YOUR DREAM JOB IS JUST A CLICK AWAY!
Access the Career Center and prepare to find the perfect job!

- Manage your job search
- Access hundreds of job postings
- Post an anonymous resume
- Advanced Job Alert system
- Never let another job opportunity pass you by!

Job search today at National Cancer Registrars Association!
Access NCRA’s Job Bank at http://www.ncra-usa.org

POWERED BY JOBTARGET
Contents

Original Articles

Application of Automated Pathology Reporting Concepts to Radiology Reports .............................................. 95
Suzanne March, MBA, CMC; George Cernile, BSc, CKE, PMP; Kim West, BS; Diane Borhani, MBA, CMC; April Fritz, BA, RHIT, CTR; Peter Brueckner, MD, FRCP, MBA

Evaluating Factors Associated with Unknown SEER Summary Stage 2000 Derived from Collaborative Stage at Central Registry Level .................................................. 101
Mei-Chin Hsieh, MSPH, CTR; Qingzhao Yu, PhD; Xiao-Cheng Wu, MD, MPH, CTR; Brad Wohler, MS; Ying Fan, MS; Baozhen Qiao, PhD; Ahmedin Jemal, DVM, PhD; Umed A. Ajani, MBBS, MPH

Coding of Specific Subgroups of Myelodysplastic Syndromes in a Population-Based Cancer Registry: Prospects for Improvement .......................... 107
Anthony P. Polednak, PhD; Cathryn Phillips, CTR

Leukemia as a Cause of Death among Patients with Myelodysplastic Syndromes (MDS) in a Population-Based Cancer Registry: Improving Estimates of MDS-Related Mortality in the Population .................................................. 115
Anthony P. Polednak, PhD; Cathryn Phillips, CTR

Evaluation of Primary/Preferred Language Data Collection .......................................................... 121
Linh M. Duong, MPH; Simple D. Singh, MD, MPH; Natasha Buchanan, PhD; Joan L. Phillips, CTR; Ken Gerlach, MPH, CTR

Comparison of the NCRA and NAACCR Strategic Management Plans ............................................. 133
Herman R. Menck, BS, MBA, CPhil, FACE

Features and Other Journal Departments

Raising the Bar: Four Secrets to a Meaningful Life ........................................................................ 136
Michele Webb, CTR

CoC Outstanding Achievement Award: Recognizing Exceptional Programs ........................ 138
Deborah C. Roberson, MSM, CTR; Denise Harrison, BS, CTR

Summer 2012 Continuing Education Quiz Answers .............................................................. 139
Deborah C. Roberson, MSM, CTR; Denise Harrison, BS, CTR

Fall 2012 Continuing Education Quiz ......................................................................................... 140
Deborah C. Roberson, MSM, CTR; Denise Harrison, BS, CTR

Special Birth Defects Call for Papers ....................................................................................... 142

Call for Papers ..................................................................................................................... 143

Information for Authors .................................................................................................. 144
Editors
Vicki Nelson, MPH, RHIT, CTR, Editor-in-Chief
Vonetta Williams, MPH, CTR, Associate Editor
Reda J. Wilson, MPH, RHIT, CTR, Editor Emeritus
Ginger Yarger, Copy Editor
Janice Ford, Production Editor

Contributing Editors
Faith G. Davis, PhD—Epidemiology
April Fritz, BA, RHIT, CTR—Cancer Registry
Denise Harrison, BS, CTR—CE Credits
Deborah C. Roberson, MSM, CTR—CE Credits
Michele A. Webb, CTR

NCRA 2012–2013 Board of Directors
President: Sarah Burton, CTR
President Elect/Secretary: Shirley Jordan Seay, PhD, OCN, CTR
Immediate Past President: Melanie Rogan, CTR
Treasurer Senior: Terri Richardson, RHIA, CTR
Treasurer Junior: Dianna L. Wilson CTR, RHIA
Educational Director: Ellen Kolender, RHIA, CTR
Professional Development Director: Deidre M. Watson, CTR
Public Relations Director & Liaison to JRM: Theresa M. Vallerand, CTR, BGS
Recruitment & Retention Director: Joyce Ritter, RHIA, CTR
ATP Director—East: Nadine Jenkins, CTR
ATP Director—Midwest: Laurie Hebert, RHIA, CCS, CCS-P, CTR
ATP Director—West: Linda Fine, CTR

Production and Printing
The Goetz Printing Company

Indexing
The Journal of Registry Management is indexed in the National Library of Medicine's MEDLINE database. Citations from the articles indexed, the indexing terms (key words), and the English abstract printed in JRM are included and searchable using PubMed.

For your convenience, the Journal of Registry Management is indexed in the 4th issue of each year and on the Web (under “Resources” at http://www.ncra-usa.org/jrm). The 4th issue indexes all articles for that particular year. The Web index is a cumulative index of all JRM articles ever published.
Abstract: Objective: This study was designed to extend the concept of automated pathology reporting to radiology reports to find central nervous system (CNS) neoplasms that may currently go undetected. Methods: Existing E-Path software was modified to account for the structure and language of radiology reports. Logic was added to allow registries to configure whether they want only new reports or if they also want history, metastatic, and/or previously known reports. Five hospital registries and 3 central registries participated. Three quality-control (QC) studies were conducted with fine-tuning taking place between the studies. The first QC study included random samples of 1,500 reports from 3 data sources. The second and third QC studies each included 1 random sample from 2 different data sources. Results: The software was able to extract reportable CNS neoplasms with a high degree of specificity and sensitivity at 99% and 100% respectively, using the original set of coding rules. This rule set was favored by our hospital registries. Participating population-based registries preferred to receive only positive-new cases. The specificity and sensitivity for this category was 96% and 94% respectively. One hospital registry compared the cases found by the software to their registry database and found 13 additional CNS neoplasm cases in a 10-month period which represented an increase of 18%. Conclusion: Automated radiology reporting is a promising method of mining a previously untapped data source to find cases of CNS neoplasms that may be missed by conventional techniques.

Key words: CNS neoplasms, brain tumors, automated cancer reporting, case ascertainment, radiology reports

Introduction

The key to improving the cancer surveillance system in the United States is the timely collection and integration of patient information from multiple sources within a medical care environment and across multiple institutions. Currently, about 95% of cancers are diagnosed through pathology reports from either hospital-based or independent laboratories. A number of initiatives underway since 2000 have applied natural language processing (NLP) techniques to automate the process of identifying and submitting reportable pathology cases to central and hospital registries. Today, automated pathology reporting is the primary tool for extraction of case information from pathology reports for this purpose. Currently, however, there is no similar software broadly available to extract potentially reportable cases from other data sources. The goal of our study was to extend automated pathology reporting to magnetic resonance imaging (MRI) of the brain and computed tomography (CT) scans of the head and neck to provide a cost-effective method of finding reports of CNS neoplasms that may currently go undetected.

We selected CNS neoplasms because the Benign Brain Tumor Cancer Registries Amendment Act (Public Law 107-260) passed October 29, 2002 imposed an unfunded mandate on all US population-based cancer registries to register these tumors. Further, according to the Central Brain Tumor Registry of the United States (CBTRUS), only about 67% of all brain tumors had a histologically confirmed diagnosis, with substantial regional variation (range for participating cancer registries: 55%-97%). Of the non-malignant brain tumors, 56% were histologically confirmed, while 41% were confirmed radiologically. The figures suggest that the historical heavy emphasis on pathology reports for casefinding could result in a substantial number of CNS neoplasms being missed.

Methods

For the study we sought the participation of 3 well-respected population-based (central) registries that would provide geographical diversity across the United States and whose data providers offered a rich source of MRIs of the brain and CT scans of the head and neck. Once a central registry was committed to participating in the project, that registry was commissioned to recruit a data provider within their catchment area. One central registry recruited 2 data providers. In addition, a not-for-profit health care provider participated but there was no participation from the corresponding central registry.

Gold Standard Reference Set

A.Fritz and Associates (AFA), highly respected experts in cancer coding, built a reference set of 10,000 de-identified radiology reports as the “gold standard” against which the software accuracy was measured. Two measurements of accuracy were used. Sensitivity quantifies the ability of the software to detect reportable CNS neoplasms, and specificity quantifies the ability of the software to detect reportable CNS neoplasms, and...
Table 1. Categorization of Selectable Radiology Reports

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Negative</td>
<td>No mention of cancer or CNS neoplasm</td>
<td>No significant intracranial abnormality</td>
</tr>
<tr>
<td>1</td>
<td>History of cancer</td>
<td>Mention of history of tumor or cancer</td>
<td>History of prostate cancer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No other neoplasm</td>
</tr>
<tr>
<td>2</td>
<td>Metastatic disease</td>
<td>Reference to the presence of metastases</td>
<td>History of renal cancer with lesions on frontal lobe</td>
</tr>
<tr>
<td>3</td>
<td>Positive, but previously known</td>
<td>A report of any primary CNS neoplasm with wording to indicate that it was known to be present when the examination was ordered</td>
<td>Resected medulloblastoma</td>
</tr>
<tr>
<td>4</td>
<td>Positive, new case</td>
<td>A report of any primary CNS neoplasm or other primary malignant tumor that did not contain reference to history of tumor or metastases</td>
<td>Pituitary gland microadenoma</td>
</tr>
</tbody>
</table>

Table 2. Sensitivity and Specificity Results for 3 Rounds of QC

<table>
<thead>
<tr>
<th></th>
<th>First QC study results by site</th>
<th>Second QC study</th>
<th>Third QC study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Central 1</td>
<td>Central 2</td>
<td>Hospital 1</td>
</tr>
<tr>
<td>True positives</td>
<td>140</td>
<td>247</td>
<td>112</td>
</tr>
<tr>
<td>True negatives</td>
<td>1304</td>
<td>1109</td>
<td>1256</td>
</tr>
<tr>
<td>False positives</td>
<td>44</td>
<td>132</td>
<td>117</td>
</tr>
<tr>
<td>False negatives</td>
<td>12</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Sample total</td>
<td>1500</td>
<td>1500</td>
<td>1500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>97%</td>
<td>92%</td>
<td>89%</td>
<td>95%</td>
<td>91%</td>
<td>88%</td>
<td>99%</td>
<td>96%</td>
<td>99%</td>
<td>100%</td>
</tr>
</tbody>
</table>
measures its ability to reject reports that do not describe CNS neoplasms. Ideally, the sensitivity of the system should approach 100% to be sure that no reportable cases are missed. Alternatively, to the extent that the specificity of the system is below 100% reports will not be rejected by the system and will be false positives. Since we expect that all selected reports will be reviewed by a certified tumor registrar (CTR), the false-positive rate represents additional work for that person.

In preparation for developing the reference set, AFA and the development team collaborated to establish the set of coding rules that would be used for the manual review and that the developers would incorporate into the software.

**Software Design**

The system used in the study was based on E-Path software components from Artificial Intelligence In Medicine, Inc. The processing of diagnostic imaging reports required the development of a lexicon and knowledge base specific to that discipline and extensive heuristic testing followed by adjustments to enable the system to perform usefully. Radiology subject matter experts assisted by elucidating the meaning of terms that are used in reporting diagnostic imaging studies. Modifications were also made to the utilities used for reviewing, annotating and scoring reports.

Figure 1 is a schematic depiction of the automated system for reporting CNS neoplasms. The paragraph numbers below correspond to the labels in the diagram.

1. Diagnostic imaging systems such as MRI and CT scanners were used to image the head and neck to discover lesions that may be neoplasms of the central nervous system. Radiologists interpreted these images and prepared narrative reports describing the findings. The imaging reports were stored in a reporting system or database.

2. The modified ePath software analyzed the text content of the imaging reports to determine if a reportable finding is in the text. This required standardization of the format of the imaging report and an interface to the source data. A standard format allowed the report data to be passed from the integration engine to the coding and filtering component.

3. The coding engine included 2 components: a natural language processing (NLP) engine and a lexicon describing the concepts and vocabulary of a particular domain of interest, in this case radiology. The NLP engine used information in the lexicon to appropriately parse and codify concepts from their natural language expression. Once the narrative text of an imaging report had been converted to machine-readable form, the integration engine could easily discern which imaging reports should be forwarded to a cancer registry. These were transmitted over a secure network to a receiving system.

4. Development of a CNS neoplasm lexicon was the objective of this project. A lexicon is a collection of data that describes the language, concepts, and vocabulary of a particular domain of interest. It also contains coding system(s) to convert concepts to machine-readable values.

5. Imaging reports were received at the registry using an integration/messaging engine, much like the engine used to extract the reports from their original source. However, before passing the data onto the cancer registry system, the software allowed a human at the registry to make the final determination as to whether or not it should be included in the registry database.

6. The cancer registry software/database is the ultimate destination of the imaging reports. Here, the reports would be stored in a “suspense” file to be further incorporated into cancer case abstracts. A standard input format utilizing HL7 and/or XML technologies would be used to achieve compatibility with the different varieties of cancer registry software.

**Software Development**

During the iterative testing period, the developers compared the results from the software to the reference set built by AFA. Using groups of approximately 2,000 reports at one time, they reviewed every discrepancy between the manual and computer coding. Facilitated meetings were held with the CTRs from AFA to discuss and resolve the discrepancies. In some cases, the CTRs decided to change their coding to better comply with a coding rule, but most of the time the software was modified to improve its accuracy. Once the accuracy was high for a set of 2,000, a new set was selected and a new baseline of sensitivity and specificity was established. After 3 datasets and a total of 6,000 test records processed with 17 test iterations executed, the sensitivity and specificity were 99% and 96% respectively. This was deemed an acceptable level to begin testing at the participating registry sites.

Registry implementations took place in September 2010. Once the system was installed, the registries began receiving files containing the reportable radiology reports selected by the system in accordance with the established selection rules. Registries were asked to review the reports to determine if they agreed with the software and inform the developers of disagreements. They were also asked to record the amount of time they spent reviewing the reports and to compare the positive cases found by the software to their registry database to determine if the tool picked up cases not found by existing registry procedures.

After the registries had used the software for a few weeks, it became apparent that there were significant differences in how registries would eventually use the tool in a production environment, and that 1 universal set of coding rules would not be sufficient. In some cases, the registry wanted only cases that represent new cases of CNS neoplasms. Other registries wanted to receive reports of all neoplasms, whether or not they had been previously diagnosed and regardless of the tissue of origin. To accommodate these differing requirements, an automated classification system was introduced to assign reportable reports into 4 categories as described in Table 1. The negative (not reportable) reports constitute a fifth category.

New logic was added to the software to allow registries to select which categories of reports they wished to receive. To maintain comparability during the study period,
however, all participating registries continued to receive and process all categories. The remaining 4,000 test reports in the reference set were used to perform iterative testing of the new functionality. After the registries had used the revised software for 2 months, we began the quality control phase of the project.

QC Studies

Three sites participated in the first QC study. Each site provided a random set of 1,500 reports from MRIs of the brain and CT scans of the head and neck. A CTR at the site reviewed the reports and categorized them, assigning codes 0-4 in accordance with the classification described in Table 1. Then, after the same reports were processed by the software, a Microsoft Excel spreadsheet was created to enable comparison of the 2 sets of results. Using the CTR coding as the benchmark for correctness, the reports were defined as true positives, true negatives, false positives and false negatives and the sensitivity and specificity were calculated. For this calculation, categories 1-4 were considered positive. Separately, it was determined how successful the software was at distinguishing between the 4 categories of positive.

The developers reviewed every report where the software’s result differed from that of the CTR to determine the causes of the discrepancy and also attempted to identify recurring patterns. Facilitated sessions were held to discuss the findings and agree on the solutions, which included a combination of changing the manually applied code to better conform to the selection rules or modifying the software by changing the lexicon, casefinding logic, or which sections of the report were processed by the system. Software modifications were introduced in 6 iterative stages to make it easier to determine if accuracy improved. To determine if the improvements in sensitivity and specificity achieved by modifying the software were general (rather than specific to the data set used to make the adjustments), a second QC study was conducted using new data. More modifications were made to the software and then a third QC study was conducted, again with a new dataset.

Results

Sensitivity and Specificity Results

The number of true and false positives and negatives and the calculated specificity and sensitivity results for the 3 QC studies are documented in Table 2.

Table 3 shows the software’s ability to categorize selectable reports into the 4 categories of positive. The column label Manual is the CTR’s categorization. Auto is the software’s categorization and Match is the number of times the software classification agreed with the manual review.

At each site, the total number of cases categorized using the software exceeded the total number categorized manually. The difference is due to overselection (false positives) of reports by the software. Significant development work was expended to improve the initial match results. After 6 rounds of iterative modification of the software, the average matching results for history of cancer, metastatic disease, positive but previously known and positive new were 77%, 98%, 65%, and 52%, respectively.

To further quantify the software’s ability to categorize reports, during the second and third QC studies, each of the 4 positive categories was isolated and the sensitivity and specificity for each category was computed. Results are documented in Table 4.

Case-Ascertainment Results

Hospital 1 provided the most extensive comparison between cases found by the software and cases found by their traditional methods. Table 5 documents the performance over a 10-month period from November 2010 through August 2011. The figure in the rightmost column of the total line is the average new cases found per month.

Discussion

Conclusions regarding the tool’s ability to select reportable CNS neoplasms with a minimum of false-positives depend on the way in which the tool will be used in a production environment. It is clear from the study that it is realistic, using the original set of coding rules, to extract reportable CNS cases with a high degree of specificity and sensitivity at 99% and close to 100%, respectively. The coding rules include selecting histories, metastases and positive previously known as well as new cases, a configuration that appealed to our hospital registry participants.

The central registries in the study, on the other hand, were interested only in new cases. By the third QC study, significant improvement was achieved for the positive-new category with specificity at 96% and sensitivity at 94%. For the other categories the modifications continued to produce mixed results. Sensitivity improved for positive, but previously known and metastatic disease specificity decreased. In the third QC study, history of cancer resulted in lower scores for both. Further work is required to accomplish...
consistently high sensitivity and specificity for individual categories in isolation. However, we believe the most typical configurations will be either selection of positive new only or selection of all categories. The high degree of accuracy for these configurations suggests that the software can be a useful tool for registries to find CNS neoplasms.

With regard to the tool’s ability to discover CNS cases that might go undetected by current methods, the results were very promising. The 10-month study conducted by Hospital 1 found 13 new CNS cases not found by their traditional methods of using E-Path and monthly hospital discharge lists. One case that was found traditionally was not found by the software, for a net increase of 12 additional cases. Overall, this represented an 18% improvement in CNS case ascertainment. While it is important to be cautious about generalizing the Hospital 1 experience, preliminary data is becoming available from other study participants.

For example, after reviewing 91 reports selected by the software during the test period, Registry 1, a central registry, identified 20 reports (22%) they expected to see in their registry database but did not. The timing of the review was approximately 5 months after the end of 2011 and reports were still being received. A second review will be necessary to determine if they would have ultimately been found.

| Table 4. Second and Third QC Results Isolated by Classification (Hospital 2 and Central 3, Respectively) |
|-------------------------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Positive new | Positive known | History of cancer | Metastatic disease |
| QC2 | QC3 | QC2 | QC3 | QC2 | QC3 | QC2 | QC3 |
| True Positive | 22 | 291 | 8 | 273 | 46 | 184 | 16 | 101 |
| True Negative | 1382 | 1452 | 1390 | 1353 | 1352 | 1460 | 1400 | 1678 |
| False Positive | 4 | 59 | 10 | 67 | 21 | 120 | 4 | 37 |
| False Negative | 14 | 20 | 14 | 129 | 3 | 58 | 2 | 6 |
| Sample Total | 1422 | 1822 | 1422 | 1822 | 1422 | 1822 | 1422 | 1822 |
| Specificity | 100% | 96% | 99% | 95% | 98% | 92% | 100% | 98% |
| Sensitivity | 61% | 94% | 36% | 68% | 94% | 76% | 89% | 94% |

| Table 5. Hospital 1 Comparison of Cases Found By Method |
|------------------------------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Month | Found by software | Found traditionally | Missed by software | Additional cases found by software | Net new cases | Percent increase in cases found |
|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| November 2010 | 7 | 5 | 0 | 2 | 2 | 40 |
| December 2012 | 7 | 6 | 0 | 1 | 1 | 17 |
| January 2012 | 6 | 4 | 0 | 2 | 2 | 50 |
| February 2012 | 9 | 7 | 0 | 2 | 2 | 29 |
| March 2012 | 9 | 8 | 0 | 1 | 1 | 13 |
| April 2012 | 8 | 6 | 0 | 2 | 2 | 33 |
| May 2012 | 9 | 9 | 0 | 0 | 0 | 0 |
| June 2012 | 17 | 14 | 0 | 3 | 3 | 21 |
| July 2012 | 3 | 4 | 1 | 0 | -1 | -25 |
| August 2012 | 5 | 5 | 0 | 0 | 0 | 0 |
| Totals | 80 | 68 | 1 | 13 | 12 | 18 |

Conclusion and Future Possibilities

Automated reporting of CNS neoplasms from radiology reports is a reliable method of mining a previously untapped data source. As was the case for the automated pathology reporting software, increased use of the tool will provide the experience necessary to further refine the lexicon and continuously improve the software’s performance.

Further research is required to determine if the use of radiology reports for casefinding can be extended to other cancers. During the study, we explored extending the software to pancreas and biliary tract radiology reports because, like CNS, we suspected that advanced cases could be missed if no follow up pathology report was ordered. However, because the radiology reports included many other organs, from colon, bowels, bladder, spleen, heart, lungs, liver, etc, the linkage between the diagnosis and the affected organ was difficult to determine programmatically.

The Radiological Society of North America (RSNA) set a goal in 2005 to collaborate with radiology professional organizations to develop a comprehensive set of anatomic and pathologic terms called RadLex.7 The first version of the RadLex Playbook was recently released on November 1, 2011.8 Since the 2007 American College of Radiology Inter...
Conference, work has been underway to use structured radiology reports to improve communication of radiology procedures employing both consistently ordered sections and standardized language. For a variety of reasons, adoption of standardized radiology reports has been slow. But as the obstacles to standardization are overcome, we believe automated radiology reporting can be successfully applied to cancers beyond CNS.

Acknowledgements

This project was funded in part with federal funds from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services under Contract No. HHSN261200900040C. We would especially like to acknowledge our hospital and central registries for their willingness to support our project with domain expertise, technical support, data review and analysis. We know you had many competing demands on your time and your commitment to our project was greatly appreciated.

References

Evaluating Factors Associated with Unknown SEER Summary Stage 2000 Derived from Collaborative Stage at Central Registry Level

Mei-Chin Hsieh, MSPH, CTR; Qingzhao Yu, PhD; Xiao-Cheng Wu, MD, MPH, CTR; Brad Wohler, MS; Ying Fan, MS; Baozhen Qiao, PhD; Ahmedin Jemal, DVM, PhD; Umed A. Ajani, MBBS, MPH

Abstract: Background: Cancer stage is critical for treatment planning and assessing disease prognosis. The percentage of unknown staged cancer cases varies considerably across state cancer registries; factors contributing to the variations in unknown stage have not been reported in the literature before. The purpose of this study was to examine whether these variations were influenced by demographic and/or clinical factors as well as the type of reporting facility. Methods: Invasive colorectal, lung, female breast, and prostate cancers diagnosed between 2004 and 2007 were obtained from the North American Association of Central Cancer Registries (NAACCR); 47 population-based cancer registries in the United States were included. The unknown stage was based on Summary Stage 2000 codes derived from Collaborative Stage Version 1 (CSV1). Relative importance analysis was used to identify variables that were essential in predicting unknown stage. Using state central registries as analytical units, multiple linear regression was used to evaluate factors associated with the percentage of unknown stage by cancer site; potential outlier registries with a high percentage of unknown stage cases were identified using boxplots and standardized residuals. Results: Overall, lung cancer had the highest percentage of unknown stage (8.3%) and prostate cancer had the largest variation of unknown stage among registries (0.6%-18.1%). The percentages of neoplasms not otherwise specified (NOS) histology, non-microscopic confirmation, and non-hospital reporting source were positively associated (p<0.05) with percentage of unknown stage for all studied cancer sites before adjustment. Variables that retained a positive association with unknown stage including all demographic and clinical variables, year of diagnosis, and type of reporting source were black race, metropolitan area <1 million population, histologies of neoplasms NOS or epithelial neoplasms NOS, diagnosis year 2005, and non-hospital reporting source for colorectal cancer; metropolitan area <1 million population, neoplasms NOS histology, and non-hospital reporting source for female breast; and diagnosis year 2005 and non-hospital reporting source for prostate. After adjustment, none of the predictors were significant for lung cancer. We observed 1 potential outlier registry each for colorectal, lung and female breast cancers. Conclusions: Factors associated with unknown stage differ by cancer site; however, the type of reporting source is an important predictor of unknown stage for all cancers except lung after adjustment. Central registries with high percentage of unknown stage should be made aware of their data quality issue(s). As a result, these registries can investigate those factors and provide training to registrars to improve their cancer data quality.

Key words: cancer, summary stage, collaborative stage, cancer registry

Introduction

Cancer staging is essential for assessing the effectiveness of early detection and intervention as well as for treatment planning and predicting outcomes. Formed in 1998, the Collaborative Staging (CS) Task Force, now known as the CS Governance Committee, developed the CS Data Collection System for use with cases diagnosed in 2004 and after to address the issue of discrepancies in staging guidelines among the 3 major cancer staging systems: American Joint Committee on Cancer (AJCC) staging, Surveillance, Epidemiology, and End Result (SEER) Summary Stage, and SEER Extent of Disease (EOD). Using the CS Data Collection System could also reduce the percentages of unknown stage. A study using data in the Centers for Disease Control and Prevention’s (CDC’s) National Program of Cancer Registries (NPCR) Cancer Surveillance System (CSS) observed a decreased percentage of unknown SEER Summary Stage 2000 (SS2000) from 2001-2003 period to 2004-2005 period. The percentage of unknown stage is a strong quality indicator of stage data as well as the quality of abstraction and the availability of source data. The Data Assessment

Original Article

Journal of Registry Management 2012 Volume 39 Number 3

101
Workgroup of the North American Association of Central Cancer Registries (NAACCR) Data Use and Research Committee (DURC) found that the percentage of unknown stage cases varies substantially by central cancer registry and cancer site.5 The common approach in research for handling unknown stage cancer cases is to exclude them from the statistical analysis; however, this could create biased results.6 To help central registries implement appropriate strategies to improve stage data quality, it is important to identify and quantify the factors associated with unknown stage. Unfortunately, such studies are lacking in the literature. Previous studies on cancer stage primarily focused on finding factors associated with advanced stage only.7-11 The purposes of this study were to identify factors associated with unknown stage cases at the central registry level for colorectal, lung, female breast, and prostate cancer as well as to identify those registries with unusually high percentages of unstaged cases.

Data and Methods

Data Source

Data from 47 US population-based cancer registries were obtained from the NAACCR Cancer in North America (CINA) Analytic file 1995-2007 data. Two states declined to be included in the CINA analytic file and 1 state did not submit the Rural/Urban 2003 Continuum code12 thus eliminating them as a data source. We included invasive colorectal (ICD-O-313 Topography: C18.0-C18.9, C19.9, C20.9), lung (C34.1-C34.9), female breast (C50.0-C50.9), and prostate (C61.9) cancer cases diagnosed in the years 2004-2007. Autopsy or death-certificate-only cases, lymphomas originating in the sites of interest, and cases with unknown rural/urban information were excluded from the analysis.

The outcome of interest was the percentage of cases with unknown SS2000 derived from the Collaborative Stage Version 1 (CSv1) at the central registry level. Independent variables were demographic (race, sex, age, and rural/urban residence) and clinical (histology, grade, and diagnostic confirmation) data, as well as the year of diagnosis and the type of reporting source. Residential regions were defined based on the 2003 Rural/Urban Continuum codes12: county metropolitan areas with populations of 1 million or more, county metropolitan areas with populations of less than 1 million, and non-metropolitan areas. Histologies were categorized as: neoplasms, not otherwise specified (NOS) (ICD-O-313 morphology 8000-8005), epithelial neoplasms NOS (8010-8046), adenocarcinoma NOS (8140), and specified histologies. The type of reporting source was grouped into 2 categories: hospital and non-hospital facilities. Beginning with cases diagnosed in 2006, 2 new categories, radiation treatment centers/medical oncology centers (hospital-affiliated or independent) and other hospital outpatient units/surgery centers, were introduced as types of reporting sources in NAACCR records.14-15 The hospital group included hospital inpatient, radiation treatment centers/medical oncology centers, and hospital outpatient units/surgery centers. Non-hospital facilities included physician’s office, nursing home/hospice, and laboratory only.

Statistical Analysis

Relative importance analysis16 was adopted to describe the contribution of each predictor variable in explaining the variance of the outcome. Predictor variables were sequentially added to linear regression models, from which the increase in R square was recorded for each predictor variable. We permuted the order of the predictors to be added in the linear models. The relative importance of the variables...
Table 1. Measures of Central Tendency and Dispersion on Unknown Derived Summary Stage 2000 by Cancer Site From 47 United States Population-Based Cancer Registries, 2004-2007

<table>
<thead>
<tr>
<th></th>
<th>Colorectal</th>
<th>Lung</th>
<th>Female breast</th>
<th>Prostate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean for all cases combined</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(N=552,242)</td>
<td>7.17%</td>
<td>8.25%</td>
<td>3.51%</td>
<td>6.59%</td>
</tr>
<tr>
<td><strong>Based on individual registry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>6.40%</td>
<td>7.30%</td>
<td>3.11%</td>
<td>5.34%</td>
</tr>
<tr>
<td>25-75% Percentile (IQR†)</td>
<td>4.61%-7.27%</td>
<td>4.95%-9.03%</td>
<td>1.56%-3.27%</td>
<td>2.59%-6.44%</td>
</tr>
<tr>
<td>Minimum</td>
<td>2.40%</td>
<td>2.44%</td>
<td>0.98%</td>
<td>0.61%</td>
</tr>
<tr>
<td>Median</td>
<td>5.48%</td>
<td>6.55%</td>
<td>2.29%</td>
<td>3.78%</td>
</tr>
<tr>
<td>Maximum</td>
<td>18.80%</td>
<td>18.68%</td>
<td>13.68%</td>
<td>18.06%</td>
</tr>
<tr>
<td>Upper Whisker‡</td>
<td>11.25%</td>
<td>15.16%</td>
<td>5.83%</td>
<td>12.23%</td>
</tr>
<tr>
<td>Lower Whisker§</td>
<td>2.40%</td>
<td>2.44%</td>
<td>0.98%</td>
<td>0.61%</td>
</tr>
<tr>
<td>Number of registry outside of upper whisker</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

†IQR: Interquartile range.
‡Upper Whisker: extends to largest data point within the boundary of Q3 + 1.5*(Q3-Q1).
§Lower Whisker: extends to smallest data point within the boundary of Q1 - 1.5*(Q3-Q1).

is defined as the average increase in the R square from the permutation of each predictor variable. Predictors, which are treated as continuous variables, with a high relative importance index indicate that the predictor is comparatively important in predicting the percentage of unknown stage. The relative importance index of the predictors varies by cancer site.

Multiple linear regression was used to evaluate the relationship between the predictors and the percentage of cases with unknown stage, and regression models were run separately for each cancer site. The covariate was the percentage of patients falling into each particular subgroup of a predictor at the central registry level. Therefore a predictor with k categories forms k-1 covariates. To identify registries with unusually high percentages of unknown stage, we used the upper whisker of the boxplots for unadjusted percentages and the standardized residuals generated from the multiple linear regression model that adjusted for important factors. The upper whisker of the boxplots is defined as the largest data point of unknown stage percentage within the boundary of following formula: third quartile + 1.5 X (third quartile – first quartile). Any registries with standardized residuals of unknown stage percentage higher than the upper bound of the 95% confidence interval (CI) were considered to have unusually high percentage of unstaged cases. The statistical significant level was set at 0.05. All analyses were carried out using SAS version 9.2 (SAS Institute, Cary, NC).

Results

Lung cancer had the highest percentage of unknown stage (8.3%), followed by colorectal cancer (7.2%) and prostate cancer (6.6%); female breast cancer had the lowest percentage of unstaged cases (3.5%) for all cases combined (Table 1). At the registry level, percentage ranges of unknown stage varied substantially for all studied cancer sites, particularly for prostate cancer varying from 0.6%-18.1% (Table 1). The 3 most important predictors based on relative importance analysis† for colorectal cancer were histology, type of reporting source, and tumor grade; for lung cancer, type of reporting source, diagnostic confirmation, and age; for female breast cancer, type of reporting source, histology, and diagnostic confirmation; and for prostate cancer, tumor grade, type of reporting source, and urban/rural area (Figure 1).

Colorectal Cancer

Registries with higher percentages of patients residing in non-metropolitan areas or with higher percentages of moderately differentiated tumor grade had lower percentages of unknown stage (Table 2), whereas registries with higher percentages of neoplasms NOS histology, unknown diagnostic confirmation, or non-hospital cases had higher percentages of unknown stage in the univariate analysis. After adjustment for all predictors, the percentages of unknown stage differed significantly by race, rural/urban residence, histology, diagnostic confirmation, diagnosis year and type of reporting source. The difference was particularly strong for histology. Compared with specific histology group, a 1% increase in neoplasms NOS or epithelial neoplasms NOS histology was related to an average of 2.7% or 3.7% increase respectively in unstaged rate after controlling for other predictors (Table 2).

Lung Cancer

In univariate analysis, older age groups were related to lower percentages of unknown stage than those aged less than 50 years old (Table 2). Registries with higher percentages of neoplasms NOS, unknown microscopic confirmation or non-hospital reporting source were associated with higher percentages of unknown stage. However, in multivariable analysis no single predictor was significantly associated with unknown stage.

<table>
<thead>
<tr>
<th>Variable§</th>
<th>Colorectal Unadjusted</th>
<th>Adjusted</th>
<th>Lung Unadjusted</th>
<th>Adjusted</th>
<th>Female breast Unadjusted</th>
<th>Adjusted</th>
<th>Prostate Unadjusted</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-0.348</td>
<td>-0.216</td>
<td>-0.348</td>
<td>-0.123</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.038</td>
<td>0.131*</td>
<td>0.027</td>
<td>-0.058</td>
<td>0.037</td>
<td>0.039</td>
<td>0.070</td>
<td>0.072</td>
</tr>
<tr>
<td>Other races</td>
<td>-0.028</td>
<td>0.032</td>
<td>-0.052</td>
<td>-0.114</td>
<td>-0.009</td>
<td>0.031</td>
<td>0.030</td>
<td>-0.042</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤49</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>50-64</td>
<td>-0.565</td>
<td>-1.866</td>
<td>-3.214*</td>
<td>-2.107</td>
<td>-0.390</td>
<td>-0.608</td>
<td>-2.575</td>
<td>3.330</td>
</tr>
<tr>
<td>65-74</td>
<td>-0.321</td>
<td>-0.910</td>
<td>-1.901*</td>
<td>-1.250</td>
<td>0.001</td>
<td>-0.063</td>
<td>-2.719</td>
<td>1.744</td>
</tr>
<tr>
<td>75+</td>
<td>-0.320</td>
<td>-1.027</td>
<td>-2.347*</td>
<td>-1.895</td>
<td>-0.360</td>
<td>-0.292</td>
<td>-2.367</td>
<td>2.845</td>
</tr>
<tr>
<td><strong>Rural/urban residence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metro area ≥ 1 million pop</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>Metro area &lt; 1 million pop</td>
<td>0.003</td>
<td>0.108*</td>
<td>-0.038</td>
<td>-0.039</td>
<td>0.002</td>
<td>0.067*</td>
<td>-0.073</td>
<td>-0.017</td>
</tr>
<tr>
<td>Non-metro area</td>
<td>-0.040*</td>
<td>-0.009</td>
<td>-0.031</td>
<td>-0.017</td>
<td>-0.024</td>
<td>-0.020</td>
<td>-0.071*</td>
<td>-0.028</td>
</tr>
<tr>
<td><strong>Histology type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specified histologies</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>Neoplasms, NOS†</td>
<td>2.486*</td>
<td>2.718*</td>
<td>0.719*</td>
<td>0.925</td>
<td>2.629*</td>
<td>3.168*</td>
<td>2.031*</td>
<td>-1.670</td>
</tr>
<tr>
<td>Epithelial neoplasms, NOS†</td>
<td>1.227</td>
<td>3.653*</td>
<td>0.159</td>
<td>0.505</td>
<td>0.364</td>
<td>0.395</td>
<td>-0.475</td>
<td>1.033</td>
</tr>
<tr>
<td>Adenocarcinoma, NOS†</td>
<td>0.197</td>
<td>0.102</td>
<td>0.173</td>
<td>0.684</td>
<td>0.636</td>
<td>-0.245</td>
<td>0.007</td>
<td>-0.235</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well differentiated</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td>-0.507*</td>
<td>0.311</td>
<td>-0.836</td>
<td>-0.095</td>
<td>-0.277</td>
<td>-0.291</td>
<td>0.320</td>
<td>-0.624</td>
</tr>
<tr>
<td>Poorly differentiated</td>
<td>-0.121</td>
<td>-0.129</td>
<td>-0.090</td>
<td>0.475</td>
<td>0.012</td>
<td>-0.061</td>
<td>0.327</td>
<td>-0.680</td>
</tr>
<tr>
<td>Undifferentiated</td>
<td>-0.939</td>
<td>1.290</td>
<td>-0.249</td>
<td>0.228</td>
<td>-0.761</td>
<td>-0.434</td>
<td>0.361</td>
<td>-0.517</td>
</tr>
<tr>
<td>Unknown</td>
<td>0.066</td>
<td>0.317</td>
<td>-0.300</td>
<td>0.152</td>
<td>0.184</td>
<td>-0.247</td>
<td>1.080</td>
<td>0.410</td>
</tr>
<tr>
<td><strong>Diagnostic confirmation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>-1.132</td>
<td>-3.863*</td>
<td>-0.146</td>
<td>-0.397</td>
<td>1.282</td>
<td>-1.197</td>
<td>-2.541</td>
<td>-2.357</td>
</tr>
<tr>
<td>unknown</td>
<td>2.910*</td>
<td>-1.549</td>
<td>1.076*</td>
<td>0.218</td>
<td>5.999*</td>
<td>-0.553</td>
<td>3.114*</td>
<td>2.112</td>
</tr>
<tr>
<td><strong>Diagnosis year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>1.491</td>
<td>1.497*</td>
<td>-0.377</td>
<td>-0.218</td>
<td>0.482</td>
<td>0.102</td>
<td>-0.331</td>
<td>1.541*</td>
</tr>
<tr>
<td>2006</td>
<td>-0.508</td>
<td>-0.088</td>
<td>-0.626</td>
<td>-0.665</td>
<td>0.446</td>
<td>0.675</td>
<td>-0.865</td>
<td>-0.152</td>
</tr>
<tr>
<td>2007</td>
<td>-0.766</td>
<td>-0.0004</td>
<td>-1.538</td>
<td>-0.741</td>
<td>-0.683</td>
<td>-0.484</td>
<td>0.426</td>
<td>0.391</td>
</tr>
<tr>
<td><strong>Type of reporting source</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>Non-hospital</td>
<td>0.422*</td>
<td>0.356*</td>
<td>0.380*</td>
<td>0.238</td>
<td>0.302*</td>
<td>0.223*</td>
<td>0.240*</td>
<td>0.293*</td>
</tr>
</tbody>
</table>

§Predictor of k categories forms k-1 covariates and covariate is the proportion of patients falling in the subgroups of each predictor at each registry.
~Not applicable.
†NOS indicates Not Otherwise Specified.
*Statistical significance at p <0.05.
Female Breast Cancer

The percentages of cases with histologies of neoplasms NOS, unknown diagnostic confirmation or non-hospital reporting source had positive associations with unknown stage in univariate analysis. This is most striking for unknown diagnostic confirmation. Compared with microscopically confirmed cases, a 1% increase in the unknown diagnostic confirmation related to about an average of 6% increase in the percentage of unknown stage (Table 2). After adjusting for other predictors, the percentage of patients residing in a metropolitan area with populations less than 1 million, the percentages coded to neoplasms NOS, and abstracted at non-hospital reporting source were positively associated with unknown stage. Compared to specific histologies, a 1% increase in neoplasms NOS was related to an average of 3.2% increase in the percentage of unknown stage after other covariates were controlled.

Prostate Cancer

Although tumor grade yielded the highest index in the variable relative importance test (Figure 1), no statistically significant differences resulted when the percentages of well-differentiated cases were compared with other grade groups in the model that were not adjusted for other predictors (Table 2). However, the F-test showed that tumor grade was significantly associated with the percentage of unknown stage (p-value=0.014). We further examined this association using unknown grade as baseline and found that moderately to poorly differentiated grades had negative associations with the percentage of unknown stage (coefficients -0.731 and -0.579, respectively). After adjustment, registries with higher percentages of either cases diagnosed in 2005 or cases reported by non-hospital reporting source were more likely to have a higher percentage of unstaged prostate cancer cases.

Identifying Registries with High Percentage of Unknown Stage

Before adjusting for all important predictors, we identified 2 registries as outliers with a high percentage of unknown stage (above upper whisker) for colorectal cancer, 1 for lung, 6 for breast and 6 for prostate cancer (Table 1). After adjustment, we found that only 2 registries had unusually high percentages of unknown stage (outside the upper bound of 95% CI), 1 for both colorectal and female breast cancers and 1 for lung cancer.

Discussion

To our knowledge, this is the first study assessing the factors associated with unknown stage and utilizing a central registry as an analytical unit. We observed that the percentage of unknown stage varied by registry and cancer site; factors related to unknown stage also differed by cancer site. Overall, female breast cancer had the lowest percentage of unknown stage cases among the 4 cancer sites. In the univariate analysis, neoplasms NOS, unknown diagnostic confirmation, and non-hospital reporting source were positively associated with the high percentages of unknown stage for all 4 cancer sites. After adjustment, the type of reporting source remained as an important predictor of unknown stage for all cancers except lung. Although race has been associated with advanced stage for colorectal, female breast, and prostate cancer,

105
for tumor stage except for cancer sites with high percentages of missing stage information. Therefore, it is critical that cancer registries reduce the unknown stage percentage as well as unknown percentages for other important variables to generate more valid results when using registry data.

This study has several strengths. First, it included 47 of the 50 US state central cancer registries; therefore, the results are generalizable to the population as the whole United States. Second, data were collected and coded uniformly based on CS and NAACCR rules, ensuring that coding schemes for variables used in this study were compatible across registries. Third, we used central registry level data as an analytical unit, instead of the individual cancer case level; this could identify issues occurring within individual registries and also identify potential outlier registries. Despite these strengths, there are a few limitations to this study. The registries’ experience of coding SEER Extent of Disease (EOD) was not taken into account in the analysis. Coding rules from SEER EOD for tumor size, tumor extension and lymph node involvement were adapted and modified in the CS system. Registries that collected SEER EOD for cases diagnosed before 2004 may have been more likely to accept the concept of coding CS data items. In addition, the effect of cancer registrar characteristics, such as certification and/or years of experience, is not evaluated. Central registries with a high proportion of non-certified tumor registrars in their state may impact the results of unknown stage. Another limitation is that class of case (NAACCR item 160) is not available in our study. Non-analytical cases (class of case codes 3-9), which include cases not diagnosed and/or treated at reporting facility, provide very limited information on CS-related information; therefore, registries with high percentages of non-analytical cases may increase the unknown stage percentages.

For cancer registries having a higher percentage of unknown stage, further investigation of factors that caused the higher unknown stage percentage is needed. Population-based cancer registry data are a valuable source for cancer research; any effort to reduce the percentage of unknown stage cases will improve the quality of cancer data and produce more reliable cancer research results. A future study analyzing changes in unknown stage rates over time is needed as more data become available.

**Acknowledgements**

The authors would like to thank the members of the NAACCR Data Assessment Workgroup, particularly Patricia A. Andrews, Maria J. Schymura, Hannah K. Weir, Missy Jamison, Bin Huang, and Xiangrong Li, for their general support.

**References**

Coding of Specific Subgroups of Myelodysplastic Syndromes in a Population-Based Cancer Registry: Prospects for Improvement

Anthony P. Polednak, PhD; Cathryn Phillips, CTR

Abstract: Myelodysplastic syndromes (MDS), reportable to US cancer registries for diagnoses since 2001, are a group of myeloid neoplasms heterogeneous in prognosis and treatment, and of growing importance in an aging population. In US registries that have reported incident MDS cases by subgroup, about 50%-67% of cases have been coded as MDS “not otherwise specified” (NOS) in the International Classification of Diseases for Oncology Version 3 (ICD-O-3). For this study, MDS cases diagnosed in 2001-2009 and reported to the population-based Connecticut Tumor Registry (CTR) were analyzed. MDS was coded as NOS for 573 (56.7%) of 1,011 cases, but the proportion varied among reporting facilities hospitals (ie, from 0 to 100%), with several statistical outliers. In pathology reports obtained for 130 CTR patients diagnosed with MDS in 2008-2009, 84% of the 62 patients coded as NOS had information on a key element (ie, % of blasts in bone marrow) in ICD-O-3 coding and other classifications of MDS subgroups. These findings suggest that central cancer registries may want to work with hospital tumor registrars in improving reporting of specific MDS subgroups using ICD-O-3. The addition of % blasts from pathology reports to the site-specific factors for MDS in the Collaborative Staging System could be proposed.

Key words: cancer registries, cancer surveillance, leukemia, myelodysplastic syndromes

Introduction

Myelodysplastic syndromes (MDS) are a group of neoplastic (clonal) myelod disorders, involving ineffective hematopoiesis with persistent and unexplained dysplasia in at least 1 myeloid bone-marrow lineage (erythroid, granulocytic and/or megakaryocytic), but most frequently involving unexplained refractory anemia (RA). MDS is most commonly diagnosed in the (growing) elderly population.

There is under-reporting of diagnosed cases to cancer registries, and also underdiagnosis of MDS (especially in elderly persons with unexplained anemia). MDS cases diagnosed since 2001 are reportable to US central cancer registries as malignant neoplasms, with the advent of the International Classification of Diseases for Oncology Version 3 (ICD-O-3). This includes registries in the National Cancer Institute’s Surveillance, Epidemiology and End Results (SEER) Program and the Center for Disease Control and Prevention’s (CDC’s) National Program of Cancer Registries (NPCR). NPCR reports and online data systems on cancer incidence by site/type, have grouped MDS under “miscellaneous” sites. For MDS (defined by ICD-O-3 Morphology codes M9980-89) in SEER registries, the annual age-adjusted incidence rate has been around 3-4 per 100,000 per year. A special report on MDS in a hematology journal used incidence data for 2001-2003 from both SEER and NPCR registries, but survival data were limited to SEER registries (which routinely follow-up all patients to ascertain vital status and causes of death).

MDS are heterogeneous with regard to prognosis, including survival rates and degree of risk of transformation (progression) to leukemia, and also with regard to the types of therapy or supportive care involved, and ICD-O-3 codes for specific MDS subgroups capture some of this heterogeneity. “Nonprogressive” forms of MDS have been defined as refractory anemia (RA) and mild decreases (cytopenias) in peripheral blood counts for other cells (neutrophils and platelets), but without increases in myeloblast (“blast”) cells in bone marrow, while more “progressive” forms have more severe cytopenias and high % blasts, with higher risk (40%-50%) of transformation of MDS to acute myeloid leukemia (AML).

Central cancer registries may want to assess the need for improving the coding of specific MDS subgroups in their databases, for use in surveillance of trends in survival (ie, SEER registries) and/or treatment (SEER and NPCR) by MDS subgroups. In those US central cancer registries that have published incidence data on MDS and its subgroups (as coded in ICD-O-3), about 50%-67% of incident cases of MDS have been coded as “not otherwise specified” (NOS) (ICD-O-3 M9989). The SEER Program involves data from 17 population-based registries, covering about 26% of the US population. Some 54.7% of all 16,230 MDS cases diagnosed in 2004-2008 were coded as NOS. This proportion is similar to that (13,908/40,948 or 56.1%) based on data on diagnoses in 2001-2003 from all high-quality US registries combined that cover about 82% of the US population, in the...
North American Association of Central Cancer Registries (NAACCR). The proportions of cases with specific MDS subgroup codes using ICD-O-3 in central cancer registry databases are probably not representative of all MDS patients. In a SEER database, for example, only a small proportion of MDS cases (323/16,230 or 2.0%) diagnosed in 2004-2008 were coded to the MDS subgroup “isolated 5q deletion” (ie, deletion of the long arm of chromosome number 5) (ICD-O-3 M9986), or 4.4% (323/7353) after excluding cases coded to M9989. In large clinical series, however, 5q deletion is the most common cytogenetic abnormality, found in about 10%-15% of MDS patients, and is important because of its association with better prognosis and with a specific treatment (ie, the drug lenalidomide, which is a thalidomide analog). In an analysis of SEER data for MDS patients diagnosed in 2001-2003, median observed survival after diagnosis (with up to 3 years of potential follow-up) was 28 months for RA and >36 months for RA with ring sideroblasts, vs only 11 months for RA with excess myeloblasts (“blasts”) (RAEB, or M9983). Small numbers of patients in MDS subgroups (eg, only 64 with 5q deletion) precluded analysis for certain subgroups. The largest subgroup of MDS, however, was NOS (1,687 patients), which had an intermediate observed survival rate, suggesting that it may comprise an (unknown) mixture of specific subgroups. In a study using 1,841 MDS cases diagnosed in 1995-2006 in the Veterans Affairs Central Cancer Registry, the median observed survival was 2.1 years but highest for RA (3.4 years) and 5q deletion (3 years) subgroups and lowest for RAEB (0.7 years); MDS NOS (median survival 1.85 years) comprised 67% of all 1841 cases.

The relative survival rate (RSR) of cancer patients in registries is calculated as survival relative to that expected from mortality rates in the general US population, taking into account age, gender, race, and time period. The 3-year RSR was only 45% for MDS patients diagnosed in 2001-2003 in the population-based cancer registries of the SEER Program, but RSR was not examined by subgroup of MDS. The limited data on specific MDS subgroups in national cancer registry databases complicates the interpretation of trends in survival rates (RSRs), because increases in the RSR could occur simply from earlier diagnosis of cancer in patients or (for MDS) changes in the distribution of prognostic subgroups, rather than improvements in treatment.

The reporting of MDS cases as NOS (M9989 code) vs specific subgroups to a central cancer registry could vary among central cancer registries, and also among facilities reporting to a single central registry. Variation could reflect differences in such factors as patient characteristics (eg, age at diagnosis of MDS), availability of pathology reports to hospital registrars, and practices of local pathologists. The present study examined variation in the proportion of MDS cases coded as NOS among Connecticut facilities reporting to the population-based Connecticut Tumor Registry (CTR) which is part of the SEER Program. In addition, pathology reports in the CTR were sought for a sample of MDS cases, to assess the presence of certain information (mainly, % blasts in the bone marrow sample) relevant to the classification of MDS subgroups, especially for those cases coded to M9989. The results of these analyses could be useful in assessing the prospects for improving information on specific MDS subgroups in central cancer registries.

### Methods

Using ICD-O-3 Morphology codes M9980-89, all cases of MDS diagnosed in 2001-2009 in Connecticut residents and reported to the CTR were identified in a data file prepared in September 2011. The CTR, located in the Connecticut Department of Public Health (DPH), receives data on reportable tumors as required by state regulations covering all state-licensed hospitals and clinical laboratories. Confidentiality of CTR data is also protected under state regulations and this study was part of a project approved by the DPH Human Investigations Committee.

The small number ascertained by death certificate only were excluded, because year of diagnosis was unknown. Included were refractory anemia (RA) (M9980), RA with sideroblasts (M9982), RA with excess of blasts (EB) or

### Table 1. Distribution of the Proportion of all MDS Cases Diagnosed in 2001-2009 that were Coded as Unspecified MDS (ICD-O-3 Morphology code M9989) for 30 Connecticut Facilities Reporting to the Population-based Connecticut Tumor Registry

<table>
<thead>
<tr>
<th>% of Cases Coded M9989</th>
<th>Number of facilities</th>
<th>Number of MDS cases</th>
<th>Coded as M9989</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>% of all 30</td>
<td>Per facility</td>
</tr>
<tr>
<td>&lt;40.0</td>
<td>6</td>
<td>20.0</td>
<td>1-77</td>
</tr>
<tr>
<td>40.0-49.9</td>
<td>4</td>
<td>13.3</td>
<td>9-106</td>
</tr>
<tr>
<td>50.0-59.9</td>
<td>5</td>
<td>16.7</td>
<td>2-68</td>
</tr>
<tr>
<td>60.0-69.9</td>
<td>7</td>
<td>23.3</td>
<td>6-151</td>
</tr>
<tr>
<td>70.0-79.9</td>
<td>4</td>
<td>13.3</td>
<td>22-64</td>
</tr>
<tr>
<td>80.0-100</td>
<td>4</td>
<td>13.3</td>
<td>5-18</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>100</td>
<td>1-151</td>
</tr>
</tbody>
</table>

Note: For these 30 facilities, the range for % cases coded as M9989 was 0%-100%, with a mean of 57.6%, standard deviation 22.6%, and median of 57.6%.
RAEB (M9983), RAEB in transformation (RAEB-T) (M9984), refractory cytopenia with multilineage dysplasia (M9985), MDS with isolated 5q deletion (M9986), therapy-related MDS (M9987) and MDS “not otherwise specified” (NOS) (M9989).

The French-American-British Cooperative Group (FAB) classification (published in 1976 and revised in 1982) was dominant among clinicians until around 2000 when the World Health Organization (WHO) system prevailed.12 These systems defined subgroups that differ in prognosis. The FAB system included RAEB-T under MDS, defined as 20%-29% blasts in bone marrow, while 30%+ blasts indicated leukemia. Leukemia, however, was redefined as 20%+ blasts in bone-marrow in a 1999 WHO publication and thereafter.12,13 The RAEB-T subgroup is considered “obsolete” in ICD-O-3, although small proportions of MDS cases are coded to this subgroup (M9984) in SEER2,3; there is some controversy regarding RAEB-T in the clinical literature, however, and some clinicians may still use the FAB classification.12 Therapy-related MDS (M9987) involves a history of treatment with various chemotherapy agents (eg, alkylating agents and topoisomerase inhibitors) and/or radiotherapy for a variety of cancer types (eg, lymphomas, multiple myeloma and certain solid cancers including respiratory system and breast), and poor prognosis.1 It is grouped not with MDS but with AML and related neoplasms in the WHO system.13

For the present study, the proportion of MDS cases coded as NOS (M9989) was examined for the 1,011 MDS cases reported by each of the 30 reporting facilities in Connecticut, after excluding the small number of cases ascertained by death certificate only. Variation in this proportion was examined among the facilities, and in relation to the number of total MDS cases reported by each facility.

Pathology reports for MDS cases diagnosed in Connecticut facilities in the latest 2 years (2008 and 2009) were sought for manual review by a certified tumor registrar (CP); information on % blasts in the bone marrow sample was of main interest.

Data analyses were conducted using Microsoft Excel. The 95% confidence intervals (CIs) on each proportion were calculated based on the normal distribution.

**Results**

_Variation in Proportion of Unspecified (Code M9989)_

**MDS Cases among Facilities**

The proportion of MDS cases with code M9989 among all 1,011 MDS cases at the 30 reporting facilities in Connecticut combined was 56.7% (573/1011) but varied from 0%-100% (Table 1). The mean for the 30 facilities was 57.6% (median also 57.6%), but the wide variation is indicated by the standard deviation of the distribution (22.6%).

Some statistical variability is expected, in view of the small sample size. Of the 9 other facilities with <50% of cases coded as M9989, 2 had <10 (3 and 9) total MDS cases, and 95% CIs were wide (data not tabulated). For 3 of the 7 facilities that had >20 total cases, however, the upper limit of the 95% CI (not tabulated) was lower than the total figure of 56.7% for all cases from the 30 facilities combined.

Seven facilities had 60%-69.9% of cases coded as M9989, and 1 of these proportions was based on large numbers (103/151 cases or 68.2%, 95% CI=60.8-75.6%). Of the 4 facilities with 70%-79% of cases coded as M9989, all had >20 total MDS cases (22-64) (Table 1); the lower limit of each 95% CI was >56.7% (data not tabulated). All 4 facilities with 80-100% of cases coded as M9989 had 18 or fewer total MDS cases (5-18, Table 1). These high proportions could reflect either statistical (sampling) variability or an impact of small MDS case-loads on the reporting of specific MDS subgroups on pathology reports.

The proportion of cases coded as NOS was higher among patients diagnosed at age 85+ years (109/175 or 62.3%, 95% CI=55.1-69.5) vs <85 years (464/836 or 55.5%, 95% CI=52.1-58.9). This would be consistent with less use of invasive procedures such as bone-marrow aspirate and biopsy in older patients for confirmation of diagnosis. The 6 facilities with <40% NOS for all ages combined (Table 1), however, also had low proportions of NOS cases within the age group <85 years at diagnosis (0%-40%), and the 8 facilities with 70%+ NOS for all ages also had high proportions within age <85 years (ie, all >70%) (data not tabulated).

**Information in Pathology Reports**

Pathology reports were sought for the 222 MDS patients diagnosed in 2008-2009 (98 in 2008 and 124 in 2009), of which 197 (88.7%) were microscopically confirmed. A pathology report was available in the CTR for 130 (66.0%) of the 197 patients. Pathology reports are not routinely submitted to the CTR by 3 Connecticut hospitals (1 small, 1 mid-size and 1 large, defined by average number of incident cancers reported annually to the CTR).

Data on % blasts in bone marrow were found in pathology reports for 113 (86.9%) of the 130, including 52 (84%) of the 62 coded as M9989 (Table 2). Of these 52 cases, 2 with 5% blasts and 10 with >5% blasts (Table 2) could be tentatively classified as RAEB (M9983); that is, aside from therapy-related MDS, all other specific MDS subgroup coded require <5% blasts. Two of the 10 cases with elevated blasts had 20%+ blasts, and could be regarded as RAEB-T (M9984), but this is an obsolete code for MDS in ICD-O-3 (ie, redefined as leukemia, as in the WHO system).

Using the data for the 113 cases with some information on % blasts (Table 2), the proportion with blasts either <5% or “not elevated” was 75.0% (39/52) among for M9989 cases vs 49.2% (30/61) among those coded to a specific MDS subgroup (p<.01 for difference between these 2 groups).

Although the main purpose of reviewing pathology reports was to assess information on cases coded as MDS NOS (M9989), the data on % blasts (Table 2) are also useful for quality-control by the registry. That is, of the 26 cases coded with M9983 and M9984, which are both defined as high % blasts in bone marrow, a high % blasts was confirmed for all 25 (Table 2) that had information available on % blasts.
Information in the CTR indicated that cytogenetic tests were performed for some patients, but only a few pathology reports had the results of these tests.

Of the 130 patients with a pathology report, the SEER tumor sequence code for MDS was 02 or greater for 40 (30.8%), indicating that a reportable tumor(s) was known to have been previously diagnosed. Of these 40, 3 were already coded as therapy-related (M9987) and 22 had information that chemotherapy and/or radiotherapy had been received. Some other study findings, however, highlight certain ties in Connecticut reporting to the CTR (Table 1), however, several had relatively high or low proportions of NOS cases that were unlikely due to chance or differences in the age distribution of all MDS cases at the facility. Similar studies are needed comparing MDS subgroups by reporting facility in other population-based central cancer registries.

The finding that a large proportion of pathology reports had information on % blasts in the bone marrow for patients coded to the NOS subgroup (Table 2) has implications for potential improvement in the specificity of reporting of MDS subgroups to the CTR. Studies are needed with larger samples of pathology reports, from central registries that collect data on patient survival and/or treatment, which differ by MDS subgroup. Studies are especially needed from central cancer registries that obtain pathology reports for all patients (eg, through electronic reporting). Supplemental pathology reports or addenda should be sought for the results of cytogenetic tests done for MDS patients, which may be delayed for several months during the process of diagnosis.23 Cytogenetic findings are needed to distinguish the MDS subgroup 5q deletion syndrome (M9986 in ICD-O-3), which also requires <5% blasts in bone marrow findings under the WHO classification system.13

The higher proportion with <5% or “not elevated” blasts among patients coded as NOS than among those coded to a specific MDS subgroup (Table 2) suggests that routine data on survival rates for MDS cases by subgroup that are reported by cancer registries may be biased.

Some Issues in Quality Control and Interpretation of Data, Including Data in Pathology Reports

In the present study, all cases coded to MDS subgroups that are defined by a high % blasts were confirmed to have high % blasts in the pathology report (Table 2), which supports the high quality of the data routinely coded in the registry. Other study findings, however, highlight certain

### Table 2. Findings on Proportion of Blast Cells (% Blasts) in Path Reports Obtained for 130 MDS Patients Diagnosed in 2008-2009 and Reported to the Connecticut Tumor Registry

<table>
<thead>
<tr>
<th>Bone marrow blasts</th>
<th>All patients</th>
<th>Selected subgroups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5% or Not Elevated</td>
<td>42</td>
<td>32.3</td>
</tr>
<tr>
<td>“No increase,” “not elevated”</td>
<td>27</td>
<td>20.8</td>
</tr>
<tr>
<td>Subtotal</td>
<td>69</td>
<td>53.1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥5% or Elevated</td>
<td>5</td>
<td>3.8</td>
</tr>
<tr>
<td>&gt;5%</td>
<td>33</td>
<td>25.4</td>
</tr>
<tr>
<td>20%+</td>
<td>(3)</td>
<td>2.3</td>
</tr>
<tr>
<td>“Increase(d), “elevated,” etc.‡</td>
<td>6</td>
<td>4.6</td>
</tr>
<tr>
<td>Subtotal</td>
<td>(44)</td>
<td>33.8</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No information in pathology report</td>
<td>17</td>
<td>13.1</td>
</tr>
<tr>
<td>Total Patients</td>
<td>130</td>
<td>100</td>
</tr>
</tbody>
</table>

*RAEB: Refractory anemia with excess blasts (M9983); includes two patients coded as RAEB in transformation (RAEB-T) (M9984).

‡Subgroup of cases coded to M9989 that also had a SEER tumor sequence code “02” or higher, indicating a previously diagnosed reportable tumor, for which CTR records showed that chemotherapy and/or radiotherapy had been received.

‡Other terms found in path reports include “very high” and “excess”.

Note: Most reports with mention of blasts had the specific % of blasts, but some had a range (“<1,” “5-10%,” “15-19%,” etc.).
issues related to the quality of pathology reports and their interpretation by those who might use data on MDS subgroups from central cancer registries. Blast percentages at or close to certain specific levels such as 5% or 20% are especially relevant to the proper classification into MDS subgroups, as well as to treatment decisions. In the present study of pathology reports, 5 patients reportedly had 5% blasts in bone marrow (Table 2) and several others were reported only as “<5." Manual microscopic counts of blasts have been recommended using a sample of 200 total cells for peripheral blood and 500 microscopic counts of blasts have been recommended using 2) and several others were reported only as ‘<5." Manual microscopic counts of blasts have been recommended using a sample of 200 total cells for peripheral blood and 500 microscopic counts of blasts have been proposed by an international working group of hematologists and hematopathologists specializing in MDS.

The 95% CI on the % blasts in bone marrow for each patient should be included in the pathology report. Alternatively, the number of cells examined should be stated, so that the CI can be calculated. Clinicians may want to be aware of the statistical imprecision in the reported value for % blasts, in making decisions about classification and treatment; statistical uncertainty is separate from inaccuracy in distinguishing blasts morphologically from other cells examined.

These issues of statistical precision of reported data on % blasts, and variation in morphological criteria used to define blasts, should be examined in future studies involving larger samples of pathology reports for MDS cases in central cancer registries.

For erythroid-predominant MDS (ie, at least 50% of nucleated bone-marrow cells are erythroblasts) counting blasts from nonerythroid marrow cells rather than total marrow cells has been proposed, and this is another issue that might be addressed in future studies using pathology reports.

The risk of transformation to leukemia or other adverse outcomes is low for MDS patients or subgroups with <5% blasts in bone-marrow samples but is moderately influenced by adding cytogenetic findings. As already noted, obtaining information on cytogenetic findings from supplemental pathology reports also would be useful in future studies.

The Potential Role of Tumor Registrars and Registries in Improving Reporting of Specific MDS Subgroups

Surveys are needed to assess whether tumor registrars at hospital and central cancer registries ever change M9989 codes, based on their examination of findings in pathology reports or queries sent to pathologists or clinicians (for more-specific codes). The list for queries could also include MDS cases not identified as therapy-related, but found to have a history of chemotherapy and/or radiotherapy for certain previously diagnosed cancers recorded in the tumor registry.

The list of reportable tumors is periodically updated by the CTR, as stipulated in state regulations, and could include a note mentioning the importance of reporting of specific ICD-O-3 codes for MDS specific subgroups vs NOS. This matter also could be addressed at meetings of hospital cancer committees, state cancer registrars, SEER and NAACCR, state medical societies, and hematologists who treat MDS patients.

Suggestions for Modifying the Collaborative Staging System’s Site-Specific Factors, Specifically for MDS

The Collaborative Staging System (CSS) was implemented for cases diagnosed starting in 2004, primarily to collect data consistent with the American Joint Committee on Cancer’s (AJCC) Cancer Staging Manual, Seventh Edition. CSS was a joint effort of NPCR, AJCC, the National Cancer Registrar’s Association, SEER, NAACCR, the National Cancer Institute of Canada, and the American College of Surgeons Commission on Cancer, for use in US and Canadian cancer registries.

In the second version of CSS, up to 24 site-specific factors (SSFs) are collected as additional clinical and/or prognostic data that may be of interest to researchers. The only SSF collected for MDS, however, involves a specific acquired genetic mutation (JAK-2 positivity or negativity) in the Janus kinase 2 (JAK-2) protein found in peripheral blood. Some cancer registries receive test results directly from laboratories. Information on % blasts in the bone marrow could be proposed for addition to the SSFs for MDS, along with mention of cytogenetic findings such 5q deletion (with vs without other chromosomal anomalies), found in a pathology report or addendum.

Also included for consideration in SSFs for MDS could be a history of chemotherapy and/or radiotherapy for a previously diagnosed cancer (from the pathology report and/or the central registry database), in order to identify therapy-related MDS (M9987). Patient clinical history may be a section of a pathology report, including certain treatment for a previous cancer(s). A variety of cancer types and treatments have reportedly been associated with therapy-related MDS. The proportion of MDS cases coded as therapy-related has been low in central cancer registry reports. In a special analysis of a SEER database, however, 26% of MDS patients were found to be “secondary” (ie, diagnosed after a previous cancer); survival was 13% lower than for MDS patients without a previous cancer.

Specific bone-marrow cytogenetic findings also may be useful in identifying possible treatment-related MDS. In the present study (Table 2), the 10 MDS patients coded as M9989 but with a history of chemotherapy and/or radiotherapy for a previous cancer could be considered as therapy-related (M9987). Therapy-related cases in clinical series reportedly have poor outcomes (median survival) regardless of % blasts. For clinical uses and research studies, however, cytogenetic findings (used in prognostic scoring systems) may affect prognosis of patients with therapy-related MDS. Again, supplemental reports with cytogenetic findings would be relevant to defining prognostically relevant subgroups of MDS using ICD-O-3 codes, or in research studies using other classifications.

Potential Role of Central Cancer Registries in Research Studies on Prognosis of MDS Patients

Some central registries may want to conduct
population-based studies on prognosis of MDS, using data on specific subgroups other than those defined by ICD-O-3 coding.

The International Prognostic Scoring System (IPSS) and the WHO Performance Scoring System (WPSS) (Table 3), developed for de novo (ie, not secondary or therapy-related) MDS, involved 4 categories of risk of death or transformation to AML. A total numeric score was based on % blasts in bone marrow, number of cytopenias (1-3) and type of chromosomal abnormalities (good, intermediate and poor groups).25,26 Survival of MDS patients in clinical series declines progressively, and risk of transformation to leukemia increases, as this prognostic score increases.19,25,26

Most prognostic scoring systems include % blasts in bone marrow.22 WPSS26 includes a patient’s blood transfusion requirement, cytogenetic findings, and WHO classification subgroup (using % blasts). Medicare databases have been used to identify elderly MDS patients who had blood transfusions and/or cytopenias.27,28 About 20% of newly diagnosed MDS patients, however, are diagnosed before age 65 years.23,30 Data on transfusions were formerly collected by SEER registries as “other treatment.”29

The “hematopoietic stem cell transplantation comorbidity index” (Table 3) modifies the well-known Charlson comorbidity index by capturing additional comorbid conditions such as obesity and psychiatric disorders.30,31 Studies using SEER-Medicare linked databases also have shown the prognostic importance of certain comorbid conditions present around the time of MDS diagnosis.28,32 Central cancer registries also could examine the ICD-9-CM (and/or ICD-10-CM) comorbidity codes included as data items (up to 10 data fields) in the NAACCR record used to report incident cases diagnosed at any age.33-35 Smoking around the time of MDS diagnosis has been reported as an adverse prognostic factor,36 and some central cancer registries have attempted to collect smoking history.37,38

A comorbidity severity scale developed at the MD Anderson Cancer Center has been shown to predict outcomes of MDS patients, independent of IPSS risk score and patient age.39 The prognostic value of this scale, as a complement to IPSS, has been confirmed for MDS patients at another large US cancer center.40

Data on supportive care for MDS patients, which would not be defined as cancer-directed treatment (and not coded) by SEER29 but may influence survival, also could be included in registry-based studies of prognosis. For elderly MDS patients, for example, SEER-Medicare linked databases have been used to estimate the use of erythropoiesis-stimulating agents (intended to control anemia and reduce the need for blood transfusions).41 These agents are more effective in “low-risk” MDS patients, but 50.4% of patients in the entire sample were coded as MDS NOS and could not be classified in low- vs high-risk subgroups using ICD-O-3 codes.41 Use of the hypomethylating drugs azacitidine and decitabine (approved by the Food and Drug Administration for MDS, starting in 2004) in elderly MDS patients was shown to be more common in those with RAEB or multilineage dysplasia vs RA.42 These agents (not routinely coded in SEER databases), considered by clinicians as low-intensity chemotherapy, inhibit DNA methyltransferase and prevent DNA synthesis in abnormal hematopoietic cells in bone marrow, and may decrease risk of transformation to leukemia and improve survival in some MDS patients with otherwise poor prognosis.9 The limitations of registry data on MDS subgroups must be addressed by other methods (as suggested in this report) by registries that link with administrative databases such as Medicare for data on treatment.

Recent Changes in Coding and Other Developments Relevant to Coding of MDS Subgroups

Recent changes to ICD-9-CM (Clinical Modification) used by hospital discharge databases include codes for “low-grade” (code 238.72) and “high-grade” (238.73) MDS, MDS with 5q deletion (238.74) as well as MDS NOS (238.75). These codes have been included in the ICD-9-CM case-finding code list for abstracting of reportable tumors for the SEER Program, effective October 1, 2006,43 and for NPCR.

### Table 3. Criteria for MDS Prognostic Scales or Scoring Systems, and Other Potential Prognostic Indicators: Potential Sources of Relevant Information

<table>
<thead>
<tr>
<th>Criterion or predictor</th>
<th>Scale(s) involved*</th>
<th>Potential source(s) of information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevant to ICD-O-3 coding of MDS subgroups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BM blasts (%)</td>
<td>1†</td>
<td>Path report</td>
</tr>
<tr>
<td>BM cytogenetic features</td>
<td>1,2</td>
<td>Path report, and supplement/addendum report</td>
</tr>
<tr>
<td>Cytopenia/dysplasia in BM, PB</td>
<td>1,2</td>
<td>Path report; lab report; hospital discharge databases; Medicare files (age 65+ years)</td>
</tr>
<tr>
<td>Relevant to prognostic studies using cancer registries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood transfusions</td>
<td>2</td>
<td>Medicare files (age 65+ years); hospital discharge databases (patients of all ages); information reported to cancer registries</td>
</tr>
<tr>
<td>Comorbid conditions</td>
<td>3</td>
<td>Hospital record; some tumor registries (using items on electronic report from hospitals); Medicare files; hospital discharge databases</td>
</tr>
</tbody>
</table>

*1=International Prognostic Scoring System, for MDS (IPSS); 2=WHO-based Prognostic Scoring System, for MDS (WPSS); 3=Hematopoietic stem cell transplantation comorbidity index.† % blasts is not included as separate indicator in WPSS, but is included in the WHO classification codes that is used in WPSS.26

BM=bone marrow.
therapies. Therapy-related MDS, however, is coded as “low-grade,” despite the association with poorer prognosis in most studies.  

Studies are needed on how often the specific codes (other than 238.75 for NOS) are available and are being used by hospital tumor registrars in reporting to central cancer registries in the NPCR and SEER Program, in casefinding and quality control efforts. Quality-control studies should assess accuracy of these ICD-9-CM codes, in comparison to medical records and pathology reports.

In response to the 2008 WHO classification system for myeloid neoplasms, a hematopoietic working group led by NCI-SEER, with representation from several national organizations, was formed to develop rules, guidelines and an interactive database (along with the Coding Manual, at http://seer.cancer.gov/tools/heme). NAACCR lists these items under “resources and training programs for cancer registries” (http://www.naaccr.org). Online training is approved by National Cancer Registrars Association (NCRA) for continuing education units (http://seer.cancer.gov/heme/training). Links to the 2012 hematopoietic database and manual are also available (at http://ncra.usa.org), and sessions are periodically held at national meetings of cancer registrars/surveillance organizations.

The aim of the working group was to assist registrars in determining case reportability, coding of multiple primary cancers in the same patient, and coding of site, histology and grade for hematopoietic neoplasms. The issue of transformation was discussed thoroughly, in view of changes in clinical knowledge; transformations from chronic to acute phase occurring after 21 days are considered 2 different primaries for hematopoietic tumors diagnosed starting in 2010. The Manual and database were cited with regard to coding of multiple primaries and also casefinding lists (ICD-9-CM and ICD-10-CM), in the 2010 NAACCR implementation guidelines and recommendations, under “hematopoietic and lymphoid neoplasm rules.” Previous coding practices artificially reduced the ascertainment of acute diagnoses (eg, leukemias). Thus, leukemias that would previously have been considered disease progression/transformation (not reportable or sequenced as new primaries) are now coded as new primaries.

The 2010 Hematopoietic and Lymphatic Coding Manual also states that “the most specific histology” recorded in the pathology report or addendum should be reported by the registrar. This could lead to increased coding of more-specific MDS subgroups (vs NOS), and it will be important to assess how hospital tumor registrars and central registries in the SEER Program and in the NPCR use this Manual and follow its recommendations. Although SEER registries must meet high data-quality standards not only for casefinding but also for coding of data including percentage unknown for various items (eg, unknown or ill-defined primary site and “malignant NOS” morphology), items such as M9989 codes have not been included in Quality Profiles. Nevertheless, individual central registries could track any improvements in coding of specific MDS subgroup over time for each reporting facility, as well as query facilities on their use of the Hematopoietic and Lymphatic Coding Manual.

These new hematopoietic coding rules will result in improvement in the registration of incident leukemia cases. These rules will also facilitate registry-based studies on the risk and predictors of leukemia after diagnosis of MDS; however, improvement in the coding of specific MDS subgroups in central registries would enhance such studies. Efforts to quantify the extent of under-registration of incident MDS cases in SEER registries, by using Medicare databases for elderly persons, may lead to the registration of substantially larger numbers of MDS cases. The impact of these efforts on the coding of MDS subgroups appears uncertain.

Conclusion
Analyses of data from the CTR of the SEER Program indicate substantial variation in the proportion of MDS NOS cases among reporting facilities. Efforts by hospital and central cancer registries, including querying clinicians regarding MDS cases coded as NOS, could lead to improvement in specificity of coding of MDS subgroups. The addition of certain site-specific factors (eg, % blasts in bone marrow) for MDS to the Collaborative Staging System might be advocated. In a sample of pathology reports, information was found on % blasts in bone marrow, which would be useful in coding specific MDS subgroups. Similar studies, preferably also including cytogenetic findings from bone-marrow samples, are needed in other central cancer registries.

References


Leukemia as a Cause of Death among Patients with Myelodysplastic Syndromes (MDS) in a Population-Based Cancer Registry: Improving Estimates of MDS-Related Mortality in the Population

Anthony P. Polednak, PhD*; Cathryn Phillips, CTR®

Abstract: Myelodysplastic syndromes (MDS), a heterogeneous group of myeloid neoplasms diagnosed mostly in elderly persons, are of increasing interest in an aging population and are associated with variable risk of progression to acute myeloid leukemia (AML). The numbers of deaths related to MDS in the population are underestimated in routine US cancer mortality statistics which are based on only the underlying cause (UC) rather than multiple causes (MCs) of death recorded on death certificates. Additional MDS-related deaths, however, may be missed if some MDS patients die with mention of leukemia but not MDS on their death certificate. This requires studies of MCs of death among all MDS patients in population-based tumor registries. This study examined MCs of death among patients diagnosed with MDS in 2001-2009 and reported to the population-based Connecticut Tumor Registry. MDS was the UC for 199 deaths (25.7% of all 773) and was coded as other than UC for 160 (20.7%). Another 121 (15.7%) death records, however, had leukemia without mention of MDS; the majority were coded to AML and most of the others as unspecified type of acute leukemia. If these 121 deaths are added to the 359 with mention of MDS, the total of MDS-related deaths would be 480 (or 62.1% of all 773 deaths). A total of 178 deaths (23.0% of all 773) were coded to leukemia as the UC, and would be included with leukemia (not MDS) in routine cancer mortality statistics. Leukemia diagnosed since 2010 in MDS patients is reportable to registries as a new primary cancer. This new rule will help central cancer registries to confirm leukemia diagnoses coded on death records, as part of the process of improving surveillance of cancer mortality rates in the population.

Key words: cancer mortality, cancer registries, cancer surveillance, leukemia, multiple causes of death

Introduction

Myelodysplastic syndromes (MDS), ranging from refractory anemia (with or without other cytopenias) with an indolent course to high risk of rapid progression to acute myeloid leukemia (AML),¹ have been reportable to US cancer registries for diagnoses since 2001, with incidence rates of about 3-4 per 100,000 per year but higher in older age groups.²³⁴

MDS is of increasing relevance in view of the aging of the population; the problem of undiagnosed MDS (especially among elderly persons with unexplained anemia);⁵ ongoing efforts to improve MDS diagnosis and therapy;⁶⁷⁸⁹; and efforts at surveillance of outcomes among MDS patients.¹⁰¹¹

Surveillance of cancer mortality based solely on the underlying cause (UC) of death rather than multiple causes (MCs) recorded on death certificates may be justified for most (but not all) types of cancer.¹²¹³ Among all 18,304 deaths with mention of MDS in 2005-2006 among US residents, 8,309 (45%) had MDS coded as other than the UC; in addition to leukemia, various chronic diseases (eg, of the circulatory and respiratory systems) were coded as UC when MDS was mentioned on the same death record.¹⁴

However, some MDS patients may die from MDS-related causes (especially leukemia) without MDS being mentioned at all on the death certificate, and this should be examined by follow-up studies of causes of death among MDS patients in population-based cancer registries. In a clinical series of “low risk” (ie, of progression to leukemia) MDS patients who had been referred to MD Anderson Cancer Center, MDS-related deaths were commonly attributed to transformation of MDS to leukemia, infections (including pneumonia) and certain other conditions.¹⁵ The sample was not population-based and deaths were not coded according to the rules of the International Classification of Diseases (ICD) used for statistical reports on cancer mortality in the US population. Infection due to MDS (eg, with neutropenia), for example, should be coded to MDS as the UC in ICD.

The present study examined MC of death, including deaths with leukemia but not MDS on the death certificate, among MDS patients in the population-based Connecticut Tumor Registry (CTR), and the relevance of the findings to improving surveillance of estimated MDS-related deaths in populations.
Methods

In September 2011, a data file was obtained from the CTR after approval of the study by the Human Investigations Committee of the Connecticut Department of Public Health. The file included data on all 1085 Connecticut residents diagnosed from 2001 through 2009 (the most recent year with complete data) with MDS. The CTR is part of the US National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) Program of high-quality population-based cancer registries.2,3

MDS was defined by International Classification of Diseases for Oncology Version 3 (ICD-O-3) Morphology codes M9980-99892,3 including: refractory anemia (RA) (M9980), RA with sideroblasts (M9982); RA with excess blasts (EB) or RAEB (M9983); RAEB in transformation (RAEB-T) (M9984); refractory cytopenia with multilineage dysplasia (M9985); MDS with isolated 5q chromosome deletion (M9986); therapy-related MDS (M9987); and MDS not otherwise specified (NOS) (M9989). RAEB-T, originally defined as 20%-29% blasts in the bone marrow, and therapy-related MDS (ie, after treatment of a previous cancer with certain chemotherapy agents and/or radiotherapy) were classified with leukemia in the World Health Organization system.5,16 Differences and temporal changes in classification systems complicate surveillance of MDS.

Connecticut state regulations require reporting of newly diagnosed reportable tumors from all state-licensed hospitals and clinical laboratories. Through reciprocal reporting arrangements with cancer registries in all 3 contiguous states and several others (including Florida), Connecticut residents diagnosed and/or treated are identified by the CTR. Because MDS can occur (sometimes along with leukemia) after treatment for a previous cancer, and transformation to leukemia occurs,16 patients diagnosed with multiple tumors were included.

Starting with deaths in 2005, the CTR added ICD-10 codes (found anywhere on a death record) for MDS (ICD-10 code D46) to the list of tumors to be included from Connecticut state death-record files for “death clearance” and follow-back to ascertain reportable tumors not found to be registered (yet) in the CTR database. The 41 MDS cases eventually coded (after follow-back efforts) in the CTR as ascertained by “death certificate only” (DCO) were excluded from this study because year of diagnosis of MDS was unknown. No MDS cases were ascertained by autopsy alone. Of the 1,044 remaining cases, 951 (91.1%) were coded as having been microscopically confirmed.

Data on vital status were from follow-up conducted by the CTR through linkages with Connecticut death records and motor vehicle license files, the Social Security Administration, and the Center for Medicare and Medicaid Services. Of the 1,044 patients, 773 (74.0%) deaths were included in the CTR database, and most decedents (721 or 93.3%) were still residents of Connecticut at the time of death. Of the 271 patients not known to have died, year of last contact was 2010-2011 for 199 (73.4%); however, it was coded as 2009 for 22 (34%) of the 65 decedents who had been diagnosed with MDS in 2009. Deaths occurred as late as 6-7 years after diagnosis (ie, for 9 of all 80 deaths in patients diagnosed in 2001).

Because the distribution of causes of death could differ by length of time after MDS diagnosis (and also by calendar year of death), causes were examined separately for patients diagnosed in 2001-2004 and 2005-2009. These 2 groups differed little in age at diagnosis of MDS (85/429 or 19.8%, vs 142/615 or 23.0% diagnosed at <65 years, and median age 76 vs 77 years), and gender (% male 245/429 or 57.1% vs 341/615 or 55.4%) (data not tabulated).

Causes of death, coded to ICD Version 10 (ICD-10), for deceased patients in the CTR had been obtained through linkages with Connecticut state vital record (death-certificate) files and the National Death Index (NDI Plus). Only the UC is reported to SEER. The SEER Data Management System (DMS)16 used by the CTR, however, has ICD-10 codes for 17 causes per decedent. Copies of death certificates are not routinely sought by the CTR, which include only conditions regarded by the certifying physician or coroner as being part of the chain of events leading to death (Part I) or as contributing to death in some other fashion (Part II). Deaths coded with any mention of MDS (ICD-10 code D46) and/or leukemia (ICD-10 C91-C95) were identified, along with selected other causes.

Leukemia occurring as transformation (progression) of MDS was not reportable to the CTR (or SEER) as a separate primary cancer until diagnoses in 2010; nevertheless, some may have been reported, and CTR records for the MDS patients were searched for any reports of leukemia (ICD-O-3 M9800-9948) that were in the CTR files. Differences in the distribution of causes of death and other characteristics between groups were tested by chi-square (using Microsoft Excel).

Results

Of the 429 patients diagnosed with in 2001-2004, 360 (83.9%) had died, vs 413 (67.2%) of the 615 diagnosed in 2005-2009, but the distribution of causes of death using 8 mutually-exclusive categories was similar for the 2 time periods (p=0.22, degrees of freedom or df=7), although the frequency of unknown cause was higher among those diagnosed in 2005-2009 (Table 1).

Of all 773 deaths, 199 (25.7%) had MDS coded as the UC and another 160 (20.7% of 773) had MDS mentioned elsewhere on the death record, or a total of 359 with mention of MDS (Table 1), all but 7 (1.9%) of which were coded as MDS not otherwise specified (NOS) (ICD-10 D46.9). MDS was mentioned for only 17 (34.7%) of all 49 deaths in 2001-2002, vs (45.2%) of 122 deaths in 2003-2004, and for 83-106 deaths (42.8-50.3% of all deaths) for each time period of 2005-2006, 2007-2008 and 2009-2011 (data not tabulated).

Leukemia was mentioned on none of the 199 death records with MDS coded as the UC and another 57 (10.2% of 560) had leukemia coded as the UC. Among the deaths without any mention of MDS on the death record, leukemia was coded for 121 (14.2% of 773) including 110 as UC (Table 1).

Chronic myeloproliferative disease (ICD-10 D47.1), also known as myeloproliferative neoplasms (MPNs), was
Table 1. Causes of Death Based on Underlying Cause (UC) and Multiple Causes Recorded on Death Certificates* for 773 Deaths Identified Among 1044 Connecticut Residents Diagnosed with Myelodysplastic Syndromes (MDS)† in 2001-2009, in the Database of the Connecticut Tumor Registry

<table>
<thead>
<tr>
<th>Cause of death category</th>
<th>Diagnosed in 2001-2004</th>
<th>Diagnosed in 2005-2009</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
</tr>
<tr>
<td>1 Deaths with mention of MDS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCS as UC</td>
<td>85</td>
<td>23.6</td>
<td>114</td>
</tr>
<tr>
<td>MDS as other than UC</td>
<td>81</td>
<td>22.5</td>
<td>79</td>
</tr>
<tr>
<td>Leukemia* as UC</td>
<td>(33)</td>
<td>(9.2)</td>
<td>(35)</td>
</tr>
<tr>
<td>2 Deaths without mention of MDS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leukemia as UC</td>
<td>55</td>
<td>15.3</td>
<td>55</td>
</tr>
<tr>
<td>Leukemia as other than UC</td>
<td>(4)</td>
<td>(1.1)</td>
<td>(7)</td>
</tr>
<tr>
<td>CMPD§ as UC</td>
<td>7</td>
<td>1.9</td>
<td>11</td>
</tr>
<tr>
<td>CMPD as other than UC</td>
<td>(3)</td>
<td>(0.8)</td>
<td>(2)</td>
</tr>
<tr>
<td>Infectious disease as UC*</td>
<td>11</td>
<td>3.1</td>
<td>10</td>
</tr>
<tr>
<td>Anemia as UC</td>
<td>5</td>
<td>1.4</td>
<td>4</td>
</tr>
<tr>
<td>Other known UC</td>
<td>102</td>
<td>28.4</td>
<td>107</td>
</tr>
<tr>
<td>3 Unknown UC</td>
<td>14</td>
<td>3.9</td>
<td>33</td>
</tr>
<tr>
<td>Total deaths</td>
<td>360</td>
<td>100</td>
<td>413</td>
</tr>
<tr>
<td>Deaths as % of all patients</td>
<td>360/429</td>
<td>83.7</td>
<td>413/615</td>
</tr>
</tbody>
</table>

*Multiple causes includes the underlying cause (UC) and as many as 17 additional causes from the death certificate, available in the registry database (see text).
‡ICD-10 codes C91-95; 2 other death records with MDS mentioned had leukemia coded as other than the UC (see text), for a total of 70 (see also Table 2).
§CMPD (chronic myeloproliferative disease), ICD-10 code D47.1.
||Includes code D47.1, and 1 additional death with mention of essential thrombocytopenia (D47.3) but leukemia coded as the UC, and 2 deaths with anemia as UC (see text).
¶ICD-10 codes A00-B99.
#ICD-10 codes for unspecified aplastic anemia (D61.9) or unspecified anemia (D64.9).
Note: p=0.22 for chi-square test (7 degrees of freedom) for difference in distribution of the 8 mutually exclusive categories of cause of death for patients diagnosed in 2001-2004 vs those diagnosed in 2005-2009.

Table 2. Leukemia Type Coded (ICD-10 Codes) on the Death Certificate for Decedents with Leukemia Mentioned on the certificate, with vs without mention of myelodysplastic syndromes (MDS), among patients diagnosed with MDS in 2001-2009

<table>
<thead>
<tr>
<th>Leukemia type on death certificate</th>
<th>Death certificate (using multiple causes)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leukemia without MDS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leukemia with MDS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>Myeloid (C92)</td>
<td>72</td>
<td>59.5</td>
</tr>
<tr>
<td>Acute myeloid (C92.0, C92.5)</td>
<td>(67)</td>
<td>(55.4)</td>
</tr>
<tr>
<td>Acute leukemia unspecified type (C95.0)</td>
<td>24</td>
<td>19.8</td>
</tr>
<tr>
<td>Other and unspecified*</td>
<td>23</td>
<td>19.0</td>
</tr>
<tr>
<td>Total</td>
<td>121</td>
<td>100</td>
</tr>
</tbody>
</table>

*All but 2 were unspecified leukemia (C95.9). No deaths were coded to lymphoid leukemia (ICD-10 code C91).
df: Degrees of freedom in chi-square test.
the UC for 18 of the 160 deaths without mention of MDS on the death record. MPN and was mentioned as other than the UC for an additional 5 deaths, including 1 with essential thrombocytopenia (ET) (ICD-10 D47.3) which had leukemia as UC. MPN, including ET, can transform (progress) to MDS.20 Aplastic anemia (ICD-10 D 61.9) or unspecified anemia (D64.9) was mentioned on the death certificate for 9 of these 23 deaths with MPN, including 7 with MPN as UC and 2 with MPN mentioned elsewhere (data not tabulated).

Small numbers of additional deaths possibly related to MDS (among the 367 without mention of MDS on the death record) included 21 coded with UC as infectious diseases (2.7% of 773), 9 as anemia (1.2% of 773) as UC (Table 1), and 14 as pneumonia (which is coded under respiratory diseases in ICD-10) (data not tabulated). Among the other deaths with known UC (n=209, Table 1) the most common were cardiovascular diseases (n=61), cancers (n=58) and respiratory diseases other than influenza-pneumonia (n=23) (data not tabulated).

Additional analyses focused on the large numbers of deaths (n=121) coded with leukemia but without mention of MDS, which could represent misreporting of MDS ("preleukemia") as leukemia by certifiers. Because MDS involves risk of progression (varying from about 5% to >50% among different MDS subgroups) specifically to AML,1,10,21 leukemia subtype was examined using ICD-10 codes on death records. The distribution of subtypes differed little between the 121 leukemia deaths without mention of MDS vs the 70 with mention of MDS, and AML was most common (Table 2). The distribution of codes for MDS subgroups (in the CTR database) also was similar for the 2 leukemia groups but not informative because NOS (M9989) was most common (ie, 60/121 or 49.5% vs 35/70 or 50.0%) (data not tabulated).

Review of CTR records could confirm 31 leukemia diagnoses on death records (ie, 31/191 or 16.2%, including 19/121 or 15.7% without mention of MDS and 12/70 or 17.1% with mention of MDS); 5 were coded (in ICD-O-3) to unspecified leukemia and 26 to myeloid leukemia, not reportable to SEER and not sequenced as multiple primary cancers. Because reporting of multiple primaries for these myeloid neoplasms was not required until diagnoses starting in 2010,29 the ability to confirm leukemia diagnoses coded on death certificates was limited.

**Discussion**

The limitations of using death certificates in assessing cancer mortality are well known, including misclassification or non-specific coding of cancer type as UC (compared to hospital diagnoses),22 although the present study included attempted confirmation of the leukemia diagnosis for a subgroup of deaths coded to leukemia. In addition, patients diagnosed with MDS may be under-ascertained by cancer registries, and MDS patients aged 65+ years at diagnosis who were registered in SEER registries reportedly had better overall survival compared to unregistered MDS patients in Medicare databases, but some non-registered MDS patients were registered only with leukemia (AML) and had poor short-term survival.6 Therefore, the distribution of causes of death of registered MDS patients (Table 1) may not be representative of that for all MDS patients in the population covered.

MDS diagnoses prior to 2001 were not ascertained by the CTR, and the present study included relatively few deaths with mention of MDS among patients diagnosed (while Connecticut residents) in the early calendar years covered. For this reason, and due to other factors such as possible erroneous diagnoses of MDS on death certificates in national databases, the present study would not have included all deaths coded with mention of MDS for Connecticut residents (at the time of death) recorded in a national database with multiple causes of death.23

The distribution of causes of death, however, differed little between patients diagnosed in 2001-2004 vs 2005-2009, and the difference in proportion with unknown cause in the CTR database (Table 1) reflects the delay in obtaining causes of death for recent deaths. Although 74% of all MDS patients had died (Table 1) according to the version of the CTR database used, continued follow-up of the cohort will be needed to assess future changes in the distribution of causes of death including leukemia. The use of drugs approved since 2004 by the Food and Drug Administration (ie, DNA hypomethylators)10,21,24 may delay blastic transformation of MDS to leukemia in some patients29 and could affect the distribution of causes of death (including leukemia). Also, studies are needed in other cancer registries that collect data on MCs of death in geographic areas that may differ in certifier practices.

Within the study limitations, 160 (44.6%) of all 359 deaths with mention of MDS were coded with MDS as other than the UC, and an additional 121 (15.7% of all 773) deaths had leukemia but not MDS on the death record (Table 1). Among these 121 leukemia deaths, however, AML predominated, as for leukemia deaths with mention of MDS (Table 2). Lack of mention MDS on the death record for these leukemia deaths may represent difficulties (among certifying physicians and coroners) related to differential diagnosis (and the changing classification systems) of MDS vs leukemia, and lack of certifier awareness of the previous diagnosis of MDS or of the importance of mentioning MDS (for cancer surveillance purposes).

For the purpose of estimating probable MDS-related deaths in the population, all 121 deaths with leukemia (the majority coded to AML) but without mention of MDS on the death record of MDS patients (Table 2) could be regarded as MDS-related. The probability of leukemia being MDS-related is highest for certain MDS subgroups (with >50% risk of progression to leukemia) regarded as preleukemic states.21

Combining the 359 deaths with any mention of MDS on the death record, and the 121 deaths with leukemia but not MDS mentioned, would result in 480 MDS-related deaths (62.1% of all 773). In routine cancer mortality statistics based on the UC of death, however, the 178 deaths with leukemia coded as UC (Table 1) would be included with leukemia, and not with MDS, although MDS has not been on the list of specific cancer sites/types in certain reports with US cancer mortality statistics.25
An additional 23 deaths (3.0% of 773) had mention of a MPN but not MDS on the death record; 1 of these had leukemia coded as UC (Table 1). Future studies should be able to confirm transformations of MPN to MDS, because of the reportability of such multiple primary cancers to cancer registries for diagnoses since 2010.18 Mention of anemia on death records for 9 of these 23 deaths could suggest misreporting of MDS as MPN, although anemia also commonly occurs among patients with MPNs as well as certain solid tumors. Some MDS subgroups involve thrombocytosis.1 Also, mixed MDS/MPN neoplasms, a separate entity in several classification systems for myeloid neoplasms, can be difficult to diagnose or to differentiate from other myeloid neoplasms.26

Elderly persons diagnosed with MDS have higher risks of developing not only leukemia (AML) but also various other conditions (eg, cardiac events and cardiac deaths, respiratory conditions and infections), compared to the general Medicare population.27 Panel review of medical records would be needed to assess whether MDS, using data such as severity of neutropenia and anemia, in individual patients contributed to death from various other UC such as infections and pneumonia, as well as chronic diseases (such as cardiovascular and respiratory conditions) that may be exacerbated by the presence of MDS.

Some central cancer registries already link their databases with state mortality files for selected calendar years, as part of the death clearance process. These linkages, however, also could be used to assess the accuracy of death-certificate diagnoses of cancer.18 A record-linkage project in 3 states estimated rates of confirmation of cancer diagnoses coded as UC on death records vs cancer site codes in statewide registries, but MDS was not included in the report.28 Registered MDS patients who died from leukemia without any mention of MDS on the death record could be added to the total number of deaths (among residents) coded to MDS using MCs of death, in order to generate better estimates of MDS-related deaths in the population covered by the registry. These estimates could be published and compared with official statistics based solely on UC of death.

Although the death clearance by central registries for recent years would have started by identifying all deaths in the region with MDS coded anywhere on the death record,17 follow-back may not have been conducted if a decedent with MDS already had leukemia but not MDS coded in the registry database. A SEER-Medicare linkage study found that some non-registered elderly MDS patients in Medicare databases were registered only as AML in SEER.8 With the advent of new rules for reporting and sequencing of multiple primary myeloid neoplasms,19 however, death clearance may result in more cases with both MDS and leukemia eventually being registered for patients with progression of MDS to leukemia.

Because of the reporting and coding rules for central cancer registries for multiple primary hemato poetic neoplasms diagnosed through 2009, only a small proportion of leukemia diagnoses on death records for MDS patients could be confirmed in the CTR. The new coding rules starting with diagnoses in 2010,19 however, will enable central cancer registries to routinely confirm leukemia diagnoses on death records (including those without mention of MDS) for MDS patients, and thus provide better estimates of MDS-related deaths in the population. These new hematopoietic coding rules19 also should result in improved completeness of registration of incident leukemia cases by central cancer registries.29

Accuracy of registration of deaths attributed to leukemia, however, is another issue. In the present study the majority of the 178 deaths with leukemia coded as UC of death (Table 1) were coded to AML, and those coded to unspecified type (Table 2) also could have been AML; thus, these leukemia deaths may actually have been largely or entirely due to MDS, which had progressed to AML.

SEER-Medicare linked databases on elderly cancer patients diagnosed since 2001 have included only the UC of death.30 Adding data MCs of death to these databases would be useful for estimating MDS-related deaths in SEER areas. For the entire US population aged 65+ years (about 97% of whom are covered by Medicare), estimating rates of MDS-related deaths would require identifying all MDS diagnoses (regardless of whether or not registered in SEER) in Medicare databases.6 The challenge of estimating the true burden of MDS-related deaths in populations, however, is further complicated by undiagnosed MDS, especially among elderly persons with unexplained anemia that can contribute to death.

In conclusion, surveillance of MDS-related deaths in populations clearly requires using MCs (not just UC) of death, but data from the CTR indicate this will still not capture other MDS-related deaths (especially with leukemia). Similar studies are needed from other central cancer registries, using MCs of death. Studies are also needed on how UC is assigned on death records by certifiers for individual patients who were diagnosed with both leukemia and MDS in different geographic areas. With the advent of new coding rules for multiple primary hematopoietic cancers,19 central cancer registries can contribute to the challenge of improving surveillance of not only MDS incidence rates but also the burden of MDS-related mortality in populations. Improved surveillance is needed to help avoid underestimating the importance of MDS (and other myeloid neoplasms) in prioritizing or planning research on the causes and treatment of specific cancers.

References
null
Original Article

Evaluation of Primary/Preferred Language Data Collection

Linh M. Duong, MPH; Simple D. Singh, MD, MPH; Natasha Buchanan, PhD; Joan L. Phillips, CTR; Ken Gerlach, MPH, CTR

Abstract: A literature review was conducted to identify peer-reviewed articles related to primary/preferred language and interpreter-use data collection practices in hospitals, clinics, and outpatient settings to assess its completeness and quality. In January 2011, Embase (Ovid), MEDLINE (Ovid), PubMed, and Web of Science databases were searched for eligible studies. Primary and secondary inclusion criteria were applied to selected eligible articles. This extensive literature search yielded 768 articles after duplicates were removed. After primary and secondary inclusion criteria were applied, 28 eligible articles remained for data abstraction. All 28 articles in this review reported collecting primary/preferred language data, but only 18% (5/28) collected information on interpreter use. This review revealed that there remains variability in the way that primary/preferred language and interpreter use data are collected; all studies used various methodologies for evaluating and abstracting these data. Likewise, the sources from which the data were abstracted differed.

Key words: data collection, interpreter use, medical records, primary language, preferred language

Introduction

Vulnerable populations—including racial/ethnic minorities, older adults, and those with low income—are at risk for poorer health and adverse health communication outcomes when they have low health literacy.1-3 “the inability to obtain, process, and understand health information to make appropriate decisions.”1 Studies have demonstrated that low literacy and low health literacy are associated with impaired patient-provider communication, patient non-adherence, increased hospitalization, and poorer health.4-7 Similarly, research examining the effects of patient-provider language discordance on the quality of care found that language barriers are associated with less health education, worse interpersonal care, and lower patient satisfaction.8

Access to a translator may facilitate transmission of health education, but having an interpreter present does not serve as a substitute for language concordance between patient and provider.9 To accurately quantify health disparities due to low health literacy level, a standardized measure would be helpful for primary/preferred language data collection practices. The need for a standardized measure becomes increasingly important as the United States grows more linguistically diverse.

According to a 2010 Census Bureau report, the United States is becoming more linguistically diverse.9 The number of people 5 years of age and older who speak a language other than English at home has more than doubled in the last 3 decades, a growth rate that is 4 times greater than that of the overall US population.9 Within this time frame, the number of speakers of non-English languages grew by 140% while the overall US population grew by 34%,9 highlighting the importance for culturally diverse health care in medical facilities and practices.

To address the health-care needs of this growing population in the United States, a better understanding is needed on how information on language spoken or primary/preferred language is being collected and utilized currently. In the past, data collection practices in health care facilities on language and communication have been limited by systems which were incomplete and used incongruously.7-10 However, with the growing use of electronic health records (EHR), there may be the potential to track and maintain information which is currently difficult to collect in a standardized manner. In addition, the use of an electronic health system has the potential to improve the quality and completeness of these data collection methods. Recently, standards have been developed for certification of EHR technology by the US Department of Health and Human Services (DHHS); standardization will promote the systematic collection of health data to inform care.11,12

While much remains to be done in the development of a standardized set of criteria to evaluate language as related to disease outcomes, the implementation of EHR is a step toward building a foundation for primary/preferred language data collection. Our literature review evaluates current practices on primary/preferred language data collection in hospital medical records and includes an assessment of data collection on interpreter use.

*Cancer Surveillance Branch, Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA. 1Epidemiology and Applied Research Branch, Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA.

Address correspondence to Linh M. Duong, MPH, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Mail Stop F-69, Atlanta, GA 30341. Telephone: (770) 488-3122. Fax: (770) 488-4759. Email: lduong@cdc.gov.

These findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.
Methods

The objectives of this literature review were to address the following questions: (1) Are primary/preferred language data being collected? (2) If so, how is primary/preferred language information being captured and collected? (3) What sources are collecting primary/preferred language data? (4) Is interpreter-use data being collected? (5) If so, how is interpreter-use information being captured and collected? (6) Where is interpreter-use data being collected?

This report defines completeness and interpreter use in the following manner: Completeness of primary language data is defined as data in which the majority (greater than 80%) of the language data is provided (ie, is not missing or unknown). Interpreter use data is defined broadly in this report as any data which records use of interpreter services, data which reports patient lack of English-proficiency requiring the aid of an interpreter to understand the physician, and if data was recorded on language-appropriate written material being provided to the patient.

Data Sources

A literature review was conducted to identify peer-reviewed articles related to primary/preferred language collection practices in hospitals, clinics, and outpatient settings as well as an assessment of data collection on interpreter use. Initially, our literature review (primary search) focused on studies related to cancer/neoplasms and primary/preferred language data collection practices. To increase our sample size, the search was subsequently expanded (secondary search) to include studies of all other diseases and primary/preferred language data collection practices. The primary search was conducted on January 6, 2011 using Embase (Ovid) and MEDLINE (Ovid) for the years 1988-2010. The secondary search was conducted on January 27, 2011 using PubMed and Web of Science to find eligible studies and restricted the search to articles published in the last 5 years. Search terms used in both searches can be found in Table 1. We excluded the following key words: programming language, ontology, language publication, language restriction, language articles, and natural language processing as these terms were not relevant to identifying spoken or vernacular languages of study subjects and as a result did not meet inclusion criteria. The results of both searches were then combined into a single library in EndNote and checked for duplicates, which were subsequently removed.

Study Selection and Data Abstraction

We reviewed titles, abstracts, and full texts of the identified citations and selected eligible articles based on prespecified criteria described below. The selection of eligible articles for data abstraction was based on a 2-phased approach. In the preliminary phase, one coauthor reviewed eligible articles that included primary/preferred language as an outcome variable of interest in either the title or abstract, and determined whether the quality of the data was assessed. Full texts of eligible articles were obtained for the second phase review. For the second phase, studies were selected if they included primary/preferred language data collection in medical records or used other data sources (survey, interview, US Census data, Medicaid data, health-plan data, etc) to obtain primary/preferred language information. Initially, 2 of the coauthors conducted the secondary phase independently. After each reviewer compiled a list of eligible articles, the 2 lists were evaluated by all authors. Then, a final list of eligible articles was compiled for data abstraction and determination of topic area for the tables. The following information was collected from each eligible article: first author name and publication year, methodology used (data source and primary/preferred language variable), and key findings of the study.

Results

This literature search (Figure 1) yielded 768 articles after duplicates were removed. After primary and secondary inclusion criteria were applied, there were 28 eligible articles which remained for data abstraction. These 28 articles were then divided between Table 2 and Table 3 according to data sources from which the primary/preferred language information was obtained. All 28 articles in this literature review reported collecting primary/preferred language data.

Table 2 includes 10 articles that are related to primary/preferred language data collected from medical records. Availability of information on primary/preferred language was quite high. Approximately 60% (6/10) of studies reported information on primary/preferred language.13-18 Furthermore, among these studies, data for primary/preferred language had a high level of completeness ranging from 82% to 96%.13,18 McClure et al reported that overall information on primary/preferred language was available in medical records for 86.4% of study participants.13 Of 27 facilities that did not have primary/preferred language data available to abstract electronically, 81.6% had that data in the medical records, most often in the admission records.13 Polednak found that about 8.4% (64/765) of cases had an unknown preferred language.14 Similarly, in another study
Table 1. Databases and Search Terms Used for Literature Review†

<table>
<thead>
<tr>
<th>Database</th>
<th>Search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary search—cancer/neoplasm only</strong></td>
<td>Language.hw,kw,sh,ti. AND (registry or registries).hw,kw,sh,ti. AND Cancer.hw,kw,sh,ti. or Neoplasm*.hw,kw,sh,ti. hw = heading word kw = key word sh = subject headings ti = title</td>
</tr>
<tr>
<td><strong>Secondary search—no focus on a particular disease</strong></td>
<td>Registry or registries[MeSH] or “Electronic Health Records” or “electronic medical record” or “hospital record” or “reporting hospital” or “reporting hospitals” AND “Language” or “language data” or “primary language” or “preferred language” or “native tongue” or “native language” or “language spoken” or “language proficiency” or “preferred language” or “primary spoken language” or “language proficiency” or “language proficient” or “Spanish speaking” or “native speaker” or “non-English” or “non-native speaker” or “language codes” or “linguistically” or “language barrier” or “languages spoken” or “language concordance” or “language concordant” or “translation services” or “collection of language” or “language data” or “linguistic” or “patient language” or “language of patient” or “multilingualism”[MeSH] or bilingual or bilingualism[MeSH] or multilingualism</td>
</tr>
</tbody>
</table>

†Excludes the following search terms: programming languages, ontology, language of publication, language restriction, language articles, and natural language processing as these terms were not relevant to identifying spoken or vernacular languages of study subjects and as a result did not meet inclusion criteria.

‡MeSH; Medical Subject Headings used by National Library of Medicine for indexing, cataloging, and searching for biomedical and health-related information and documents.

§Quotation marks surround words that were searched as a phrase.

||** = wildcard truncation for plurals.
<table>
<thead>
<tr>
<th>First author and publication year</th>
<th>Methodology</th>
<th>Key findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gindi et al, 2010&lt;sup&gt;19&lt;/sup&gt;</td>
<td>Data Source: EMR abstraction from Baltimore City public STD clinic. Primary/preferred language variable: Language spoken (English-speaking or Spanish-speaking) Language status (Latino English-Proficient, Latino Spanish-Speaking, or Non-Latino)</td>
<td>2% (39,728) patients were Latinos. More than half of Latino patients were Spanish Speaking (60%). This differed by gender and age group.</td>
</tr>
<tr>
<td>McClure et al, 2010&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Data Source: MR from Greater Bay Area Cancer Registry and California Cancer Registry. Primary/preferred language variable: English or non-English</td>
<td>Overall, information on spoken language was available in MR for 86.4% of study participants. For 27 facilities for which language data was not abstracted electronically: 81.6% had language information in MR, most often in admission records. Information on interpreter use was collected in consent forms and nurses’ notes. Significant differences by race, year of diagnosis, and advanced stage were found.</td>
</tr>
<tr>
<td>Parker et al, 2010&lt;sup&gt;19&lt;/sup&gt;</td>
<td>Data Source: EMR Minnesota multispecialty care delivery organization. Primary/preferred language variable: Patient’s Preferred Language</td>
<td>Language interpreter data exist for 93% of this sample, and country of origin data exist for 53%. Total number of patients is not provided for language interpreter data or country of origin data.</td>
</tr>
<tr>
<td>Polednak, 2007&lt;sup&gt;a14&lt;/sup&gt;</td>
<td>Data Source: MR from Connecticut population-based registry. Primary/preferred language variable: Preferred Language (English, Other, Unknown)</td>
<td>Prevalence of comorbid diabetes was 25.1% (192/765). About 15.1% of 166 preferred English vs 30.3% of 535 who preferred a non-English language (predominantly Spanish). About 8.4% (64/765) cases had an unknown preferred language.</td>
</tr>
<tr>
<td>Polednak, 2008&lt;sup&gt;18&lt;/sup&gt;</td>
<td>Data Source: MR from Connecticut population-based registry. Primary/preferred language variable: Preferred/primary language (English, Spanish, Bilingual, Asian, Inconsistent, and Other, Unknown)</td>
<td>Recent weight loss was mentioned for only 21.5% and was less frequent (12.7%) among 237 preferring English vs 28.2% of 418 preferring Spanish and 28.6% preferring an Asian language; the association with language category persisted when other variables were considered. No indication of completeness on language data provided.</td>
</tr>
<tr>
<td>Polednak, 2009&lt;sup&gt;a15&lt;/sup&gt;</td>
<td>Data Source: MR from Connecticut population-based registry. Primary/preferred language variable: English or Other (Spanish, bilingual, inconsistent)</td>
<td>Only 9.7% of 992 Hispanic or Asian patients had no information on language preference. Information on use of an interpreter was missing for 36.1% of 653 probable non-English-prefering patients. Missing information was more frequent in Asian than Hispanic American patients.</td>
</tr>
<tr>
<td>Polednak, 2009&lt;sup&gt;b16&lt;/sup&gt;</td>
<td>Data Source: MR from Connecticut population-based registry. Primary/preferred language variable: Primary language (Spanish/bilingual vs English)</td>
<td>Prevalence of a comorbid mental disorder declined with age but did not differ by primary language (Spanish/bilingual vs English). Primary/preferred language was not recorded in 8.4% of records abstracted. Total number of patients is not provided for primary language.</td>
</tr>
</tbody>
</table>
he noted that records for all but 9.7% of 992 Hispanics or Asian patients had information on language preference and that this information was missing more often in Asian than in Hispanic-American patient records. In a third study by Polednak found that primary/preferred language was not recorded in 8.4% of records abstracted. Solberg et al reported that primary/preferred language data was missing for 6,972 (3.7%) of cases in their study sample. In another study, Solberg et al found that primary/preferred language data was 96% complete. In regards to the collection of interpreter use information, as presented in Table 2, 67% (2/3) reported data completeness information (64% and 93%; respectively) and 33% (1/3) reported where interpreter use data were collected (consent forms and nurses notes).

Table 3 includes 18 articles that used other data sources (eg, surveys, interviews, US Census data, Medicaid data, health-plan data) to obtain primary/preferred language information. Only 6% (1/18) provided data on the availability of primary/preferred language and this study did not specify a percentage missing; rather, researchers just stated that preferred language was usually recorded. There were 11% (2/18) of studies that collected interpreter use information; however, neither study reported the completeness of data on the use of an interpreter.

Overall, 25% (7/28) of studies reported completeness of information on primary/preferred language while 18% (5/28) of studies reported the completeness of data on the use of an interpreter (Table 2, Table 3).

A number of studies shown in Table 3 combined primary/preferred language data obtained from surveys with data from medical records, or in health plan data was not limited to a single disease (Table 2, Table 3). In fact, it appears that physicians from a range of specialties recorded information on primary/preferred language including physicians who treat cancer, mental health disorders, sexually transmitted diseases, diabetes, nutrition, and tobacco use. Thus, primary/preferred language affects many diverse specialties. This is further supported by a study by Hasnain-Wynia et al in which 20 practices nationwide, each with 5 or fewer physicians, were interviewed, found that primary/preferred language data was collected across several disciplines. These practices

---

**Table 2, cont. Studies Related to Language Data Collection in Medical Records**

<table>
<thead>
<tr>
<th>First author and publication year</th>
<th>Methodology</th>
<th>Key findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solberg et al, 2008[17]</td>
<td>Data Source: EMR from Minnesota HealthPartners Medical Group (HPMG) multi-specialty care delivery organization. Primary/preferred language variable: Preferred language (English, Other, or No Data)</td>
<td>Overall, 19.7% with recorded status were tobacco users as were 8.5% of those whose preferred language was other than English. Language data was missing for 6,972 (3.7%). Underreporting of tobacco status appears to be correlated with the absence of other data such as insurance information (84.4%), ethnicity (80%), and preferred language (73.6%).</td>
</tr>
<tr>
<td>Solberg et al, 2010[18]</td>
<td>Data Source: EMR from Minnesota HealthPartners Medical Group (HPMG) multi-specialty care delivery organization. Primary/preferred language variable: Language preference (English, Spanish, or Other)</td>
<td>Groups receiving fewer [tobacco cessation prescription] orders than their comparison groups included those with non-English preference. The same groups were less likely to fill that prescription, except patients with non-English preference or Medicaid. Language data was 96% complete. Total number of patients is not provided for primary language.</td>
</tr>
<tr>
<td>Toccher et al, 1998[40]</td>
<td>Data Source: EMR from clinical and administrative databases at the University of Washington Medical Center and Harborview Medical Center. Primary/preferred language variable: Language type (English-speaking or Non-English-speaking)</td>
<td>At these institutions, the quality of diabetes care for non-English speaking patients appears as good as, if not better than, for English-speaking patients. Physicians may be achieving these results through more frequent visits and laboratory testing. No indication of completeness on primary language.</td>
</tr>
</tbody>
</table>

EMR=electronic medical record. MR=medical records. STD=sexually transmitted disease.
<table>
<thead>
<tr>
<th>First author and publication year</th>
<th>Methodology</th>
<th>Key findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gomez et al, 2004&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Data Source: MR from Greater Bay Area Cancer Registry. Language data obtained from interview. Primary/preferred language variable: Preferred Language (English, Half English/Half Not English, Not English)</td>
<td>Among US-born Asians, those misclassified as foreign-born were more likely than those correctly classified to prefer a non-English primary language. Asian subgroups varied by preferred language. The multiple-race Asian group was most likely to prefer to use English (79%), followed by Japanese (86%), other Asian (61%), and Filipinos (56%), while the majority of Vietnamese (62%) preferred not to use English.</td>
</tr>
<tr>
<td>Gomez et al, 2005&lt;sup&gt;10&lt;/sup&gt;</td>
<td>Data Source: Greater Bay Area Cancer Registry. Language data obtained from interview. Primary/preferred language variable: Preferred language (English, Half English/Half Not English, Not English, or Not Asked/Refused)</td>
<td>About 40% preferred to use English as a primary language, and 30% preferred another language. Patients who preferred speaking a language other than English were half as likely to have unrecorded birthplace, although the magnitude of this association was diminished somewhat in the adjusted model.</td>
</tr>
<tr>
<td>Hamilton et al, 2009&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Data Source: Los Angeles Cancer Surveillance Program. Language data obtained from survey. Primary/preferred language variable: Short Acculturation Scale for Hispanics 5-point scale) – Only English, English better than Spanish, both equally, Spanish better than English, Only Spanish). Respondents answered the following questions using 5-point scale: (1) What language(s) do you read and speak? (2) What language(s) do you usually speak at home? (3) In what language do you usually think? (4) What language do you usually speak with your friends?</td>
<td>Greater than 50% of the self-identified Latinas indicated that they preferred to speak Spanish over English. The Short Acculturation Scale for Hispanics results suggests that those strongly preferring Spanish reported the lowest levels of education, being born in the United States, and having either parent born in the United States.</td>
</tr>
<tr>
<td>Hasnain-Wynia et al, 2010&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Data Source: Data collected from 20 practices nationwide and were from medical practices with 5 or fewer physicians. Language data obtained from interview. Primary/preferred language variable: Preferred or primary language Use of interpreter</td>
<td>Of the 20 practices surveyed, 9 reported collecting either race, ethnicity, or primary language; 3 collected race/ethnicity and primary language data; 5, only race/ethnicity; and 1, only primary language. Only 1 practice feature facilitated demographic data collection: use of EMR system (7 of 10 practices with an EMR collected data). When patient information on language is collected, it is rarely used to schedule interpreters or to guide the translation of patient materials, even when these services are offered by the practice.</td>
</tr>
<tr>
<td>First author and publication year</td>
<td>Methodology</td>
<td>Key findings</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Hawley et al, 2008&lt;sup&gt;24&lt;/sup&gt;</td>
<td>Data Source: MR from Los Angeles metropolitan SEER registries data. Language data obtained from survey and merged to SEER data. Primary/preferred language variable: Race/ethnicity (Latina-Spanish speaking, Latina-English speaking, African-American, Caucasian) Health literacy (low, moderate, high) Translation (did not need, family or friend, doctor or staff)</td>
<td>The analytic sample included 877 women: 24.5% Latina-Spanish speaking (Latina-SP), 20.5% Latina-English speaking, 24% African-American and 26.6% Caucasian. Approximately 28% of women in each ethnic group reported a surgeon-based, 36% a shared, and 36% a patient-based surgery decision. Spanish preferent Latina women had the greatest odds of high decision dissatisfaction and regret controlling for other factors. Low health literacy was independently associated with dissatisfaction and regret and slightly attenuated associations between Latina-SP ethnicity and decision outcomes.</td>
</tr>
<tr>
<td>John et al, 2005&lt;sup&gt;32&lt;/sup&gt;</td>
<td>Data Source: Greater Bay Area Cancer Registry. Language data obtained from interview. Primary/preferred language variable: Acculturation index based on language usage and generational status</td>
<td>Among long-term foreign-born residents, breast cancer risk was lower among Hispanics who moved to the United States at age ≥20 years and those who spoke mostly Spanish.</td>
</tr>
<tr>
<td>Johnson-Kozlow, 2010&lt;sup&gt;32&lt;/sup&gt;</td>
<td>Data Source: 2005 California Health Interview Survey (CHIS). Language data obtained from survey. Primary/preferred language variable: Acculturation was score composed of seven status variables that measure English language use and proficiency, nativity and citizenship, and years lived in the US</td>
<td>Approximately 18% said that only English was spoken at home; 5% said they had difficulty understanding their doctor at their last doctor visit. Of those 82% said they had difficulty understanding the doctor due to language and 66% said they needed another person to help them understand the doctor.</td>
</tr>
<tr>
<td>Kaplan et al, 2011&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Data Source: Eight California Cancer Registry regions and linked to survey data about patient treatment decision making. Language data obtained from survey. Primary/preferred language variable: Ethnicity language group (White women, English-speaking Latinas, or Spanish-speaking Latinas)</td>
<td>English-speaking Latinas (ESL) were more likely to receive radiation than their Spanish-speaking or white counterparts, controlling for demographic and other factors. A greater proportion of white women had a college education compared to ESL and Spanish-speaking Latinas (SSL) women. The majority of white and ESL women were privately insured, but this was not true for SSL women. A larger proportion of white and ESL women reported having a relative with a history of breast cancer compared with SSL women.</td>
</tr>
<tr>
<td>Karter et al, 2000&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Data Source: Kaiser Permanente Northern California Region health survey. Language data obtained from survey. Primary/preferred language variable: Language measure (prefer to communicate in non-English language) English language difficulty (Yes or No?)</td>
<td>Among Hispanics and Asian/Pacific Islanders, 26 and 30%, respectively, were identified as having difficulties communicating in English or as preferring languages other than English. However, only 1% of non-Hispanic Caucasian and African-American members with diabetes had language difficulties. In most cases, those patients with language difficulties were less likely to practice self-monitoring of blood glucose (SMBG) at recommended levels compared with subjects who were fluent in English.</td>
</tr>
<tr>
<td>First author and publication year</td>
<td>Methodology</td>
<td>Key findings</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Kouri et al, 2010&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Data Source: SEER population-based data. Language data obtained from US Census. Primary/preferred language variable: Race, Ethnicity, Birthplace</td>
<td>Foreign-born Hispanic women in the United States have a lower probability of being diagnosed at earlier stages of breast cancer and, for women with early-stage disease, of receiving radiation following breast conserving surgery compared to US-born Hispanics and whites. Adjusted rates of stage at breast cancer diagnosis included an adjustment for Spanish language proficiency. Adjusted rates of breast-conserving surgery (BCS) without radiation, BCS with radiation and mastectomy included an adjustment for Spanish language proficiency.</td>
</tr>
<tr>
<td>Napoles-Springer et al, 2007&lt;sup&gt;27&lt;/sup&gt;</td>
<td>Data Source: Population-based SEER registry. Language data obtained from telephone survey. Primary/preferred language variable: Ethnicity Language of Interview (English or Spanish)</td>
<td>Results suggest that families play an important role in promoting use of support groups among Latina breast cancer survivors, and that spirituality may offer an alternative source of support. More effort should be directed toward providing culturally and linguistically appropriate support services to breast cancer survivors, and increasing awareness of these services among oncologists, patients, and family members.</td>
</tr>
<tr>
<td>Polednak, 2005&lt;sup&gt;20&lt;/sup&gt;</td>
<td>Data Source: 30 acute care hospitals were surveyed. Language data obtained from survey. Primary/preferred language variable: Race, Ethnicity, Birthplace</td>
<td>At least one staff member at 86% of 28 responding hospitals reported a hospital policy to ask patients about their race, vs 25% for ethnicity and 57% for birthplace, and patient self-reports were reportedly used to obtain race in 100% of hospitals vs 54% for ethnicity. Ethnicity was rarely recorded on any specific type of document, although preferred language was usually recorded.</td>
</tr>
<tr>
<td>Polednak, 2007b&lt;sup&gt;38&lt;/sup&gt;</td>
<td>Data Source: MR from Connecticut population-based registry. Language data obtained from Physician Profile Survey (PPS). Primary/preferred language variable: Data on primary/preferred language of each Hispanic patient was not available in this study.</td>
<td>Having a Follow-up physician (FUP) with a Spanish-language practice (SLP) was statistically significantly associated with receipt of radiotherapy for breast cancer but not for prostate cancer. This methodology should be explored in states with larger Hispanic populations, and future efforts should include efforts to obtain data on other cancer treatments (eg, chemotherapy and hormone therapy).</td>
</tr>
<tr>
<td>Ramsey et al, 2010&lt;sup&gt;46&lt;/sup&gt;</td>
<td>Data Source: Washington State Cancer Registry (WSCR) and Medicaid enrollment and claims records. Language data obtained from Medicaid (per Scott Ramsey via email on 08/31/2011). Primary/preferred language variable: Primary language (English or Other)</td>
<td>Factors associated with not receiving radiation included in situ disease and non-English as a primary language.</td>
</tr>
</tbody>
</table>
represented a diverse set of specialties including internal medicine, obstetrics and gynecology, family medicine, pediatrics, geriatrics, and pulmonary medicine. Results of this study found that of the 20 practices interviewed, 9 reported collecting demographic data (eg, race, ethnicity, and/or language). Some studies had a primary language.21

Table 3, cont. Use of Other Data Sources (Survey, Interview, US Census, Medicaid, Health Plan) to Get Language Info

<table>
<thead>
<tr>
<th>First author and publication year</th>
<th>Methodology</th>
<th>Key findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith et al, 201013</td>
<td>Data Source: Brain Attack Surveillance in Corpus Christi (BASIC) – stroke surveillance study. Language data obtained from interview. Primary/preferred language variable: Language (self-reported language fluency and dichotomized as “Spanish” or “English”; English speakers included subjects fluent in both languages)</td>
<td>Mexican Americans were less likely than non-Hispanic whites to arrive by emergency medical services (odds ratio, 0.6; 95% CI, 0.4, 0.8). Men were more likely than women to present to the hospital within 3 hours (odds ratio, 0.7; 95% CI, 0.5, 0.9); language was not associated with study outcomes.</td>
</tr>
<tr>
<td>Sweeney et al, 200714</td>
<td>Data Source: Utah Cancer Registry and New Mexico Tumor Registry. Language data obtained from interview. Primary/preferred language variable: Surname, Ethnicity</td>
<td>Hispanics who were correctly classified differed from those who were misclassified, reporting lower language acculturation and education attainment. The authors conclude that a surname search efficiently identifies Hispanics, although individuals identified using this method are not completely representative. Recruitment of Hispanic cases and controls does not appear to be affected by selection bias related to community characteristics.</td>
</tr>
<tr>
<td>Traylor et al, 201017</td>
<td>Data Source: Kaiser Permanente’s Northern California Diabetes Registry of 2005. Language data obtained from health plan, at plan level separate from the registry, and available through automated clinical data (per co-author Julie Schmittiel via email on 09/07/2011 and 09/09/2011). Primary/preferred language variable: Race/Ethnicity/Patient language</td>
<td>Patients who chose their physicians were more likely to have a same race/ethnicity physician with OR of 2.2 (95% CI 1.74-2.82) for African American patients, 1.71 (95% CI 1.44-2.04) for Hispanic patients, 1.11 (95% CI 1.04-1.18) for white patients, and 1.38 (95% CI 1.23, 1.55) for Asian patients. Limited English language was a strong predictor of concordance for Hispanic patients (OR 4.81; 95% CI 4.2-5.51) and Asian patients (OR 9.8; 95% CI 7.7, 12.6)</td>
</tr>
<tr>
<td>Yoon et al, 200824</td>
<td>Data Source: MR from Los Angeles County SEER Registry. Language data obtained from survey. Primary/preferred language variable: Race/Ethnicity (Black, Hispanic English speaker, Hispanic Spanish speaker, Other, White)</td>
<td>Multi-variate analysis controlling for patient characteristics and treatment showed that older, black, Hispanic Spanish-speaking, widowed or never married, and working women were less likely to report severe symptoms than other women.</td>
</tr>
</tbody>
</table>

Primary/preferred language variables differed for all studies (Table 2, Table 3). Some studies had a single primary/preferred language variable. Other studies had 2 or more primary/preferred language variables. And, others still developed their own intrinsic scale used to measure primary/preferred language or used a scale already developed. Examples of some of the single primary/preferred language variables used include the following variable categories: English or non-English, patient’s preferred language, preferred language (English, Other, and Unknown), preferred/primary language (English, Spanish, Bilingual, Asian, Inconsistent, and Other, Unknown), English or Other (Spanish, bilingual, inconsistent), primary language (Spanish/bilingual vs English), preferred language (English, Other, or No Data), language type (English-speaking or Non-English-speaking), and preferred language (English, Half-English/Half Not English, Not English). Likewise, studies with 2 or more primary/preferred language variables also had wide-ranging categories. Examples include a study by Gindi et al which had 2 primary/preferred language variables: language spoken (English-speaking or Spanish-speaking) and language status (Latino English-Proficient, Latino Spanish-Speaking, or Non-Latino) and a study by Hawley et al which had the following primary/preferred language variables: race/ethnicity (Latina-Spanish speaking, Latina-English speaking, African American, Caucasian) and health literacy (low, moderate, high). Studies that developed their
own primary/preferred language variable include a study by John et al that created an acculturation index based on language usage and generational status and a study by Johnson-Kozlow which developed an acculturation scale with scores that were composed of 7 status variables that measured English language use and proficiency, nativity, citizenship, and years living in the United States. Lastly, Hamilton et al used a scale already developed called the Short Acculturation Scale for Hispanics to assess primary/preferred language.

Discussion

Our literature review indicates that although the completeness of primary/preferred language data collection is high, 96% completeness for primary/preferred language data collection in 1 study, there remains variability in the way this information is collected. For example, investigators used different protocols for evaluating the collection of primary/preferred language and the sources used to collect this information also varied. In addition to using hospital medical records to obtain primary/preferred language information, investigators used surveys, interviews, US Census data, Medicaid data, or health plan data. This information was then combined with the medical records to assess disparities in health outcomes based on primary/preferred language concordance. These findings show that primary/preferred language data collection occurs in multiple ways within various settings. Moreover, the collection of data from various sources reduces the ability to make comparisons across studies as well as limits the possibility of aggregating primary/preferred language data study results to obtain a global review of health outcomes. Similarly, the lack of a common definition and standard codes may impede research efforts. Therefore, a standardization of primary/preferred language collection practices may be warranted.

Furthermore, the collection of interpreter service data differed between studies. Of the studies that reported collecting primary/preferred language data, only 18% (5/28) collected information on interpreter use. While a study by McClure et al that stated interpreter use information was found in consent forms and nurses’ notes while Polednak reported that information on interpreter use was missing for 36.1% of 653 probable non-English-prefering patients. And, even when interpreter use data are collected, it is unclear how this information is used to improve services for these patients.

Our findings regarding the variability in the collection of primary/preferred language data is similar to what was found in a 2007 Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) report. In 2006, JCAHO required the maintenance of records on patients “language and communication needs.” These standards are intended to support the provision of care, treatment, and services in a manner that is conducive to cultural, language, literacy, and learning needs of individuals. For example, these provisions include standards for respecting values and beliefs of the patient, appropriate communication, including interpreter and translation services, effective communication throughout organization, ensuring that orientation and ongoing staff education is appropriate to the needs of patient population, and the collection of data, documentation of needs and access to data. However, a review in 2007 of these records reveals that there still remains much work in improving this system as noted by an evaluation of 60 representative US hospitals. In this review, JCAHO found that systems for collection of required data on language and communication were “underdeveloped” and used “inconsistently.” However, this may change as a result of implementation of the certification criteria for EHR technology.

In 2010, the DHHS issued a final rule to complete the adoption of an initial set of standards, implementation specifications, and certification criteria for EHR technology. Stage 1 criteria for EHR certification states the minimum elements required in support of meaningful use by eligible professionals, eligible hospitals, and critical access hospitals under the Medicare and Medicaid EHR Incentive Programs. Specifically, the collection of demographic data includes a record of preferred language and demonstration of meaningful use of this technology. However, a data collection field in the EHR does not necessarily indicate that the data will be collected by physicians. Providing an appropriate incentive may be needed to assist physicians in collecting these data.

In addition, a report by the Institute of Medicine (IOM) stated that data on a person’s language and communication needs should be a part of any minimum data set related to health care delivery and quality improvement. The IOM subcommittee for this report recommended identifying spoken language need in a stepwise approach: first by determining how well the individual believes he/she speaks English and second by asking what language he/she needs for a health-related encounter. This will allow for improved quality of services in subsequent encounters, in analysis of health disparities, and in system-level planning (determining the needs for interpreters and matching patients to language-concordant providers). A study by Karliner et al, which adds support to this recommendation, found that a screening question asking how well a patient speaks English followed by language preference for medical care was most inclusive and accurate for identifying patients likely to benefit from language assistance. Certification of EHR did not include recommendations from this IOM report.

Policies and initiatives to strengthen health literacy across the nation have also begun to take root. For example, Healthy People 2020 has begun incorporating objectives to improve health literacy and provider communication. Specifically, these 2 objectives seek to “improve health literacy of the population (HC/HIT-1)” and to “increase proportion of persons who report that their health care providers have satisfactory communication skills (HC/HIT-2).” The overall goals are to improve health outcomes, health care quality, and achieve health equity through the use of health communication strategies and health information technology.

Similarly, recent federal policy initiatives have begun...
to bring health literacy to the forefront of the health care discussion, including the Affordable Care Act of 2010, the DHHS’ National Action Plan to Improve Health Literacy, and the Plain Writing Act of 2010. The Affordable Care Act addresses health literacy by integrating training on health literacy for health professionals (section 5301) and requiring that health plans and insurers provide consumers with a summary of health information, benefits, and coverage options that is clear and consistent and that can be compared to other plans (section 2715). The National Action Plan to Improve Health Literacy provides a consolidated structure with which to unite health literacy goals and strategies for the nation. And, the Plain Writing Act of 2010 specifies that federal agency documents must be written clearly so the public will be able to understand them.

As uniform data collection becomes the norm, the strengthening of health information technology across the nation may have direct implications for medical records. Medical records at the hospital level may benefit from the development of standardized protocols for primary/preferred language data collection practices. These protocols may improve the ability of hospitals to help address patients’ linguistic needs and support studies related to reducing health disparities. The standardization of primary/preferred language data would allow an accurate assessment of differences across hospitals, clinics, and outpatient settings.

The ultimate goal of primary/preferred language concordance between patient and provider is providing access to health services that are required for effective treatment, especially for patients with complex illnesses such as cancer. Such patients commonly require access to multiple specialists, effective coordination of care, accurate information about disease and treatment options, and timely attention to symptoms.

The need for high-quality data that are complete and accurate is not unique to primary/preferred language data collection. In fact, obtaining data on variables such as race/ethnicity, socioeconomic status, and stage at diagnosis also have been difficult. As a result, in addition to the standardization of data collection practices, continuous quality-control activities are also needed to identify and correct errors and to ensure uniformity and accuracy of the data collected.

Our findings should be considered in light of several limitations. Although we employed a thorough and extensive search strategy and literature review, some studies may not have been identified and included in this review. In particular, since we focused on peer-reviewed publications, we did not examine unpublished documents or reports on this topic. In addition, due to variability in the manner that the primary/preferred language data were collected across the studies included in this review, we were not able to aggregate studies for meta-analysis. In spite of these limitations, this literature review is among the first assessments, to our knowledge, to examine primary/preferred language and interpreter use data collection practices in hospital medical records, to explore the completeness of these data, and to identify areas in need of improvement.

Conclusions

As the United States moves toward improving the health literacy of its population by strengthening provider communication through health information technology and by passing federal initiatives and policies to support these goals, a more uniform protocol may emerge for collecting information on primary/preferred language. This is especially important in light of the Healthy People 2020 objectives to improve population outcomes related to health-care quality and health equity through the use of health information technology. The development of a standardized protocol to collect data on primary/preferred language may improve research methods used to analyze health disparities related to language spoken and utilization of interpreter services which has the potential to impact disease outcomes. The ability to describe areas in which resources are lacking for vulnerable populations unable to access health care due to patient-provider discordance in language may aid in creating public health interventions targeted at improving and increasing needed resources at facilities which serve a diverse population.

Acknowledgements

The authors would like to acknowledge Onnalee Gomez and Katherine Tucker, CDC Library personnel, who conducted the literature review search and consolidation of articles for this paper. In addition, we would like to thank Qiang Ling, Carissa Holmes, Mridhula (Maya) Kumar, Megan Crawley, and Kate Allen for their comments and suggestions in the development of this paper.

References


Comparison of the NCRA and NAACCR Strategic Management Plans

by Herman R. Menck, BS, MBA, CPhil, FACE

Abstract: The Strategic Management Plans of the National Cancer Registrars Association (NCRA) and the North American Association of Central Cancer Registries (NAACCR) were compared, and differences noted. No uncovered subject areas were found.

Key words: cancer registry associations, strategic management plan, goals, plan development, education, recruitment, policy

Introduction

Both of the premier cancer registry organizations, the National Cancer Registrars Association (NCRA) and the North American Association of Central Cancer Registries (NAACCR), have developed strategic management plans (SMPs) to more closely focus their efforts and maximize their results.¹ ² NCRA’s plan was first completed in 2003, and the NAACCR’s more recently in 2011.

It is understood that there is significant overlap between the missions of NCRA and NAACCR, but there are also differences in furtherance of cancer registration. Interest in better understanding this commonality of mission, the relative organizational priorities, and discovery of any possible uncovered subjects led us to a comparison. This comparison attempts a big-picture macro-comparison, and may not comprehensively compare secondary detailed inclusions.

Methods

Both plans were the result of elaborate collaborative processes involving committee chairs, officers, staff, and other leaders of each organization. Every attempt at synergism was made. These formalizations of common mission understandings are more complex than they might appear. Numerous meetings and copious correspondence were used in completing the final products. The NCRA SMP is 4 pages long and in summary form. The NAACCR SMP is 28 pages long and inclusive of more detail. The NAACCR SMP is challenging to read and remember, but provides more of a roadmap. The level of detail in the NCRA SMP includes objectives within each strategy.

NAACCR’s SMP priority statements include much more detail, such as how, who, and when each priority is to be met. For each priority area, a major goal, measurable objectives, responsible parties, key strategies, expected outcomes, and timelines are set.

In addition, NAACCR’s SMP includes various supporting documentation, including leadership names; vision; mission; organizational values; history and background; strategic management process; market research and environmental scan; management assessment; management survey; interviews; and approvals.

Both organizations employed an outside plan development facilitator to enhance discussion and optimize the final product.

Results

For ease of understanding, I have reorganized and renumbered a combined list of NCRA strategies and NAACCR priorities, I to IX.

The following (I-IV) are a list of strategy/priorities common to both NCRA and NAACCR strategic plans.

I. Education/Professional Development

There are strong similarities between NCRA’s first strategy and the second goal of NAACCR’s fifth priority. NCRA’s first strategy, entitled Education/Professional Development, reads, “Provide comprehensive educational opportunities that are accessible, cost appropriate and forward-thinking” and includes these objectives:

1. Develop a comprehensive education plan to meet the changing demands of the profession and the CTR credentialed individual.
2. Deliver basic and advanced (post-CTR) education opportunities.
3. Expand and enhance formal education opportunities.
4. Monitor and encourage informatics role in the cancer registry profession.

In comparison, the second goal of NAACCR’s fifth priority, entitled Professional Development, reads, “Develop a comprehensive multidisciplinary training program that provides cross-training and leadership skills to ensure that professional personnel in NAACCR member organizations possess the requisite skill sets required to excel in the rapidly changing cancer surveillance environment.”

Objectives include:

1. Continue to provide educational venues for registry personnel.
2. Assist registries in the development of IT skill sets.
3. Create a comprehensive training program.

These strategies are largely similar, and both aim for comprehensiveness. It may be that the inclusion of education/training as a sub-priority, later in the NAACCR SMP, indicates a lower priority.
II. Recruitment and Retention

NCRA’s strategy reads, “Expand the workforce of the cancer registry profession by encouraging new people to enter the field and by improving retention of those currently in the field”, and includes the following objectives:
1. Implement recommendations from NCRA’s Workforce Study.
2. Increase partnerships with peer organizations.
3. Increase and develop the leadership skills and participation of members.
4. Monitor and encourage professionalism and skills.

The second objective above of NCRA’s second strategy (Recruitment and Retention), which seeks to increase partnerships, has been separated out by NAACCR as their first priority area, as further discussed below.

NAACCR’s statement on Recruitment and Retention reads, “Partner with other professional organizations to address recruitment and retention issues and delineate NAACCR’s role in the national retention and recruitment effort,” and includes these objectives:
1. Develop and implement strategies to retain personnel in central registries by enhancing career opportunities for individuals with diverse skill sets such as epidemiology, operations, statistics, and information technology.
2. Support collaborative approaches.
3. Examine opportunities to build career ladders within central cancer registries.
4. Encourage cross-training of personnel.

Here as elsewhere in their strategic plan, NAACCR has listed more specific outcomes and a timeline for this strategy. NCRA does not include outcomes or timelines in their plan.

Again, as in Education/Professional Development, NCRA and NAACCR’s Recruitment and Retention Strategies/Priorities are largely similar.

III. Advocacy/Policy Development

NCRA’s Strategy on Advocacy is similar to that of NAACCR’s Communication and Policy Development.

Advocacy: NCRA’s SMP reads, “Be a strong advocate for our members by actively engaging in processes to network and communicate to affect an opinion”, and includes the following strategic objectives:
1. Expand the advocacy process and outcomes mechanism by creating and implementing a process to develop and evaluate NCRA policy statements.
2. Enable members input regarding standard changes prior to implementation.
3. Position NCRA as part of the decision-making process.
4. Develop programs to raise awareness and improve image perception of the profession.

Policy Development: NAACCR’s Priority on Policy Development reads “NAACCR needs to develop and promulgate carefully crafted policy statements on critical issues of national relevance to its members and the broader public they serve.”

Goal 1: Serve as the voice for NAACCR members on key issues involving central cancer registries.
Objective 1: Use modern methods such as internet-based technologies to capture and share member views, opinions, and perspectives on important registry issues.

Objective 2: Develop position papers and policy statements, as appropriate, that support registries, cancer surveillance, and the NAACCR mission.

Objective 3: Serve as a united voice for policy issues and position statements that promote NAACCR’s mission or benefit central cancer registries.

IV. Member and Customer Services/Communication

Member and Customer Services: NCRA’s strategy reads, “Assure satisfaction of internal and external NCRA customers and excellence in communications” and includes the following strategic objectives:
1. Establish a communication plan.
   a. Increase communication with state and local associations.
   b. Increase communication and involvement with standard setters.
   c. Explore ways to maximize effective communications with NCRA’s other audiences.
2. Continually evaluate innovative opportunities that would improve the quality, access, and delivery of NCRA’s published information and materials to the membership and the field (includes all products, services, research, and development).
3. Increase member satisfaction with NCRA benefits and services.
4. Develop and implement a marketing plan.

Communication: NAACCR’s priority states: “Promote the sharing of expertise, knowledge, procedures, and best practices among NAACCR members to ensure efficiency and reduce redundancy of effort” and includes the following objectives:
1. Develop a resource on the NAACCR website where members may post informational items that may be of value to other NAACCR members.
2. Develop an area on the NAACCR website where members may ask other members for guidance with particular issues or suggest a problem for collaborative solution efforts.
3. Move NAACCR’s use of Web-based and technology-driven communication systems forward to improve information sharing, and promote adoption of best practices.

Next are 2 strategies that are largely unique to NCRA.

V. Credentialing

NCRA’s SMP states “Advance, administer and deliver a continually improving credentialing program to meet the needs of the profession” and includes the following strategic objectives:
1. Hold credentialing process to the highest standards.
2. Evaluate the need for new credentials to support the field.
3. Evaluate the relevance of the CTR nomenclature, and take appropriate action.
4. Promote credentials offered by NCRA.
5. Evaluate, enhance, and deliver the certification maintenance process.
VI. Administrative and Finance

NCRA's SMP states, “Administration and Finance. Maintain financial viability with an effective and efficient infrastructure” and includes the following strategic objectives:
1. Establish a 5-year financial plan and 6-month reserve.
2. Establish a financial growth plan.
3. Establish a governance philosophy to function as a knowledge-based board of directors.
4. Ensure adequate human resources: volunteers, staffing and contract services to support organizational goals.

Next are 3 priorities that are largely unique to NAACCR’s SMP.

VII. Strategic Alliances

NAACCR’s SMP states “Major Goal: Strengthen relationships with SMOs and existing partners, while establishing new alliances with key organizations to promote the NAACCR mission” and includes the following objectives:
1. Cultivate productive working relationships with SMOs by facilitating open communications and purposeful actions.
2. Enhance existing relationships and build new strategic alliances that benefit cancer registries and support the NAACCR mission.

NCRA’s SMP, under Recruitment and Retention, states, “Increase partnerships with peer organizations.” The SMP does not explicitly extend this to other NCRA strategies.

VIII. Standardization and Registry Development

“The core function of NAACCR is to support and strengthen central registry standardization and development. ... NAACCR cannot afford, if it is to remain true to its core function, to conduct business as usual but rather must embrace change.... Prepare for the cancer surveillance system of the future—a system that is more timely and adaptable to change.”

Goal 1: Prepare for the cancer surveillance system of the future—a system that is more timely and adaptable to change.

Objective 1: Explore how cancer surveillance systems will interface with electronic health records and continue to assess semantic interoperability issues.

Objective 2: Stay engaged and remain current with national/international efforts regarding electronic health records and enhance efforts to include cancer in the “meaningful use” case for public health reporting.

Objective 2: Provide consensus standards and best practices for the collection and processing of cancer and patient information.

Objective 1: Ensure the maintenance of Standards volumes and implementation guidelines that are consensus-based, reflect a comprehensive vetting process, and conform to interoperable content and transmission standards.

IX. Research and Data Use

“As research methodologies are changing, the development and use of data will also have to change to make the fullest possible contribution to North America’s future research agenda in epidemiology, cancer prevention and control, and, in particular, clinical outcomes.”

Goal 1: Develop efficient, centralized processes to improve access to North American population-based cancer registry data for cancer linkages, research, and surveillance.

Objective 1: Promote the use of quality, and timely registry data by enhancing the annual Call for Data and the various NAACCR-CINA data products.

Objective 2: Develop a voluntary process to combine limited data from multiple registries to facilitate record linkage research.

Objective 3: Facilitate the development and availability of state and province-specific relative survival data.

Objective 4: Maintain and enhance tools to inform researchers about state and province-specific research experience, interests, and processes to initiate research. Responsible Parties: Data Use and Research Committee.

Objective 5: Increase accessibility to NAACCR’s Cancer in North America (CINA) products by periodically evaluating data access policies and processes.

Objective 6: Strengthen NAACCR’s internal capacity to support research activities.

Discussion/Conclusion

It is clear that there is great commonality of mission for NCRA and NAACCR. Nonetheless, important differences exist. For NCRA, primary responsibility for individual certification (CTR) is featured. For NAACCR, there is emphasis on registry development, standardization, research and data use. NAACCR speaks primarily to central registrars. While most NCRA members are hospital based, recently approximately 20% have reported central registry employers. Other differences exist.

NAACCR is primarily an organization of central registries, while NCRA is an organization of individual registrars. In NAACCR, the lead person from each organization speaks for that organization. In NCRA, each member is considered to have an equal voice.

NCRA and NAACCR have various tools for communication and collaboration, but some workers have found interorganizational collaboration difficult. NAACCR may have an IT committee, and NCRA may have an informatics committee, but each organization’s committee activities remain quite independent.

In this analysis, no relevant subjects not covered by NCRA or NAACCR were found. It is believed that the development of these strategic management plans sets a sound baseline, and favorably prepares cancer registration for the future.

References
Are you living a meaningful life? Many of us have spent years looking for ideas, principles, and strategies to manage our lives. If you are like me, along the way you have found some inspirational nuggets here and there that have helped you get through a difficult day or a rough patch in life. But that’s really not enough. When all is said and done, almost everyone wants to have lived a life of purpose and meaning.

The secret to living a meaningful life may not be as elusive as you think. There are no reference books, standard-setting rules or guidelines, online databases, webinars, books or conferences that you need to get started.

So, what’s the secret? Simply put, it is your attitude, your philosophy about others, the choices you make each day and how you act them out. Let’s look at 4 secrets to living a meaningful life that were inspired by Jim Rohn, an American entrepreneur, author and motivational speaker:

Secret #1: Life is worthwhile if you LEARN. In this case, what you do not know can, and will, hurt you. In order to exist as a cancer registrar, you have to learn. This includes learning from negative and positive experiences; learning from your peers; learning from practice, from formal education and continuing education programs; and learning from as many other resources as you can and are available to you just for the asking. Learning is balanced, meaning you grow your technical skills and nurture your personal and leadership skills. It means you share with others what you have learned with the intent of helping them to grow and be the best they can be. Learning never stops and it is not limited to any type of resource, topic, or venue. To live a meaningful life you need to adopt a non-stop, “learn from everything you possibly can” attitude and practice it every day.

Secret #2: Life is worthwhile if you TRY. You can’t just learn and do nothing; you have to try to do something with it. Try to make a difference. Try to make some progress. Try to learn a new skill. Try to do a task or complete a project better than you did before. Try to implement a new process that will help you be a better cancer registrar and person. You have to keep trying. Will you fail? Of course. But, studies have shown that only 1%-3% of us will get back up, shake ourselves off and try again. So, you have to TRY to be in that top 3%! Try your best. Give it your all. Why not forget about all the statistics and past attempts and just go all out?

Secret #3: Life is worthwhile if you STAY. If you have accepted a position in a cancer registry, you need to stay during the good and the bad times. Stick it out and see things through. Sometimes life just happens, budgets collapse, people behave badly, and we’re asked to do more with less. But you still stay and do your best. Don’t end in the middle or leave unfinished business. Stand up, take a breath, and stay until you see it through. Don’t just walk away or turn your back on something as though it, or they, do not matter.

Because THEY matter, WE matter, and YOU matter! You deserve to stay and become your very best.

Secret #4: Life is worthwhile if you CARE. If you care, even just a little, you will get results in your work and life. If you care a lot, you will get incredible results! Cancer registrars are commissioned to care enough to make a difference. Be willing to stand up and care enough to make a difference in the life of your fellow registrars, an individual who is studying to become a cancer registrar, and the physicians and administrators in your organization. Care enough to make a difference in how you do your work, how you collect and report data so that it is nothing less than the most accurate and complete results that others rely upon us to deliver. Care enough to never leave a case unreported, never use unreliable data or resources, and never compromise your objectives or values. Care enough to do what is required of your job, no matter what the noise is that comes from all around you. Care enough to stand up for value, excellence, and quality, even if it means you’re threatened or pressured to do otherwise. Care enough to set records, to stand out from the crowd, and to change the status quo. Care enough to win!

On occasion I have been dubbed an idealist by some, criticized by others, and even dismissed by a few. I am grateful for those experiences because each time I have been challenged to understand what I really care about and why. I must follow my vision, passion, and philosophy for my work as a cancer registrar. While I have not yet achieved all of my goals and am a “work in progress,” my journey has been meaningful and brought incredible growth and learning. What is your experience and what meaning have you given to your life thus far?
Have you had a life-changing experience like this? I hope so. Because it is out of negative experiences that we know who we are and why we do this work. It’s not easy being a cancer registrar. But, I have to say, it may not be easy to work with cancer registrars either. We tend to be a baffling mystery to folks who do not understand the complexity of our work, its foundational principles, or our insatiable need for rules and guidelines.

I encourage you to take these 4 words: learn, try, stay, and care to the very core of your cancer registry heart. If you are feeling worn out, misunderstood or undervalued, you can take comfort and be inspired by them. Put them on a card and keep them in front of you. Understand how these 4 words and the secrets they carry can make a difference in your life. These 4 words: learn, try, stay and care, can change you forever if you will only let them determine your attitude and actions, and give meaning to your work as a cancer registrar.


This article is also available on video to be shared with your friends or replayed at your cancer registry educational conferences and events. To access the free video, use your web browser and navigate to: http://www.YouTube.com/CancerRegistrar. Subscribe to the channel to get updates and new videos that are released.

References

NCRA’s Workbook for the Staging of Cancer: A Companion Guide to the AJCC Cancer Staging Manual (7th edition) provides cancer registrars a tool to understand the concepts of TNM staging and explains how to apply them in a consistent manner. The workbook outlines the principles of TNM staging and offers an overview of the AJCC staging process. Ten primary cancers are described in detail and extensive exercises are provided. Rationales for the correct answers are also included.

Price:
NCRA member price: $109 • Non-member price: $119

Go to www.ncra-usa.org/stagingworkbook to order!
**Special Feature**

**CoC Outstanding Achievement Award: Recognizing Exceptional Programs**

The Outstanding Achievement Award (OAA) enables the American College Of Surgeons Commission on Cancer (CoC) to identify programs from which to draw best practice examples to be included in the Best Practices Repository, act as mentors to new programs, and to serve as invited presenters for face-to-face workshops and educational webinars.

When the CoC released upgraded standards for cancer programs in 2004, a numeric rating system was introduced that assessed the cancer program’s compliance with each of the 36 standards required for accreditation. As would be expected, ratings were defined for compliance and noncompliance with the standard along with a designation when the standard was not applicable to the program. The CoC took this opportunity to establish an additional rating to recognize performance with the standard that exceeded the minimum requirements. A commendation rating was developed for 7 standards that represent the full scope of cancer care—program leadership, quality data, clinical care, research, community outreach, education, and quality improvement. In addition to commendation levels for individual standards, the CoC established the OAA to recognize programs that met all standards at the time of survey and achieved commendation for standards eligible for commendation. Since its inception, the CoC has awarded the OAA to almost 550 programs. Each year the number of OAA recipients has increased. From small facilities to large, all categories are represented and many programs have received the award more than once.

The Accreditation Committee selects the commendation standards that form the basis for the OAA criteria each year. This enables the committee to recognize important concerns about a standard or the commendation criteria and make changes when either or both are deemed to be too restrictive or too broad to make the commendation rating meaningful. As a result, the commendation rating for standard 7.2 (registrar education) has never been part of the OAA criteria because the committee determined that enabling CTR attendance at a national meeting was beyond the reach of some programs. In addition, the 4.3 (physician staging) commendation was eliminated in 2009 when the committee updated the requirements for the standard to focus on appropriate treatment planning using evidence-based guidelines. The committee expects to continue to refine and add to commendations with the implementation of the 2012 standards.

*Cancer Program Standards 2012: Ensuring Patient-Centered Care* recognized eight standards for commendation that encompass cancer program leadership, clinical services, and data quality. The standards are:

- Standard 1.3: Cancer committee attendance
- Standard 1.9: Clinical research accrual
- Standard 1.11: Cancer registrar education
- Standard 1.12: Public reporting of outcomes
- Standard 2.1: College of American Pathologists (CAP) protocols
- Standard 2.2: Oncology nursing
- Standard 5.2: Abstracting timeliness
- Standard 5.6: Data quality

Each commendation is based on measurable criteria that present the cancer program with a challenging but achievable level of performance to be met in order to receive the higher rating. In most cases, the commendation criteria have been advanced for standards that were part of the 2004 manual. For example, the commendation level for abstracting timeliness is now set at 95% of the annual caseload each year where in the 2004 manual commendation was earned if more than 90% of the annual caseload had been abstracted within the 6-month time frame each year. The percentage of patients accrued to cancer-related clinical trials each year is another example. New standards were also selected for commendation, including the annual attendance of cancer committee members at cancer committee meetings and public reporting of outcomes. The former was established to reinforce the importance of cancer committee member participation in meetings where important cancer program decisions are made. The latter was created to encourage the cancer program to share patient or program outcomes with the public each year. In fact, the only possible rating for standard 1.12 (public reporting of outcomes) is commendation. This means that the cancer committee can choose whether or not to meet this standard. A cancer program can receive full accreditation without meeting standard 1.12 but the program will not be eligible for the OAA unless this standard is met by sharing program performance with the public each year.

Beginning with 2013 surveys and for the first time since its inception, the Accreditation Committee will use all of the commendation standards as the basis for awarding the OAA. This change will enable the CoC to identify the peak performing cancer programs and afford them the recognition they clearly deserve.

The CAnswer Forum Team are:

- M. Asa Carter, CTR
- Vicki Chiappetta, RHIA, CTR
- Anna Delev, CTR
- Debbie Etheridge, CTR
- Donna Gress, RHIT, CTR
- Karen Stachon
CORRECT ANSWERS FOR SUMMER 2012

JOURNAL OF REGISTRY MANAGEMENT CONTINUING EDUCATION QUIZ

THE CHALLENGES OF ABSTRACTING RELIABLE INFORMATION ON PATIENT RACE AND ETHNICITY: INITIAL OBSERVATIONS FROM THE DATA IMPROVEMENT PROJECT ON PATIENT ETHNICITY AND RACE (DIPPER)

1. Race and ethnicity are among the most important data used in any analysis of cancer incidence or mortality.
   a) True
   b) False

2. The collection of accurate and reliable data on race and ethnicity is crucial to understanding which of the following health/health care disparity issues?
   a) What disparities exist
   b) The extent of the disparities
   c) How to mitigate the disparities
   d) All of the above

3. Interviews with patient registration managers throughout Rhode Island revealed that:
   a) some electronic systems in Rhode Island hospitals failed to meet the Office of Management and Budget (OMB) guidelines on how to collect race and ethnicity
   b) most hospitals use the preferred 1-question format for race
   c) North American Association for Central Cancer Registries (NAACCR) allows 5 racial categories to be recorded for each patient
   d) hospital cancer registrars in Rhode Island frequently choose 3 or more race categories

4. Rhode Island hospital records accurately reflect the complexity of multiracial patients
   a) True
   b) False

5. The identification of Hispanic/Latino ethnicity is:
   a) more difficult for cancer registrars than identification of race
   b) readily available to cancer registrars in all Rhode Island hospitals
   c) most commonly listed on the face sheet in Rhode Island’s acute care hospitals
   d) made more challenging when race and ethnicity are listed side by side on the face sheet.

6. The identification of a patient’s ethnicity can be compromised by:
   a) clear procedural guidelines
   b) congruent information in the medical record
   c) discomfort with asking questions about race and ethnicity
   d) asking patients about race and ethnicity at each admission

7. In-depth interviews and direct observations at acute care hospitals in Rhode Island reveal that:
   a) race and ethnicity data are readily accessible to cancer registrars
   b) more time and attention should be devoted to the training of patient registration clerks
   c) race and ethnicity data recorded in the medical record are highly accurate
   d) all of the above

8. Cancer registrars can improve the identification of patient race and ethnicity through:
   a) alignment of electronic patient registration systems with cancer registry standards
   b) standardized training on capturing race and ethnicity data from hospital medical records
   c) guidelines for prioritizing data from different sources
   d) all of the above

9. The goal for the next phase of DIPPER (Data Improvement Project on Patient Ethnicity and Race) is:
   a) complete self-identification of race and ethnicity at each encounter
   b) complete self-identification of race and ethnicity only at the initial registration encounter
   c) visual identification of race and ethnicity by registration clerks
   d) the development of more acronyms for use in the drop-down menus used by registration clerks

10. Accurate, accessible, and complete data on race and ethnicity may be improved by:
     a) providing training on the collection of race and ethnicity
     b) encouraging physicians to report race and ethnicity when they order pathology reports
     c) asking hospitals to record race and ethnicity on the face sheet or other easily accessible document
     d) all of the above
EVALUATION OF PRIMARY/PREFERRED LANGUAGE DATA COLLECTION

Quiz Instructions: The multiple choice or true/false quiz below is provided as an alternative method of earning CE credit hours. Refer to the article for the ONE best answer to each question. The questions are based solely on the content of the article. Answer the questions and send the original quiz answer sheet and fee to the NCRA Executive Office before the processing date listed on the answer sheet. Quizzes may not be retaken nor can NCRA staff respond to questions regarding answers. Allow 4–6 weeks for processing following the submission deadline to receive return notification of your completion of the CE process. The CE hour will be dated when it is submitted for grading; that date will determine the CE cycle year.

After reading this article and taking the quiz, the participants will be able to:
• Discuss current practices on primary/preferred language data collection in hospital medical records
• Explain how low health literacy impacts health outcomes
• Describe how the standardized collection of primary/preferred language could help decrease health care disparities

1. Studies have demonstrated that low literacy and low health literacy are associated with:
   a) impaired patient-provider communication
   b) better patient compliance
   c) decreased hospitalization
   d) improved health

2. Language barriers are associated with:
   a) less health education
   b) worse interpersonal care
   c) lower patient satisfaction
   d) all of the above

3. Having an interpreter present serves as a substitute for language concordance between patient and provider.
   a) True
   b) False

4. Data from the 2010 Census Bureau indicate that during the last 3 decades:
   a) the number of people 5 years of age and older who speak a language other than English at home has decreased by half
   b) the number of speakers of non-English languages grew by 34%
   c) the number of people 5 years of age and older who speak a language other than English at home has grown at a rate that is 4 times greater than that of the overall US population
   d) the overall US population grew by 140%

5. Historically, data collection practices on language and communication have been:
   a) limited by systems that were incomplete and used incongruously
   b) augmented by systems that were complete and used congruously
   c) limited by systems that were complete and used congruously
   d) augmented by systems that were incomplete and used incongruously

6. Primary/preferred language data collection occurs in multiple different ways within various settings, resulting in:
   a) reduced ability to make comparisons across studies
   b) limited ability to aggregate primary/preferred language data study results
   c) impeded research efforts due to the lack of a common definition and standard codes
   d) all of the above

7. It is unclear how data on interpreter use is used to improve services for patients.
   a) True
   b) False

8. Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) standards:
   a) do not require facilities to maintain records on patients’ language and communication needs
   b) support the provision of care in a manner that is conducive to the cultural, language, literacy, and learning needs of individuals
   c) ensure that orientation and ongoing staff education are appropriate to the needs of facility administration
   d) do not include provisions for the collection of data

9. Primary/preferred language concordance between patient and provider:
   a) is less important for patients with complex illnesses such as cancer
   b) has no impact on the effective coordination of care
   c) provides access to health services that are required for effective treatment
   d) results in less timely attention to symptoms

10. One limitation of this study is:
    a) the use of unpublished documents and reports
    b) the use of aggregate studies for meta-analysis
    c) some studies may not have been identified and included
    d) the use of a thorough and extensive search strategy and literature review
Instructions: Mark your answers clearly by filling in the correct answer, like this ■ not like this X. Passing score of 70% entitles one (1) CE clock hour per quiz.
Please use black ballpoint pen.

1 A B C D
2 A B C D
3 A B
4 A B C D
5 A B C D
6 A B C D
7 A B
8 A B C D
9 A B C D
10 A B C D

Submit the original quiz answer sheet only! No photocopies will be accepted.

For Internal Use Only
Date Received:__________________
Amount Received:__________________
Notification Mailed:__________________
National Cancer Registrars Association

SPECIAL BIRTH DEFECTS CALL FOR PAPERS

Focus on Birth Defects Registry Methods, Spring 2013

The National Birth Defects Prevention Network (NBDPN) is pleased to announce that we are continuing our arrangement with the *Journal of Registry Management* (JRM) to provide a themed issue on birth defects surveillance, with papers focusing on birth defects program methods, approaches to assuring/evaluating data quality and completeness, and innovative approaches to improving the science and practice of birth defects surveillance. The NBDPN annual report will continue to be published in the winter in *Birth Defects Research Part A* (BDRA) and will focus primarily on birth defects epidemiology and etiologic research. This arrangement with the JRM enables the NBDPN to continue to advance the field of birth defects surveillance through publications that demonstrate methodological innovations and registry practice enhancements. The JRM birth defects themed issue (Spring 2012) can be accessed for a limited time at http://www.ncra-usa.org/i4a/pages/index.cfm?pageid=3307.

Manuscripts will be peer-reviewed and approved under JRM guidelines (http://www.ncra-usa.org/files/public/authorinfo.pdf). Manuscripts should be accompanied by a cover letter including the names, addresses and telephone numbers or email addresses for at least 2 potential reviewers uninvolved with the current work who might be willing to provide an objective and unbiased review. Manuscripts should have internal clearance, if required for agency submission, and be submitted by Wednesday, January 9, 2013 to jrm_nbdpn@nbdpn.org. Manuscripts may come from a single program, from multistate, and/or international collaborative efforts and activities.

We are especially interested in manuscripts focusing on the following areas:

1. Use of surveillance data to support public health program assessment, design, and evaluation
2. Cluster investigations
3. Statistical techniques, role and/or use of GIS, or meta-analyses
4. Quality assurance: general protocols, re-abstraction techniques, assessment methods, and/or evaluation
5. Methods for presentation and dissemination of birth defects data
6. Program methodology: role of advisory committees, model legislation, methods for planning and evaluation, and/or prenatal and stillbirth surveillance

Potential contributors may contact Dr. Russell Kirby at (813) 396-2347 or rkirby@health.usf.edu, or Dr. Wendy Nembhard at (864) 388-1737 or wnembhard@health.usf.edu, for more information or for a consultation on topics under consideration. If your topic is outside the topical areas listed above, please contact us because we would like to discuss it with you further.
CALL FOR PAPERS

Topic:
1. Birth Defects Registries
2. Cancer Registries
   - Cancer Collaborative Stage
   - Cancer and Socioeconomic Status
   - History
3. Trauma Registries
4. Recruitment, Training, and Retention
5. Public Relations

The Journal of Registry Management, official journal of the National Cancer Registrars Association (NCRA), announces a call for original manuscripts on registry methodology or research findings related to the above 5 subjects, and related topics. Contributed manuscripts are peer-reviewed prior to publication.

Manuscripts of the following types may be submitted for publication:
1. Methodology Articles addressing topics of broad interest and appeal to the readership, including methodological aspects of registry organization and operation.
2. Research articles reporting findings of original, reviewed, data-based research.
3. Primers providing basic and comprehensive tutorials on relevant subjects.
4. “How I Do It” Articles describe tips, techniques, or procedures for an aspect of registry operations that the author does particularly well. The “How I Do It” feature in the Journal provides registrars with an informal forum for sharing strategies with colleagues in all types of registries.
5. Opinion papers/editorials including position papers, commentaries, essays, and interviews that analyze current or controversial issues and provide creative, reflective treatments of topics related to registry management.
6. Bibliographies which are specifically targeted and of significant interest will be considered.
7. Letters to the Editor are also invited.

Address all manuscripts to: Vicki G. Nelson, MPH, RHIT, CTR, Editor-in-Chief, Journal of Registry Management, (770) 488-6490, vnelson@cdc.gov.

Manuscript submission requirements are given in “Information for Authors” found on the inside back cover of each Journal and on the NCRA Web site at http://www.ncra-usa.org/jrm.
Manuscripts may be submitted for publication in the following categories: Articles addressing topics of broad interest and appeal to the readership, including Methodology papers about registry organization and operation; Research papers reporting findings of original, reviewed, data-based research; Primers providing tutorials on relevant subjects; and “How I Do It” papers are also solicited. Opinion papers/editorials including position papers, commentaries, and essays that analyze current or controversial issues and provide creative, reflective treatments of topics related to registry management; Letters to the Editor; and specifically-targeted Bibliographies of significant interest are invited.

The following guidelines are provided to assist prospective authors in preparing manuscripts for the Journal, and to facilitate technical processing of submissions. Failure to follow the guidelines may delay consideration of your manuscript. Authors who are unfamiliar with preparation and submission of manuscripts for publication are encouraged to contact the Editor for clarification or additional assistance.

Submission Requirements

Manuscripts. The terms manuscripts, articles, and papers are used synonymously herein. E-mail only submission of manuscripts is encouraged. If not feasible, submit the original manuscript and 4 copies to the Editor. Manuscripts should be double-spaced on white 8-1/2” x 11” paper, with margins of at least 1 inch. Use only letter-quality printers; poor quality copies will not be considered. Number the manuscript pages consecutively with the (first) title page as page one, followed by the abstract, text, references, and visuals. The accompanying cover letter should include the name, mailing address, e-mail address, and telephone number of the corresponding author. For electronic submission, files should be 3-1/2”, IBM-compatible format in Corel WordPerfect®, Microsoft® Word for Windows®, or converted to ASCII code.

Manuscripts (Research Articles). Articles should follow the standard format for research reporting (Introduction, Methods, Results, Discussion, References), and the submission instructions outlined above. The introduction will normally include background information, and a rationale/justification as to why the subject matter is of interest. The discussion often includes a conclusion subsection. Comprehensive references are encouraged... as are an appropriate combination of tables and figures (graphs).

Manuscripts (Methodology/Process Papers). Methodology papers should follow the standard format for research reporting (Introduction, Methods, Results, Discussion), or for expository papers not reporting results (Introduction, Methods, Discussion), as well as the submission instructions outlined above.

Manuscripts (“How I Do It” articles). The “How I Do It” feature in the Journal provides registrars with a forum for sharing strategies with colleagues in all types of registries. These articles describe tips, techniques, or procedures for an aspect of registry operations that the author does particularly well. When shared, these innovations can help registry professionals improve their skills, enhance registry operations, or increase efficiency.

“How I Do It” articles should be 1,500 words or less (excluding references) and can contain up to 2 tables or figures. To the extent possible, the standard headings (Introduction, Methods, Results, Discussion) should be used. If results are not presented, that section may be omitted. Authors should describe the problem or issue, their solution, advantages (and disadvantages) to the suggested approach, and their conclusion. All submitted “How I Do It” articles will have the benefit of peer/editorial review.

Authors. Each author’s name, degrees, certifications, title, professional affiliation, and email address must be noted on the title page exactly as it is to appear in publication. The corresponding author should be noted, with mailing address included. Joint authors should be listed in the order of their contribution to the work. Generally, a maximum of 6 authors for each article will be listed.

Title. Authors are urged to choose a title that accurately and concisely describes the content of the manuscript. Every effort will be made to use the title as submitted, however, Journal of Registry Management reserves the right to select a title that is consistent with editorial and production requirements.

Abstract. A brief abstract must accompany each article or research paper. The abstract should summarize the main point(s) and quickly give the reader an understanding of the manuscript’s content. It should be placed on a page by itself, immediately following the title page.

Length. Authors are invited to contact the Editor regarding submission of markedly longer manuscripts.


Visuals. Use visuals selectively to supplement the text. Visual elements—charts, graphs, tables, diagrams, and figures—will be reproduced exactly as received. Copies must be clear and properly identified, and preferably e-mailed. Each visual must have a brief, self-explanatory title. Submit each visual on a separately numbered page at the end of the manuscript, following the references.

Attribution. Authors are to provide appropriate acknowledgment of products, activities, and support especially for those articles based on, or utilizing, registry data (including acknowledgment of hospital and central registrars). Appropriate attribution is also to be provided to acknowledge federal funding sources of registries from which the data are obtained.

References. References should be carefully selected, and relevant. References must be numbered in order of their appearance in the text. At the end of the manuscript, list the references as they are cited; do not list references alphabetically. Journal citations should include author, title, journal, year, volume, issue, and pages. Book citations should include author, title, city, publisher, year, and pages. Authors are responsible for the accuracy of all references. Examples:


Key words. Authors are requested to provide up to 5, alphabetized key words or phrases which will be used in compiling the Annual Subject Index.

Affirmations

Copyright. Authors submitting a manuscript do so on the understanding that if it is accepted for publication, copyright in the article, including the right to reproduce the article in all forms and by any means, shall be assigned exclusively to NCRA. NCRA will not return any reasonable requests by the author(s) for permission to reproduce any of his or her contributions to the Journal. Further, the manuscript’s accompanying cover letter, signed by all authors, must include the following statement: “We, the undersigned, transfer to the National Cancer Registrars Association, the copyright for this manuscript in the event that it is published in Journal of Registry Management.” Failure to provide the statement will delay consideration of the manuscript. It is the author’s responsibility to obtain necessary permission when using material (including graphs, charts, pictures, etc.) that has appeared in other published works.

Originality. Articles are reviewed for publication assuming that they have not been accepted or published previously and are not under simultaneous consideration for publication elsewhere. If the article has been previously published or significantly distributed, this should be noted in the submission for consideration.

Editing Journal of Registry Management reserves the right to edit all contributions for clarity and length. Minor changes (punctuation, spelling, grammar, syntax) will be made at the discretion of the editorial staff. Substantive changes will be verified with the author(s) prior to publication.

Peer Review

Concurrent manuscripts are peer-reviewed prior to publication, generally by 3 reviewers. The Journal Editor makes the final decision regarding acceptance of manuscripts. Receipt of manuscripts will be acknowledged promptly, and corresponding authors will be advised of the status of their submission as soon as possible.

Reprints

Authors receive 5 complimentary copies of the Journal in which their manuscript appears. Additional copies of reprints may be purchased from the NCRA Executive Office.
Beyond data collection there's knowledge

METRIQ® cancer registry software for integrated workflow, improved productivity

METRIQ does more than manage your cancer registry data; it helps increase productivity and improve data quality. Designed to be intuitive and efficient, METRIQ speeds the registrar through collection and reporting, transforming data into the knowledge clinicians need to analyze outcomes and better allocate patient care resources.

METRIQ has a dedicated team of certified tumor registrars ready to assist with registry operations and provide oncology data management support. The result: an overall customer service satisfaction rating of 97%.

You, your medical team and your patients deserve the best. Learn more at www.elekta.com/metriq.

Human care makes the future possible

ELEKTA