Leukemia Research 38 (2014) 71–75

Contents lists available at ScienceDirect

Leukemia Research

journal homepage: www.elsevier.com/locate/leukres

High rate of uncaptured myelodysplastic syndrome cases and an improved method of case ascertainment

Christopher R. Cogle[a,*], Michelle R. Iannacone[b], Daohai Yu[c,f], Ashley L. Cole[b], Iman Imanirad[a], Lulu Yan[c], Jill A. MacKinnon[d], Alan F. List[e], Dana E. Rollison[b]

[a] Division of Hematology and Oncology, Department of Medicine, College of Medicine, University of Florida, Gainesville, FL, United States
[b] Cancer Epidemiology Program, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, United States
[c] Department of Biostatistics, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, United States
[d] Florida Cancer Data System, University of Miami School of Medicine, Miami, FL, United States
[e] Division of Hematologic Malignancies, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, United States
[f] Temple University School of Medicine, Philadelphia, PA, United States

ARTICLE INFO

Article history:
Received 24 June 2013
Received in revised form 1 October 2013
Accepted 26 October 2013
Available online 5 November 2013

Keywords:
Myelodysplastic syndromes
Epidemiology

ABSTRACT

The myelodysplastic syndromes (MDS) are often diagnosed in outpatient clinics and may be under-reported to state cancer registries, which predominantly rely on hospital records and laboratory reports. We used a new method of cancer case capture to determine the rate of missed cases and estimate a more accurate incidence of MDS. Using a unique keyword algorithm, we queried all electronic pathology (E-path) reports sent to the state of Florida cancer registry in 2006 to identify potential MDS cases. A stratified, random sample of E-path reports was then reviewed to confirm diagnosis and assign MDS subtype. Characteristics were compared between captured and uncaptured MDS cases. 7111 E-path reports with MDS keyword hits were identified, of which only 18% linked to a registered MDS case, 47% linked to a different cancer, and 34% did not link with any record. Case review of a stratified, random sampling of 285 individuals led to the discovery that uncaptured cases made up 37.7% of the total true MDS cases in 2006. It is estimated that the true incidence of MDS is 5.3 individuals out of 100,000, compared to previous reports of 3.3 out of 100,000. Uncaptured MDS cases were younger and more likely to have information in the pathology report facilitating MDS subtype assignment. Only two-thirds of true MDS cases are captured in Florida using current case-finding mechanisms. Application of a keyword search strategy to identify cases among E-path reports is a feasible technique to improve MDS case ascertainment.

1. Introduction

The myelodysplastic syndromes (MDS) include a diverse group of hematological disorders characterized by dysplastic and ineffective hematopoiesis. MDS is challenging to diagnose and classify, especially in cases when blast percentage is not increased [1]. Although approximately 30% of MDS cases progress to acute myeloid leukemia (AML), most patients with MDS die of complications associated with cytopenias. In the ninth edition of the International Classification of Diseases (ICD-9), MDS was coded as a disease of the blood and blood forming organs, but was reclassified as a neoplasm in the tenth edition (ICD-10) and the corresponding ICD for Oncology Third Edition (ICD-O-3), the classification system used by population-based cancer registries. As a result, MDS became a reportable malignancy to population-based registries for the first time in 2001, the year ICD-O-3 was implemented worldwide.

Recently, data from the North American Association of Central Cancer Registries (NAACCR), which includes registries reporting to the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program, provided the first opportunity to investigate patterns in the incidence and survival of patients with MDS in the United States [2–4]. Based on these data, the annual age-adjusted incidence for MDS in the United States is estimated at 3.3 per 100,000 [4]. However, several lines of evidence suggest that reported MDS incidence is considerably underestimated. Within SEER data, incidence increased from 3.3 per 100,000 in 2001 to 3.8 per 100,000 in 2004 [4]. Since MDS only became reportable in 2001, the increase in incidence from 2001 to 2004 may reflect the acclimation of registrars to the new reporting guidelines. However, only 4% of MDS cases in NAACCR were reported to registries by physicians' offices [4]. This is a surprisingly low proportion considering that MDS is often diagnosed and managed in outpatient clinics. Furthermore, using a novel claims-based algorithm to

* Corresponding author at: University of Florida, 1600 SW Archer Road, Gainesville, FL 32610-0278, United States. Tel.: +1 352 273 7493; fax: +1 352 273 5006.
E-mail address: c@ufl.edu (C.R. Cogle).

0145-2126/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.leukres.2013.10.023

query the SEER-Medicare database, we and others have estimated that the incidence of MDS in individuals aged 65 years or older is nearly four-fold higher than the incidence estimated by population-based cancer registries [5–7]. Whereas the primary strength of a claims-based approach is its large sample size, this strategy is often limited in its ability to examine details of patient demographics and disease characteristics due to lack of clinical information in administrative datasets. Therefore, we sought to evaluate an improved method of case ascertainment that combines the power of a large, population-based sample size and the detail of patient-level clinical information.

In the United States, state cancer registries currently capture information on cancer diagnoses through paper reports sent directly from hospitals, laboratories, and physicians' offices. To comply with reporting regulations, pathology laboratories report cancer diagnoses to cancer registries via electronic pathology (E-path) reports. Although private physicians are required to report MDS cases to the state cancer registries, they may not be aware of the recent classification of MDS as a reportable malignancy and/or may not have the resources for reporting. Therefore, laboratory-generated E-path reports may represent a comprehensive and cost-effective method of identifying individuals diagnosed with MDS by physicians who send bone marrow biopsy specimens to pathology laboratories for review. However, the sheer volume of E-path reports generated and their considerable overlap with paper reports sent directly from hospitals renders case-finding through E-path reports a resource-intensive process that is not always undertaken by state cancer registries. Therefore, we created a novel method to carefully examine E-path reports in order to improve estimation of the true incidence of MDS.

2. Methods

The Florida Cancer Data System (FCDS, Miami, FL) is the population-based cancer registry for the state of Florida. Hospitals and physicians' offices report incident cancer cases through a web-based reporting tool that facilitates discrete data capture. Pathology laboratories are also required by law to report to FCDS all cancer diagnoses that may have occurred in Florida residents, although their method of reporting differs in that they send batch files of electronic pathology (E-path) reports to FCDS. Upon receipt, these pathology reports are linked to an existing incident report if one exists. These reports correspond to a wide range of laboratory tests, including routine diagnostic tests unrelated to cancer. Until recently (2010), the resources required to conduct cancer case-finding using E-path reports have outweighed the potential benefit of identifying otherwise uncaptured cancer cases. However, as described above, MDS may represent a suitable malignancy for case capture using E-path reports given that it is often diagnosed in the outpatient setting and necessitates review of tissue pathology (i.e., peripheral blood and bone marrow).

To identify potential MDS cases, the following search of E-path reports was conducted. Florida's statewide cancer registry queried all E-path reports received in 2006 using a unique keyword search strategy designed to identify pathology reports that met inclusion and exclusion diagnostic terminology to describe MDS (listed in Table 1). The year 2006 was chosen to permit acclimation of reporting practices introduced in 2001 and completeness of data [4]. Since accurate diagnosis of MDS requires a bone marrow examination [8], the E-path reports were further restricted to bone marrow biopsy reports by querying on the term "marrow". The number of search terms (Table 1) identified in each E-path report was determined, and the E-path report was then categorized into one of four keyword hit categories (≤1, 2, 3–5, or >5). For example, if an E-path report contained two keyword terms from Table 1, then that E-path report would be included in the "2" hit category.

E-path reports were cross-referenced with the FCDS database, using an established probabilistic linkage algorithm. For those cases that linked to existing records in FCDS, additional data elements were retrieved from the FCDS database, including information on age, sex, race, MDS subtype, year of diagnosis, history of previous cancers, and health insurance carrier. MDS subtypes were defined based on eight ICD-O-3 codes: 9980 (Refractory Anemia), 9982 (Refractory Anemia with Ringed Sideroblasts (RARS)), 9983 (Refractory Anemia with Excess Blasts (RAEB)), 9984 (Refractory Anemia with Excess Blasts in Transformation (RAEB-t)), 9985 (Refractory Cytopenias with Multilineage Dysplasia (RCMD)), 9986 (MDS associated with chromosome 5q deletion (del 5q)), 9987 (therapy related MDS), and 9989 (MDS not otherwise specified (NOS)).

To estimate the proportion of E-path reports identified through the keyword search that did not link to existing MDS records in FCDS and corresponded to true uncaptured MDS cases, a manual record review was conducted (CRC). E-path reports were randomly selected from each of the four keyword hit categories (i.e., ≤1, 2, 3–5, or >5) and then further stratified by whether or not the reports linked to existing

Table 1
Characteristics of captured MDS cases identified by FCDS versus uncaptured true MDS cases identified by keyword search term strategy of E-path reports. Case analysis was restricted to the year, 2006.

Characteristic	MDS cases registered in FCDS in 2006 (n = 1061)	Uncaptured MDS cases identified through E-path report review that were registered in FCDS for cancers other than MDS (n = 31)				Uncaptured MDS cases identified through E-path report review that were not registered in FCDS for any type of cancer (n = 40)			
				OR (95% CI)[c]				OR (95% CI)[c]	
	n (%)	n (%)	p[a]	Crude	Age-adjusted	n (%)	p[a]	Crude	Age-adjusted
Age									
<65	175 (16.5)	14 (45.2)		1.00 (reference)		11 (30.6)		1.00 (reference)	
≥65	886 (83.5)	17 (54.8)	<.0001	0.24 (0.12–0.50)	Not applicable	25 (69.4)	0.04	0.45 (0.22–0.93)	Not applicable
Sex									
Male	593 (55.9)	16 (51.6)		1.00 (reference)	1.00 (reference)	23 (57.5)		1.00 (reference)	1.00 (reference)
Female	467 (44.1)	15 (48.4)	0.71	1.19 (0.58–2.43)	1.19 (0.58–2.45)	17 (42.5)	0.87	0.94 (0.50–1.78)	0.80 (0.40–1.58)
Race									
White	980 (92.4)	28 (90.3)		1.00 (reference)	1.00 (reference)			Information not available	
Other	81 (7.6)	3 (9.7)	0.73	1.30 (0.39–4.36)	1.06 (0.31–3.63)				
Health insurance									
Medicare	759 (79.0)	9 (34.6)		1.00 (reference)	1.00 (reference)				
Private	129 (13.4)	10 (38.5)	<.0001	6.54 (2.61–16.40)	5.53 (1.75–17.45)				
MDS subtype (ICD-O-3)									
MDS, NOS[b]	746 (70.3)	6 (19.4)		1.00 (reference)	1.00 (reference)	7 (17.5)		1.00 (reference)	1.00 (reference)
MDS subtype specified	315 (29.7)	25 (80.7)	<.0001	9.87 (4.01–24.29)	8.72 (3.52–21.61)	33 (82.5)	<.0001	11.16 (4.89–25.50)	14.03 (5.39–36.50)
RAEB/RAEB-t	91 (28.9)	9 (36.0)	0.50	1.00 (reference)	1.00 (reference)	8 (24.2)	0.69	1.00 (reference)	1.00 (reference)
Other subtypes	224 (71.1)	16 (64.0)		0.72 (0.31–1.69)	0.81 (0.34–1.93)	25 (75.8)		1.27 (0.55–2.92)	1.20 (0.52–2.80)

[a] p-value from the Fisher exact test.
[b] NOS = not otherwise specified.
[c] Odds ratio (OR) and 95% confidence interval (CI); MDS cases registered in FCDS in 2006 served as the comparison group for both the uncaptured cases registered in FCDS for different cancers and uncaptured cases not registered in FCDS for any cancer.