al. Pilot induction regimen incorporating pharmacokinetically guided topotecan for treatment of newly diagnosed high-risk neuroblastoma: a Children's Oncology Group study. J Clin Oncol. 2011; 29(33):4351–4357. [PubMed: 22010014]
9. Brodeur, GM.; Hogarty, MD.; Mosse, YP.; Maris, JMM. Neuroblastoma.. In: Pizzo, PA.; Poplack, DG., editors. Principles and Practice of Pediatric Oncology. Lippincott Williams & Wilkins; Philadelphia: 2011. p. 886-922.
10. Pearson AP, Pinkerton CR, Lewis IJ, et al. High-dose rapid and standard induction chemotherapy for patients aged over 1 year with stage 4 neuroblastoma: a randomised trial. Lancet Oncol. 2008; 9:247–56. [PubMed: 18308250]

Table I

Three-Step Algorithm Used to Assemble a High-risk Neuroblastoma Cohort and Validation Results

		NBL* Cohort	Validation at CHOP** [Gold Standard=CHOP registry]				
			True Positives	False Positives	False Negatives	Positive Predictive Value (95% CI)	Sensitivity (95% CI)
Step 1	ICD-9 Codes (194.0, 194.8, 194.9, 158, 158.8, 158.9, 192.9, 196.XX, 199.XX)	8117	55	426	0	11.4% (8.7%-14.6%)	100% (93.5%-100%)
Step 2	Exclusion Criteria						
	• No unique MRN#	8116	55	425	0	11.5% (8.7%-14.6%)	100% (93.5%-100%)
	• No pharmacy billing records for induction chemotherapy	4670	55	170	0	24.4% (18.9%-30.6%)	100% (93.5%-100%)
	• Claim records for impermissible chemotherapy	3390	54	128	1	29.7% (23.1%-36.9%)	98.2% (90.3%-99.9%)
Step 3	Chemotherapy Review	952	49	2	6	96.1% (86.5%-99.5%)	89.1% (77.8%-95.9%)

*NBL=neuroblastoma

**CHOP=The Children's Hospital of Philadelphia

#MRN=medical record number

Table II Characteristics of Patients from the PHIS-Assembled High-risk Neuroblastoma Cohort and a European High-risk Neuroblastoma Study

	PHIS* -Assembled High-risk Neuroblastoma Cohort No (%)	European High-risk Neuroblastoma Study [10] Standard Arm No (%)
Age Group		
<12 mo	50 (5.2)	N/A
12 mo-18 mo	110 (11.5)	17 (13)
19 mo-9 yrs	742 (78)	111 (84)
10-15 yrs	37 (3.9)	2 (2)
16-20 ys	9 (1)	2 (2)
>20 yrs	4 (0.4)	N/A
Sex		
Male	524 (55)	75 (57)
Female	428 (45)	57 (43)
Race		
White	670 (70.4)	N/A
Black	136 (14.3)	N/A
Other**	146 (15.3)	N/A
Total	952 (100)	132 (100)

*PHIS=Pediatric Health Information Systems

**Other=Asian, American-Indian, Other, Unknown