# REGISTRAR PIP

Visit SEER*Educate: A comprehensive training platform for registry professionals

## Improving the Quality of Histology Coding

### Introduction

The data provided by cancer registries are critical for epidemiological research, public health surveillance, health service planning, and the evaluation of the impact of interventions on cancer incidence, treatment, and survival. To meet these requirements, the data produced by cancer registries must be of high quality to reduce the chance of introducing bias to estimates of incidence and survival rates.

Poor coding performance often results when employees don't know exactly what they're supposed to do or why following a specific procedure is necessary. Discussions involving enhancing data quality invariably touch on how it impacts productivity. Everything extra is simply that . . . it's extra. Needing to do more with less is a common state of existence in most registries. Given this is the reality for registries, there is a growing importance in providing regular and relevant training for all staff that is capable of both assessing performance and enhancing technical skills to improve productivity because training will ultimately reduce time spent problem-solving.

This article examines the issues impacting the accuracy of current histology coding, which along with primary site, is one of the two most common data items used to classify malignancies. Today, histology coding is heavily impacted by the rapid expansion of medical knowledge. Following a review of the histology coding results completed prior to the in-person 2021 Surveillance, Epidemiology and End Results (SEER) Program Workshop, it is clear there is a need to learn how to use the updated materials provided by standard setters to code this data item more consistently and accurately.

### Results - Coding Accuracy 

Prior to the release of the SEER Hematopoietic and Lymphoid Neoplasm Database, hematopoietic and lymphoid histologies were challenging for many registrars to code accurately. This database allows registrars to electronically search for a diagnosis more quickly and to find information related to alternate names used to describe the same disease process, as well as clinical information related to specific tumor markers, immunohistochemical testing, genetic testing, or other characteristics that define a diagnosis or a particular histology which can be used to improve coding accuracy.

With the release of the latest edition of the Solid Tumor Rules and the ICD-O-3.2 Coding Table, the SEER*Educate team learned the coding of solid tumor histologies has proven to be significantly problematic for registrars. Is providing repeated training in how to use these rules enough to improve coding accuracy? Perhaps the many rules associated with coding solid tumor histology for various primary sites is now sufficiently challenging as to require something more than the text version of the Solid Tumor Rules to improve histology coding for these tumors.

Prior to the 2021 SEER Workshop, 947 registrars completed histology coding for 25 cases (9 hematopoietic/lymphoid diseases and 16 solid tumors). The accuracy rate for hematopoietic/lymphoid diseases was 70% versus 53% for solid tumors. The overall accuracy rate for the 25 cases was 59% with only 2 hematopoietic/lymphoid cases achieving an accuracy rate over 85%. What were the problem areas we discovered?


CSS is funded by the National Cancer Institute's SEER Program, Contract Number HHSN261201800004I

## Error Types - Hematopoietic/Lymphoid Cases

- **Genetics and Molecular Changes**

Unfortunately, coding histology for the pathology reports that included genetic abnormalities or molecular data proved most difficult to code accurately during the workshop. There was a tendency to code the NOS histology rather than use the genetic and molecular information to code a more specific histology. This is likely due to inadequate training and practice in deciphering this type of information.

While making sense of all the genetic abnormalities, mutations, and rearrangements involving the hematopoietic and lymphoid neoplasms is challenging, **Appendix D** titled **"Introduction to Genetic Nomenclature,"** found at the end of the Hematopoietic and Lymphoid Neoplasm Coding Manual, provides a brief introduction to some of the nomenclature, genetic alterations and molecular information needed to successfully code hematopoietic histologies. **Hematopoietic and Lymphoid Neoplasm Coding Manual Published August 2021 (cancer.gov)**

Genetics and biomarker test results will continue to grow in relevance when coding the more specific histologies. One frustration in working with this type of information is that only sometimes is this information applicable to coding the more specific histologies. Learning how and when to apply genetic and biomarker evidence when coding histology takes practice. When in doubt as to the relevance of this type of information provided by the pathology department, check the medical record for clinical interpretations from the managing physician or the oncologist treating the patient as to the histology being treated.

Learning how to interpret available genetics and biomarker information to code histology is an important aspect in being able to ultimately assess the use of targeted therapy to identify and attack certain types of cancer cells and improve patient outcomes for particular histologies.

- **Ambiguous Terminology**

While the growing use of genetic and molecular information is important in coding hematopoietic/lymphoid cases, so too is learning the importance of following the existing rules associated with the use of ambiguous terminology when coding histology for these cases. The change in use of ambiguous terminology for these cases has been in place for over a decade and yet errors in applying the following rules are still problematic today.

✓ If a single histology is described by ambiguous terminology, the single histology can be coded.

✓ If the NOS histology and a more specific histology described by ambiguous terminology are given in a pathology report, further clinical confirmation of the more specific histology must be found in the medical record and documented in the abstract in order to code the more specific histology. In other words, the ambiguous terminology alone cannot be used to code the more specific histology in this situation.

## Error Types – Solid Tumor Cases

- **The Fundamentals**

Improving histology coding sometimes boils down to simply following fundamental procedures in order to improve coding accuracy. There is not an application used to code solid tumor histologies as we have when coding heme/lymphoid cases. If there were, such an application might be able to help ensure such procedures were consistently followed to assign the appropriate primary site/histology codes. Lacking such an application forces the manual application of the required priority processing of the procedures. By not following the procedures, those of us who continue to rely solely on the use of the histology dropdowns in registry software will continue to insert coding errors into the database. While the use of dropdowns may prove a quicker way to put a value in this field, it is not the most accurate for all cases.


CSS is funded by the National Cancer Institute's SEER Program, Contract Number HHSN261201800004I

**Solid Tumor Coding Steps**

The steps to coding histology using the Solid Tumor Rules include:

1. Select the appropriate schema from the Solid Tumor Rules.
2. Determine which document or source has the highest priority for the schema.
3. After choosing the document and section of the document with the highest priority, identify the provisional histology or histologies for the tumor under consideration.
4. *Before* applying the H Rules, review the schema section titled "Coding Histology."
5. *After* applying the Coding Histology rules, apply the H Rules stopping at the first applicable one.
6. If the applicable H Rule does not provide the correct histology to code, use the appropriate Table (either the histology combination codes table OR the Specific Histology, NOS, and Subtypes/Variants table) in the Solid Tumor Rules schema.
7. If the applicable Table does not have the histology for the case, use the ICD-O-3.2 Coding Table.
8. If the histology cannot be determined, refer to the SEER Inquiry System (SINQ) to determine if this situation has already been addressed and if not, submit a question regarding the histology to Ask a SEER Registrar.

With practice stepping through these rules in the above priority order will come the ability to quickly identify which pathology report/section of the pathology report or source document can be used to code histology, when to ignore histologies described by ambiguous terminology, and how to efficiently move through the schema-specific H Rules and Tables.

- **Ambiguous Terminology . . . Revisited**

Given the College of American Pathologists (CAP) indicates pathologists are no longer to use ambiguous terminology to describe an applicable more specific histology, neither can registrars. Even though changes were made over three years ago in the use of ambiguous terminology for the Solid Tumors Rules schemas, cases involving these terms still represent one of the most common coding errors for solid tumors.

The following rules represent the only three times ambiguous terminology can be used to code histology:

✓ The only diagnosis available is **one histology** described by ambiguous terminology (e.g., a final diagnosis of "favor adenocarcinoma" is coded to 8140/3 is not coded the same way as a diagnosis of "carcinoma, favor adenocarcinoma" which is coded to 8010/3).

✓ There is a NOS histology and a more specific histology described by ambiguous terminology and the specific histology is **clinically confirmed** by a physician and the registrar has documented the physician's statement in text that is transmitted to the central registry.

✓ The patient is **receiving treatment** based on the specific histology described by ambiguous terminology.

In the majority of cases, the NOS code from the pathology report, which also contains a more specific histology described using ambiguous terminology, does not need to be used if the hospital registrar documents the physician's clinical confirmation which supports the more specific histology diagnosis in the abstract. However, failing to provide that clinical text documentation will result in a large number of cases being recoded at the central registry to the NOS histology described in the pathology report when a more specific histology is described using an ambiguous term.

In our central registry we conducted a review of 100 lung cases and discovered 95% of the time the NOS histology with a more specific histology described using ambiguous terminology on a pathology report was confirmed clinically. However, only 50% of the time the abstractor documented those clinical finding in the abstract to support the coding of the more specific histology. There is clearly an abstracting training issue to be addressed. **Without the clinical confirmation and documentation** of the more specific histology described


CSS is funded by the National Cancer Institute's SEER Program, Contract Number HHSN261201800004I

using ambiguous terms on the pathology report, the abstracting omission results in what appears to be histology coding errors when viewed by those in the central registry. As with the coding of any field on an abstract, proper documentation is required to support each coding decision.

- **Need to Use References**

Staying current on new terminology, changes in behavior/histology and reportability issues associated with the ICD-O-3 proved challenging for registrars who did not routinely reference required documentation when coding histology. Many proved unaware of the changes to behavior codes resulting in reportability updates as well as changes made to histology terms previously considered non-reportable that are now reportable and vice versa. Several of the SEER Workshop cases were selected specifically to assess the awareness of the North American Association of Central Cancer Registries' (NAACCR) ICD-O-3 Work Group's recommendation to use ICD-O-3.2 jointly with the 2021 ICD-O Histology and Behavior Code Update Tables, Solid Tumor Rules, and Hematopoietic and Lymphoid Neoplasm Database. Histology/behavior coding accuracy ranged from 46% to 84% for the four applicable cases included in the Workshop with an average of 66%.

- **Biomarkers**

Currently, biomarkers are most frequently used to differentiate histologies for the CNS primaries. For example, in the Malignant CNS schema, the coding hierarchy lists biomarkers from resection pathology as the highest priority when deciding the appropriate histology to assign a case. These biomarkers not only help determine the type of treatment recommended but when accurately reflected in the registry database will allow researchers to assess whether some biomarker-defined histologies have better outcomes than others. Learning to recognize applicable biomarker test results in pathology reports and how to interpret them correctly for coding histology will only grow in importance and relevance.

- **Check SINQ . . . Submit a Question to Ask a SEER Registrar**

When having trouble coding histology because it isn't found in any of the resource material used to code this data item, check the SEER Inquiry System (SINQ) for help. SINQ is a searchable collection of questions asked by cancer registrars when they run into problems coding various data items on cancer cases. The questions and answers include both clarifications to existing and historical coding rules and often address new coding issues not covered in the existing material available to registrars. The latter oftentimes applies to new ways pathologists express histologies. If unsuccessful in finding an answer to the histology coding question in SINQ, submit a new question using the Ask a SEER Registrar form available at https://seer.cancer.gov/registrars/contact.html.

## Survey 


- **Background**

As we reviewed the histology coding results on the SEER*Educate exercises for those planning to attend the SEER Workshop, we were concerned about the potential impact on histology coding accuracy for cases registrars were entering into their databases. If rules were not being followed when coding exercises, were the rules being applied accurately when performing daily coding activities? The current coding rules have been in place for over three years. The SEER*Educate coding results indicate something isn't working. What is it?

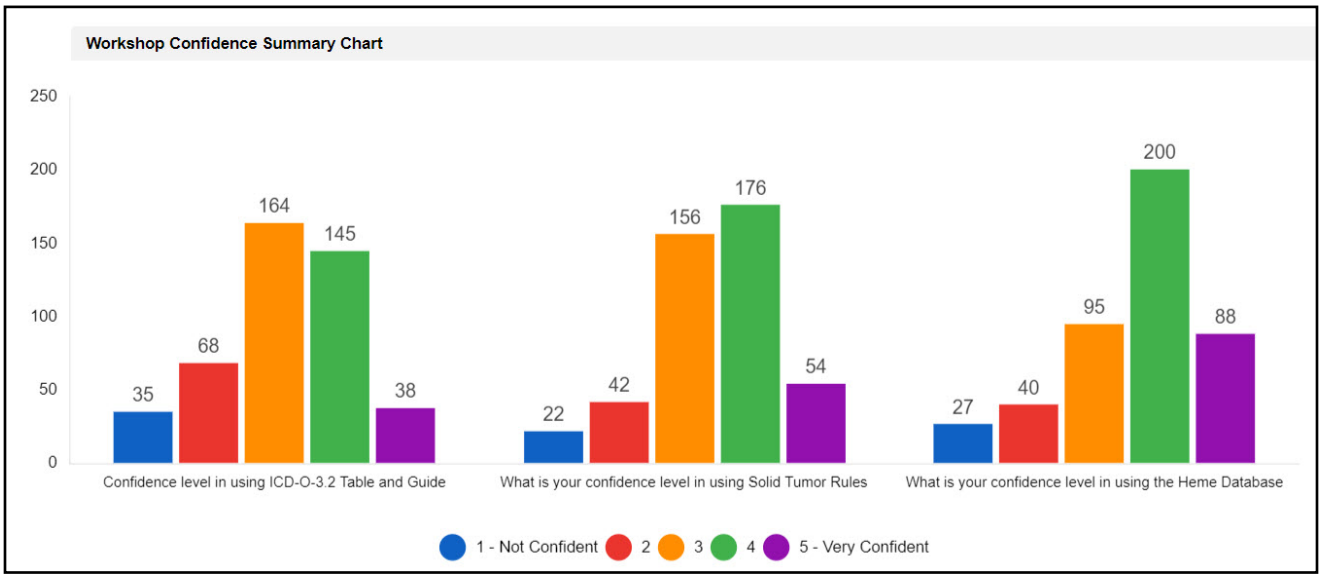
We wanted to know how confident registrars were in their ability to apply the current resource material and what they view are the problems and potential solutions to improve the accuracy and consistency of histology coding. After completing all 25 cases in the SEER Workshop training module in SEER*Educate, registrars were asked to anonymously complete a survey about the rules. The survey was prepared asking participants their confidence level in using the ICD-O-3.2 Table/Guide, Solid Tumor Rules, and the Heme Database.


CSS is funded by the National Cancer Institute's SEER Program, Contract Number HHSN261201800004I

- **Results**

The survey was distributed to the 947 participants who completed all 25 SEER Workshop cases. A total of 450 (47.5%) responses were received and the results are show in the Confidence Summary Chart. Given the types of histology coding errors observed, it is not surprising 288 registrars responded with higher confidence scores (4 and 5) associated with using the Heme Database, 230 registrars in understanding how to follow and apply the Solid Tumor Rules and only 183 registrars comprehending the need to use the ICD-O-3.2 Table/Guide to identify new histologies and changing tumor behavior associated with existing histologic types that impact reportability.

**Confidence Summary Chart**




- **Registrar Comments**

Many registrars were not only surprised by their scores but disheartened by them. Frustration levels were highest among those who have worked in registries the longest because they more keenly feel the impact of the seemingly unending cycle of unlearning and relearning necessary to be able to apply the latest rules and guidelines when coding histology.

When asked their opinion about why the histology coding accuracy was so low, registrar responses reflected very similarly themed comments:

✓ Coding rules seems fluid
✓ Difficulty in finding the time to keep up with latest changes
✓ Too many resources to review to code a single field
✓ The procedure is too time-consuming to follow for coding histology
✓ Rule formatting in manuals is difficult to follow
✓ Uniqueness of site-specific rules requires added effort to correctly apply for a given primary site

Given the differences between the histology coding accuracy rate for hematopoietic/lymphoid diseases (70%) and solid tumors (53%) and registrars' confidence level in using the Heme Database (288 with scores of 4 or 5) over the Solid Tumor Rules (232 with score of 4 or 5), it's not surprising the suggestion of creating a Solid Tumor Database was proposed. When considering histology coding at a very high level, we can see both heme/ lymphoid diseases and solid tumors:

✓ Have similar rules for using ambiguous terminology


CSS is funded by the National Cancer Institute's SEER Program, Contract Number HHSN261201800004I

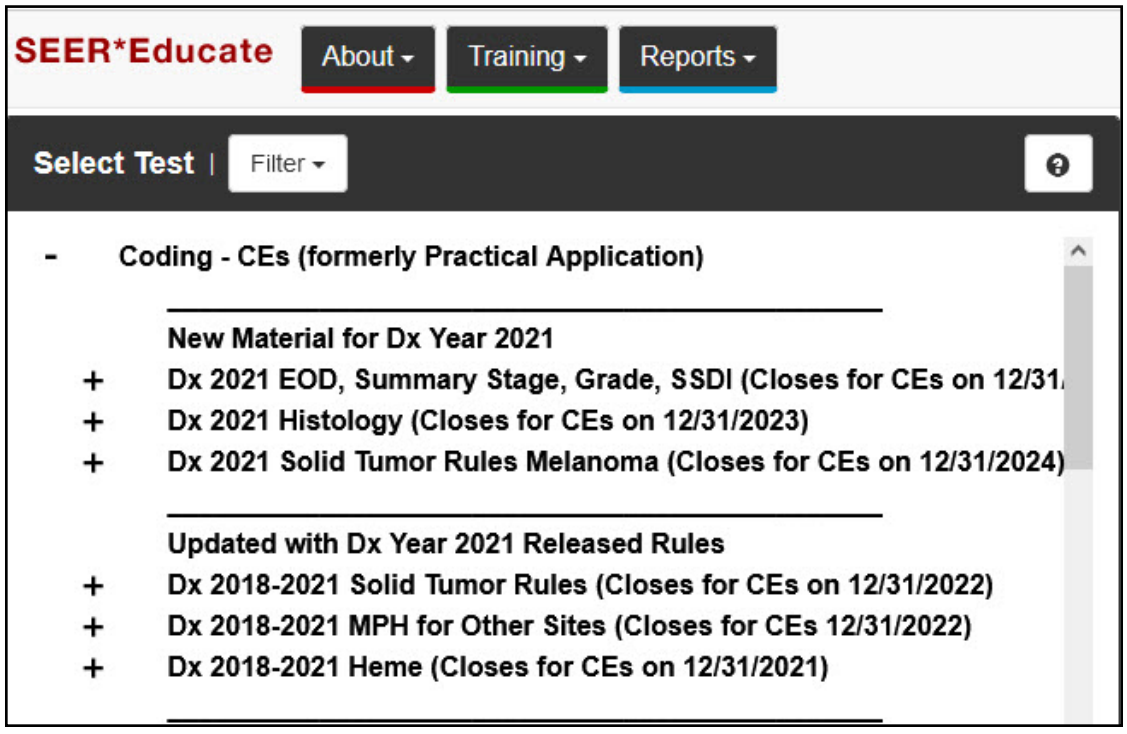
- ✓ Include instructions to code the more specific histology when specific criteria have been met
- ✓ Consider biomarker results if applicable

A database with an effective search functionality could improve histology coding for solid tumors in the same way the database for heme/lymphoid cases did. For solid tumors, manually searching multiple rule sets by primary site, tables, and a spreadsheet to code histologies has proven more than a little challenging given the low accuracy rate (53%). It would be quicker and more accurate to enter a diagnosis date, primary site and the text associated with the histology and have the database return the correct histology based on the appropriate scheme rule. The Solid Tumor Rules wouldn't go away, they become the cornerstone of a database that includes algorithms that can propose the histology based on the system's interpretation of the rules, tables, and the ICD-O.

## Conclusion

Given the importance of histology as one of the key data items used to classify malignancies, standard setters need to recognize the challenges associated with improving coding accuracy involve more than training that will continue to require manual coding of this field. An electronic application needs to be developed to enhance coding accuracy of this critical data item. One option is to explore the development of a corresponding solid tumor database similar to the current SEER Heme Database. A second option is to encourage standard setters to pursue ongoing refinements to Natural Language Processing (NLP) tools that include algorithms that allow free text pathology report documentation to be used to more accurately auto-code histology while simultaneously considering the required general guidelines, site-specific rules and tables, and the ICD-O. After three years of training and working with manual options for coding histology, it is clear we all need some automated help to improve our coding accuracy and consistency.

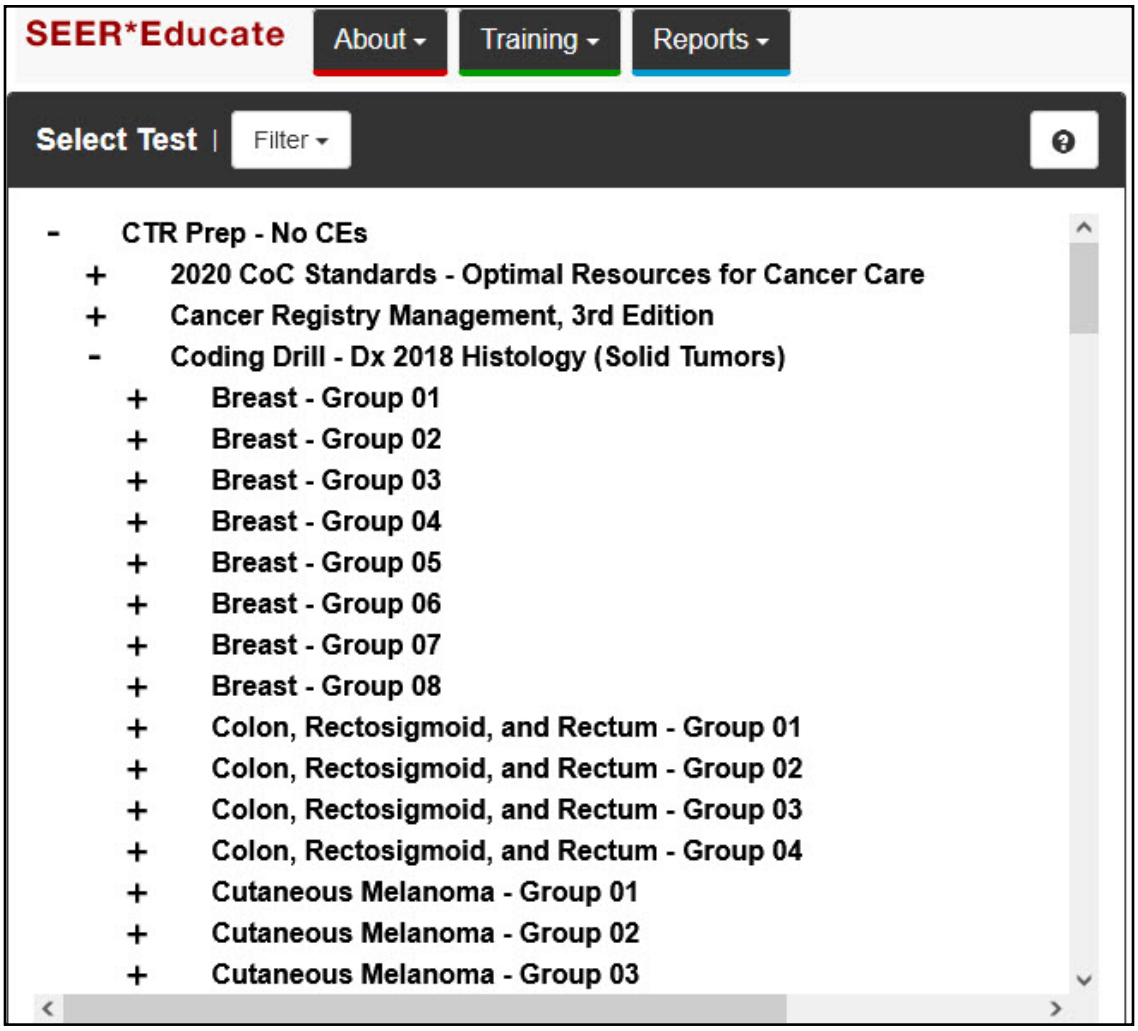
What can we, as registrars, do now to improve our coding of histology? We could identify the largest volume site we each personally abstract of breast, colon, head and neck, kidney, lung, melanoma, malignant CNS, non-malignant CNS, and urinary and then work through the 5 related solid tumor cases in SEER*Educate, perhaps doing a case per week. A case takes approximately 15 to 20 minutes to complete and read the rationales if you follow our registry's SEER*Educate practice of "do not agonize over any one case." This training activity at one case per week gets each of us focused on a specific set of STR rules with a set of preferred answers/rationales to see if we are still on the same page as the standard setter.





CSS is funded by the National Cancer Institute's SEER Program, Contract Number HHSN261201800004I

After completing the 5 STR cases, we could also look at the CTR Prep Coding Drill for Dx 2018 Histologies, which are grouped by site. These exercises are generally quicker because one is coding site, histology, and behavior. Perhaps on Mondays, work through a group of 5 cases.




The long-term solution of improved histology coding does involve the standard setter developing a better means of presenting the rules to the registrar. The near-term solution is for us to incorporate whatever amount of training, even a case a week or a case a month, into our schedules because learning to code histology is not a "once and done" activity.


CSS is funded by the National Cancer Institute's SEER Program, Contract Number HHSN261201800004I