

Chapter 9: Screening for Reportability

SEER*DMS uses a combination of automated and manual screening tasks to determine whether a source record contains data related to a reportable case. The Reportability field in the record table is set during these processes. The following is an overview of the workflow for a reportable record by record type. Wherever it says that a record will build a new CTC it means that it will build a new Patient Set with a new CTC or it will build a new CTC within an existing Patient set. Overview of the workflow for a reportable record:

If a source record's Reportability field is set to reportable:

- Abstracts – a reportable abstract builds a new CTC or is linked to an existing CTC. This work is completed via match and consolidation tasks. The workflow for a reportable abstract is not complete until it is linked to a CTC (the one it created or one that it matched).
- Pathology reports – the record may also be used to build a new CTC, like an abstract. However, not every reportable pathology report follows a "build now" workflow. The rules are based on registry criteria. An Abstract Facility Lead (AFL) is created for every reportable pathology record. As is true with all AFLs, the AFL stays open until an abstract is received, registry AFL criteria are met, or registry staff review and close the AFL. Registry staff can run a Build CFO process to create new CTCs from pathology records that are unlinked and have an open AFL.
- Case finding records – a reportable case finding record will not build a CTC automatically. An AFL is created for a reportable case finding record, per registry rules. Similar rules are applied to close AFLs as were described above for pathology reports. Registry staff can also run a Build CFO process to create new CTCs from case finding records.
- Death certificate records – an AFL is also built for each reportable death certificate record. SEER*DMS applies rules based on the NAACCR death clearance manual to match reportable death certificates to CTCs. Registry-defined AFL algorithms determine when to close the AFL. Registry staff can run a Build DCO process to create new CTCs from death certificates.

Records that are not reportable are used to update follow-up information for a patient. Some non-reportable records are purged from the database, per registry criteria.

As the workflow diagrams in Chapter 4 show, the task to Screen for Reportability occurs early in the workflow after Autocoding and Record Edits. If possible, an automatic Screening task will set the reportability flag. Otherwise, either the standard **Screening** task or the **Pathology Screening task** will be created so that a user can set the field manually.

In this chapter, you'll learn about

- Coding Site, Histology, Behavior, and Reportability
- Values of the Reportability Flag
- Automated Screening Task
- Auto-coding Pathology Reports
- Path Screening Tasks
 - Highlighting Supporting Text
 - Coding Reports with Multiple Primaries
 - Using the Auto-complete Feature
 - Display of Keywords and Dates
 - Standard Screening Tasks

Coding Site, Histology, Behavior, and Reportability

A manual screening task to set reportability is rarely required for abstracts and other records that contain coded values for site, histology, and behavior. The auto-screener is typically able to set the reportability flag based on the registry's screening algorithm. A manual screening task is only created for abstracts if the auto-screener could not determine reportability because site, histology, or behavior is not coded. This occurs in a very small percentage of abstracts.

The NCI and the Department of Energy (DOE) engaged in an interagency collaboration as part of the Cancer Moonshot, a federal initiative dedicated to ending cancer as we know it. Natural Language Processing (NLP) algorithms were developed as part of the NCI-DOE MOSSAIC project. MOSSAIC stands for Modeling Outcomes Using Surveillance Data and Scalable Artificial Intelligence for Cancer. MOSSAIC NLP algorithms are used in SEER*DMS to increase the automation of pathology data processing. The algorithms are implemented as Application Programming Interfaces (API). An API allows one application, in this case SEER*DMS, to call another application (an NLP algorithm).

- The path extraction API auto-codes site, histology, behavior, and laterality on a path report based on the narrative text fields.
- The reportability API, currently in beta testing, will be used to filter non-cancer reports and identify pathology reports related to reportable cancers.

The path extraction API is executed during the Autocoding task. A manual Pathology Screening task is created if the report could not be auto-coded. The Pathology Screening task has a specialized interface that highlights keywords in the text fields to facilitate the coding of site, histology, behavior, and laterality. The user codes site, histology, behavior, and laterality if they consider the report to be reportable. If it is a non-reportable report, then the user does not need to code those data items. They can flag the record as non-reportable, code a non-reportable reason, and enter an optional comment. If the user sets the record to non-reportable then the automated screening rules will not be applied. If the user codes site and morphology, then the reportability field will be set via the automated screening task.

The automated screening task sets the record reportability field after site, histology, behavior, and laterality are coded via the imported values, path extraction API, or a user in a manual Screening or Path Screening task. Note: imported values of these fields are available for nearly all abstracts. It is also true that a small percentage of pathology reports have coded values for these fields in the original HL7 import file. The imported values are set in the record, if available in the HL7.

A manual task is created if the auto-screener cannot determine reportability; or if, according to your registry's algorithms, non-reportable abstracts always require review. Some registries only expect to receive reportable abstracts and, therefore, require a manual review for any abstract with a non-reportable site and morphology.

Electronic Pathology Reports

Most electronic pathology reports received by SEER*DMS registries are in a structured narrative format. This is narrative text divided into labeled sections: clinical history, nature of specimen, gross observation, microscopic observation, final diagnosis, comments, etc.

Information related to site and morphology are provided in text and not in coded data items. SEER*DMS applies Natural Language Processing (NLP) algorithms to code site and morphology codes. Reports that cannot be autocoded are sent to a Path Screening task for manual coding.

Paper Pathology Reports

A small percentage of path reports are paper-based reports scanned and loaded as image files into SEER*DMS. A SEER*DMS user codes site and morphology in an Image Data Entry task. The record moves forward to the automated screening task. A manual Screening task is created if the auto-screener cannot determine reportability based on the site and morphology.

Imaging Report Records

Imaging reports contain a physician's interpretation of diagnostic images (X-ray, PET/CT scans, mammogram, ultrasound, etc.). Imaging report records are received in Health Level 7 (HL7) format like e-path. However, NLP algorithms used for e-path cannot be applied to imaging reports at this time. NLP algorithms for imaging report records may be developed in the future. If the registry's workflow requires coding of imaging reports then the reports are coded in manual tasks.

Death Certificate Records

All SEER*DMS registries receive electronic death certificates. These are data files with a record for each death certificate. A death certificate record contains patient identifiers, primary cause of death, secondary causes of death, and related information. A SEER*DMS polisher auto-codes site and morphology based on the first listed cause of death that is a reportable cancer. The auto-screener determines reportability. The Morphology polisher codes site and morphology on death certificate records; documentation is provided in SEER*DMS (go to Help > Polishers).

Health Index Records

Hospital discharge data are processed as health index records in SEER*DMS. The same polisher that codes site and morphology based on cause of death is applied to disease codes on health index record. Refer to help text for Morphology polisher in SEER*DMS (go to Help > Polishers).

Values of the Reportability Flag

The possible values for the reportability flag are described below. Some values are not relevant for all record types. The criteria used to define each value are defined in the registry's configuration settings and documented on the Screening help page.

Reportable – The record contains data for an eligible case as defined by your registry's criteria.





Auditable – Deprecated and being phased out in original SEER*DMS registries that use this value. Auditable was intended to be used if the record contains cancer-related data but is not a reportable case. A value of "auditable" was assigned to flag the record for case finding audits.

Non-reportable - The record contains data for a disease that is not collected by your registry. SEER*DMS processes the record to obtain follow-up information.

Unknown – Reportability flag has not been set. The record may not have been screened at all; or the screening task may need to be completed in a manual Screening or Pathology Screening task.

Not applicable – The record is not a medical record and does not require screening, for example, it may be a supplemental record from a linkage used to set vital status and date of last contact.

The following icons are used in SEER*DMS to show the value of a record's reportability flag:

-  Non-reportable
-  Auditable
-  Unknown (the record has not yet been screened)
-  Not applicable; the record type is not screened (e.g., supplemental records)

Automated Screening Task

The Automated Screening Task is one of the first steps in the workflow, executed after the record is imported and the auto-coding polishers are applied (see the Workflow Diagrams in Chapter 4). The screening task applies an algorithm to set the Reportability Flag.

The auto-screener algorithm is a three-step process. Step 1 applies validation rules on the fields required by the algorithm. The reportability flag is set to unknown; and a manual Screening or Path Screening task is created if one of these rules fails. These rules check for valid values in site, histology, and behavior. Registry-specific rules for other fields may also be applied.

Step 2 applies a second set of rules, the “reportable rules”. The reportability flag is set to reportable if one of these rules returns true. This set of rules is based on the SEER and/or NPCR case finding rules, however, registry specific logic is applied if the registry has additional reportable disease codes. The screen shot below shows an example of reportable rules used in 2021.

Reportable Rules. Reportability will be set to Reportable if any of the following are true:

- ICD-O-1 (only if no ICD-O-1 value is blank): year is between 1973 and 1991 (inclusive), behavior is 2 or 3 and not(site is 173XX and histology is 8000-8004, 8010-8045, 8050-8082 or 8090-8110)
- ICD-O-2 (only if any ICD-O-1 value is blank): year is between 1973 and 1991 (inclusive), behavior is 2 or 3 and
 - not (site is C44X and histology is 8000-8004, 8010-8045, 8050-8082 or 8090-8110)
 - not (site is C53X and behavior is 2 and year >= 1996)
- ICD-O-2: year is between 1992 and 2000 (inclusive), behavior is 2 or 3 and
 - not (site is C44X and histology is 8000-8004, 8010-8045, 8050-8082 or 8090-8110)
 - not (site is C53X and behavior is 2 and year >= 1996)
- ICD-O-3: year is 2004 or later, behavior is 0 or 1 and site is C70X-C72X, C751, C752 or C753
- ICD-O-3: year is 2001 or later, behavior is 2 or 3 and
 - not (site is C44X and histology is 8000-8005, 8010-8046, 8050-8084 or 8090-8110)
 - not (site is C53X and behavior is 2 and year >= 1996)
 - not (site is C619 and histology is 8148 and behavior is 2 and year >= 2001)

Step 3 is executed if the validation rules passed, and all reportable rules returned false. Step 3 determines whether the reportability flag should be set to auditable or non-reportable, based on registry-defined logic.

Auto-coding Pathology Reports

The Path Coding polisher sets primary site, histology, behavior, and laterality for electronic path reports. The polisher predicts values for the fields using Natural Language Processing (NLP) algorithms. The NLP algorithms were developed as part of the NCI-DOE Collaboration Project. The NLP algorithm returns six prediction flags (true/false):

- There is one report-wide prediction flag. This flag is true when the NLP algorithm predicted values with the highest level of confidence.
 - The accuracy rate for each individual field is greater than 97%; and the accuracy rate for all fields combined is also above 97%.
 - If the report-wide prediction flag is true, then the path report is auto-coded.
- There are five field-level prediction flags, one for each of these fields: 3-character site, 4-character subsite, histology, behavior, and laterality.
 - The NLP algorithm is extremely accurate, but not quite as confident if the report-wide flag is false but all field-level flags are true.
 - A percentage of these reports will be auto-coded. A random sample will be sent to manual path screening for QC purposes.

Auto-coding Rules:

- If the report-wide prediction flag is true:
 - Set primary site, histology, behavior, and laterality.
 - The polisher will use the 4-character subsite if it was successfully predicted (this is expected when the report-wide prediction flag is true).
- If the report-wide prediction flag is false:
 - If field-level prediction flags for site, histology, behavior, and laterality are true then set the fields.
 - The polisher will use the 4-character subsite if it was successfully predicted; otherwise, the 3-character site will be used.
 - Otherwise, do not set the fields and create a manual path screening task.
 - This happens when NLP could not predict one or more fields: (site and subsite flags are both false) or hist flag is false or beh flag is false or lat flag is false.

If the 3-character site is to be used:

- C21 is converted to C210
- C14, C22, C30, C38, C42, C48, C76 – no conversion; create a manual path screening task to code site on the record.
- For all other sites, a 9 is added as the 4th character.

All fields are validated against their lookups. A manual task is created if a value is not valid.

Path Screening Tasks

Requires system permissions: *screening* and *rec_edit*

Use the Path Screening task to code site, morphology, and reportability on path report and imaging report records. To start a Path Screening session, filter the worklist by Task Type = **Path Screen**. Sort the tasks in the order that you wish to complete them.

Tip: Sort by the patient's last name and the list will be ordered by Last Name, First Name, DOB, Event Date (descending), Path Number, and Specimen Received Date (descending). The Event Date is typically the date specimen collected. This sort order ensures that you see all records for the same patient with the latest report listed first. If you click the Last Name column header again then the sort order will be reversed. You will work on names that start with a letter later in the alphabet and the report dates will be in ascending order.

Click a task ID to open a Pathology Screening Task. As you complete each task, the next task in the list will open automatically.

The polisher to auto-code the path report using NLP algorithms is controlled by a registry option. It will be turned on in some registries in 2021 for testing. It will be turned off in other registries. The polisher only returns results for path reports and does not predict values for imaging reports. The screen show below shows a sample path report in a Path Screening task. In this example, the NLP polisher did not predict values because the polisher is turned off.

SEER*DMS Demo 21.65+ Pathology Screening add flag Task View Manage Tools System Help

HL7 E-Path

Save **Cancel**

Reportability *
Based on coded values (recommended)

Primary Diagnosis

Site *

Laterality *

Histology *

Behavior *

Add Primary

REC-54

PATTERSON, JUDY C 330-65-4321 <unk sex> MRN: 54321-6
Unlinked b. 07-11-1936 State: FAC-1000

EPath Sites

EPath Morph

Clinical Hist
Screening colonoscopy, Maternal hx of adenocarcinoma of colon age 57

Nature of Spec
2 cm polyp ascending colon 2 mm polyp in sigmoid colon

Gross Path
A. The first container is labeled "ascending colon." It contains a polypoid piece of tan mucosal tissue measuring

Final DX
A. Ascending Colon
SESSILE SERRATED ADENOMA (POLYP) WITH LOW-GRADE ADENOMATOUS DYSPLASIA.
B. Sigmoid Colon
TUBULAR ADENOCARCINOMA

Supp Rpt Add

Staging

To code site, morphology, and reportability in a Path Screening task (without polisher results):

1. Review the pathology text.
2. If the report indicates cancer:
 - a. Code values for Site, Laterality, Histology, and Behavior.
 - b. Leave the Reportability field set to "Based on coded values".
 - c. Click Save. SEER*DMS will automatically calculate reportability based on the site and morphology that you entered.
3. If this is not a reportable report:
 - a. Use the Reportability drop-down in the left panel and select Non-reportable.
 - b. Select a reason from the Non-reportable Reason drop-down.
 - c. Enter a comment.
 - d. Click Save.

Note: Most registries use two values for reportability (Reportable, Non-reportable). A few registries also categorize path and imaging reports as Auditable. This value is used for specific rules defined by registry management. If you select Auditable then you should also enter a reason code and comment.

The next screen shot shows the same report in a Path Screening task, but the polisher to auto-code the report is turned ON. The confidence level of the polisher was not high enough to auto-code the report, therefore, the report needs to be manually coded via Path Screening.

SEER*DMS 21.65+ Pathology Screening add flag Task View Manage Tools System

HL7 E-Path REC-103190001

Save **Cancel**

Reportability *
Based on coded values (recommended)

Primary Diagnosis

Site *

C187 : Sigmoid colon	69%
C182 : Ascending colon	13%
C180 : Cecum	5%

Laterality *

0 : Not Paired 88%

Histology *

8140/3 : Adenocarcinoma, NOS	71%
8210/3 : Adenocarcinoma in adenomatous polyp	13%
8480/3 : Mucinous adenocarcinoma	7%

Behavior *

3 : Malignant Primary 97%

[Add Primary](#)

PATTERSON, JUDY C 330-65-4321 <unk sex> MRN: 54321-6
Unlinked b. 07-11-1936 Region: GA FAC-100000

Clinical Hist
Screening colonoscopy, Maternal hx of adenocarcinoma of colon age 57

Nature of Spec
2 cm polyp ascending colon 2 mm polyp in sigmoid colon

Gross Path
A. The first container is labeled "ascending colon." It contains a polypoid piece of tan mucosal tissue measuring 2.0 cm in greatest dimension. The polyp margin is inked, sectioned, and submitted in cassettes A1 and A2.B.
The second container is labeled "sigmoid colon." It contains one piece of light tan mucosal tissue 0.2 cm in greatest dimension. Entirely submitted in cassette B.

Micro Desc
Microscopic Examination performed supportive of the Final Diagnosis above.

Final DX
A. Ascending Colon
SESSILE SERRATED ADENOMA (POLYP) WITH LOW-GRADE ADENOMATOUS DYSPLASIA.
B. Sigmoid Colon
TUBULAR ADENOCARCINOMA

Supp Rpt Add

Staging

The polisher's results are shown in order of a relative confidence score. In this example, the polisher determined that the site might be C187, C182, or C180. The score shown next to each possible result does not indicate a probability. It indicates that the polisher had higher confidence in the first value over the second; and higher confidence in the second value over the third. One of these values may be right or they may all be wrong. A person needs to read the path report and code the site and morphology.

To code site, morphology, and reportability in a Path Screening task (with polisher results):

1. Review the pathology text.
2. If the report indicates cancer then code site, laterality, histology, and behavior:
 - a. If the correct value is listed in the polisher results:
 - i. Click the row with the correct value. For example, if C187 then click the row on the left that says C187: Sigmoid colon.
 - b. If the correct value is not listed:
 - i. Manually enter the correct value.
 - ii. Use your mouse to highlight text supporting the value that you entered. Detailed instructions for identifying supporting text are provided later in this chapter.
 - c. Leave the Reportability field set to "Based on coded values".
 - d. Click Save. SEER*DMS will automatically calculate reportability based on the site and morphology that you entered.

3. If this is not a reportable report:
 - a. Use the Reportability drop-down in the left panel and select Non-reportable.
 - b. Select a reason from the Non-reportable Reason drop-down.
 - c. Enter a comment.
 - d. Click Save.

Other features and tools in the Path Screening Task are described in the following sections:

- Highlighting Supporting Text
- Coding Reports with Multiple Primaries
- Using the Auto-complete Feature
- Display of Keywords and Dates

Highlighting Supporting Text

To highlight text using your mouse, position your cursor at the beginning of the text you want to highlight. Press and hold your primary mouse button (typically the left-button). While holding the mouse button, drag the cursor to the end of the text and let go of the mouse button. The screen shot below shows text being highlighted in the Final DX field.

Final DX

A small population of monoclonal B-cells (Kappa) is present in the bone marrow. The antigenic profile is consistent with chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL).

You will be prompted to select data items when you release the mouse. You may select one or more data items for the same text. In this example, the same phrase is used as supporting text for site and laterality.

Final DX

A small population of monoclonal B-cells (Kappa) is present in the bone marrow. The antigenic profile is consistent with chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL).

If the highlighted text supports the coding of fields listed below - please check the fields.

Use Ctrl+C or click [copy to clipboard](#) to copy the highlighted text.

Data Item	Comment
<input type="checkbox"/> Site <input type="checkbox"/> Laterality <input checked="" type="checkbox"/> Histology <input type="checkbox"/> Behavior	

OK **Cancel**

You may highlight multiple phrases for the same data item(s). The next screen shot shows that two phrases were highlighted for site and laterality. The number of highlighted phrases is shown next to each data item.

Coding Reports with Multiple Primaries

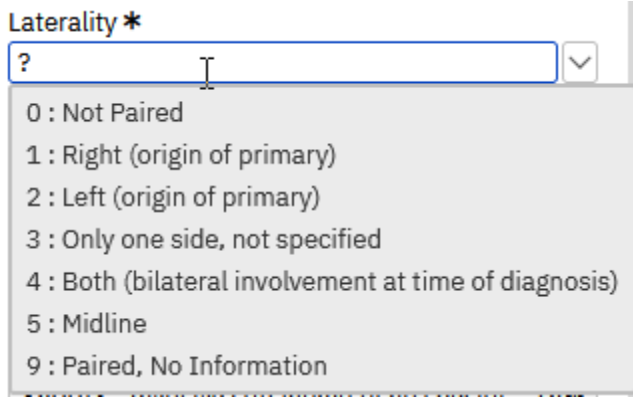
A small percentage of reports will contain information for two primary cancer diagnoses. Use the Add Primary link in the left panel to enter a second diagnosis. A maximum of five diagnoses can be entered for a report.

When you complete the task, SEER*DMS will create a separate record for each primary defined in the Path Screening task.

Using the Auto-complete Feature

You may use search text or a code to find the correct value for a field. For example, start typing "paget" or 8540 for histology. A list of possible values will be displayed. Select a value and the system will set histology and behavior based on that morphology.

Enter a question mark to see all values for a field. This screen shot shows the values for laterality.



Display of Keywords and Dates

SEER*DMS uses color-coded highlighting to emphasize dates, and significant words or phrases in text fields. A user-preference controls whether this option is turned on or off. To turn off keyword highlighting, click your username to open your account preferences. Enter "keyword" into the search box. Change Keyword Match Style to None.

Standard Screening Tasks

Requires system permission: *screening* and *rec_edit*

Records created via abstracting software are submitted to SEER*DMS with pre-coded values for site and morphology. SEER*DMS only creates a manual task to code these data items if they are missing or not valid.

The auto-screener sets the reportability flag for nearly all abstracted records; a manual Screening task is created for a very small percentage. A manual task is created if the auto-screener cannot determine reportability; or if, according to your registry's algorithms, non-reportable abstracts always require review. Some registries only expect to receive reportable abstracts and, therefore, require a manual review for any abstract with a non-reportable site and morphology.

To open a screening task, click a **Screening** link in the worklist summary on the home page. The Information column for each task will show a message from the auto-screener. For example, the auto-screener requires a valid value for site. If site was missing or blank then the auto-screener message would be "missing site". A similar message would be listed if the record failed other Auto-Screener Rules. The auto-screener rules and algorithms are described on the Screening help page.

Once you open the task, the record will be displayed in the record editor. Messages from the auto-screener will be displayed in the **Alerts** tab in the right panel. In this example, the screener could not determine reportability because the record does not have values for behavior or histology.

The screenshot shows a software interface titled "Supporting Data". At the top, there are three tabs: "Alerts" (with a red badge containing the number "1"), "Text", and "Edits". Below the tabs, the main heading is "Review Data and Set Reportability" with an information icon and a "(1)" indicator. A light blue box highlights the section "Auto-screening Issues", which contains the text "Behavior or Histology is missing." Below this, a message states: "Auto-screener could not determine reportability. Review the data and address the issues listed above. You must manually set the reportability flag." At the bottom of the interface is a prominent blue button labeled "Set Reportability".

To set a record's reportability flag in a standard Screening task:

1. Review the auto-screener messages displayed on the **Alerts** tab.
2. Set values on the CTC page for missing critical fields (site, histology, behavior, etc).
3. Review the record to ascertain reportability:
 - a. Click the links in the left navigation panel to review fields on other data pages.
 - b. If text is not displayed on the main data panel, click **Text** to open a pop-up window showing the text fields or go to the Text tab in the right panel.
4. If you can ascertain reportability for this record, set the reportability status field:
 - a. Click **Set Reportability** on the Alerts tab.
 - b. Select a **Reportability** value from the drop-down list.
 - c. If you have selected *Auditable* or *Non-Reportable*, document your findings by selecting a coded value in the **Non-Rpt Reason** field and entering text for the **Non-Rpt Comment**.
5. If you are unable to determine reportability, use one of these options.
 - a. If further information from the abstractor or physician is required, you may submit a request for follow-back information.
 - b. If you require assistance, you may reroute the screening task to your manager or a colleague by following the instructions in *Chapter 4: Using the Worklist*.
6. Click **Save**.
7. Click **Save & Exit**.