**Final Melanoma Annotation Guidelines**

**(03/13/2024)**

**General Annotation Guidelines**

* There are 21 classes of interest for identifying features pertinent to melanoma.
* Don’t mark the text within the document title.
* Don’t interpret clinical data to decide whether to mark a class. Use the author’s assertions to make the decision.
* You may look up terms found in the notes as you are annotating to ascertain whether they belong to one of the classes. We suggest this site as a good reference [melanoma - UpToDate](https://www.uptodate.com/contents/search?search=melanoma&sp=0&searchType=PLAIN_TEXT&source=USER_INPUT&searchControl=TOP_PULLDOWN&searchOffset=1&autoComplete=false&language=&max=0&index=&autoCompleteTerm=&rawSentence=)

**Span Selection Guidelines**

* “Spans” refer to a contiguous set of tokens within a sentence, where “tokens” are words and other meaningful character combinations (e.g., units of measurement and acronyms).
* **\*\*\*Discontinuous spans are discouraged in this task**. Annotate phrase as 1 or 2 separate spans.
* Spans selected for annotation should be restricted to the smallest number of tokens needed to represent the class in question.

**Attribute Assignment Guidelines**

* Each class has one or more attributes, also requiring determination.
* Keep in mind that you are looking for assertions by the author of the note. It is important that you distinguish your own inferences from those of the author,
  + Don’t mark “uncertain” for instances of text that you are uncertain whether they meet the criteria.
    - Text within the note such as “possible hypertension” is an example of expressed uncertainty.

**Assertion Status (Attribute)**

For each class, be careful to note whether the author of the note is asserting it as being present or negated or some qualifying uncertainty is expressed.

“Positive” is the default value.

If the class of interest is affirmed, or expressed as present, select “Positive” or “Yes”

If the class of interest is negated, or expressed as being absent select “Negative” or “Absent”

If there is uncertainty expressed by note author in relation to the class, select “Unable to determine’

**Relationships Between Concepts-** annotators should markeach concept to the correct (parent) cancer diagnosis.This is because some reports include multiple specimens or differing diagnoses.Specimens are also frequently sent out for biomarker testing or further review with all results noted on a single pathology report.

\*\*\*\* \*\*\*\*\*\*\*\* \*\*\*\*\*\* \*\*\*\*\*\*\* Pathology Classes \*\*\*\*\*\*\*\* \*\*\*\*\*\*\*\* \*\*\*\*\*\*

1. **0\_NOTE\_REVIEWED- Use this class only when you encounter a note lacking any relevant information. This will let NLP know the annotator didn’t inadvertently skip over note.** 
   1. **Attribute: Reviewed**

**Possible value:** Yes; Unable to complete due to missing/incomplete data (Default)

1. **1\_CANCER\_DIAGNOSIS:** mark text referring to diagnosis or evidence pathologist sent slides out for additional review
2. **Attribute: Topography**

**Possible values:** Bladder; Blood; Bone; Brain; Breast; Colorectal; GI; Liver; Lung; Lymph node; No mention; Other; Prostate; Skin

1. **Attribute: Topography detail**

**Possible values:** Axillary; Back torso; Foot; Front torso; Genitalia; Hand; Head (scalp, face, ears); Internal organ; Lower limb; Neck; Other; Shoulder; Unable to determine or missing; Upper limb

1. **Attribute: Laterality**

**Possible values:** Bilateral; Left; Midline; Right; Unable to determine, missing, or N/A

1. **Attribute: Status (+)**

**Possible values:** Definite (Default); Negation; Probable; Needs further review

1. **Attribute: Temporality (+)**

**Possible values:** Current (Default); Past

1. **Attribute: Experiencer (+)**

**Possible values:** Family; Patient

1. **Attribute: Result (If benign or no residual, you are done annotating).**

**Possible values:** Benign (NOT cancer); Malignant; Negative for residual cancer; Unable to determine or missing

**Examples:**

|  |  |
| --- | --- |
| **Attribute: Topography** | Skin |
| **Attribute: Topography detail** | Upper limb |
| **Attribute: Laterality** | Right |
| **Attribute: Status (+)** | Definite (Default) |
| **Attribute: Temporality (+)** | Current (Default) |
| **Attribute: Experiencer (+)** | Patient |
| **Attribute: Result** | Malignant |

|  |  |  |
| --- | --- | --- |
| “Skin, left anterior chest” | **“**Skin, right upper back**”** | “Right neck sentinel lymph node” |
| “Right scalp melanoma” |  | “Sentinel lymph node left groin” |
| Synoptic reports can appear to concatenate. In example below, treat as 2 annotations of cancer diagnosis. Sometimes a report may have  **1st mention:** “Final diagnosis: skin, mid back, excision, invasive malignant melanoma. See protocol (AKA Synoptic format)”    **2nd mention:** Section header says, “Protocol for Melanoma of the Skin version 7.0” If patient **has melanoma,** then this header is specific to patient and **should** be annotated. If template appears and patient **negative melanoma, then ignore** because it’s not specific to the patient.  **Possible values:** Procedure: excision Specimen laterality: midline Tumor site: Back | “Skin lesion right axilla, excision, benign compound nevus”. | Mark ‘status negated’ and result =‘benign NOT cancer’. Ignore ‘benign compound nevus’ because it’s normal. |

1. **2\_MELANOMA\_HISTOLOGY:** Definition: Refers to the growth pattern. Melanomas are asymmetrical and poorly circumscribed lesions with architectural disturbance and usually marked cytological atypia. A relationship should be annotated from melanoma histology to cancer diagnosis.
   1. **Attribute: Type**

**Possible values:** Acral letiginous melanoma; Amelanotic melanoma; Choroidal melanoma; Desmoplastic melanoma; Lentigo maligna melanoma; Melanoma in situ; Mucosal melanoma; Nevoid melanoma; Nodular melanoma; Ocular melanoma; Other melanoma or NOS; Residual cancer noted; Spitzoid melanoma; Superficial spreading melanoma; Unable to determine/needs further review

|  |  |
| --- | --- |
| “malignant melanoma superficial spreading type” | “Residual invasive malignant melanoma” |
| “lentigo maligna” | “Malignant melanoma, not otherwise classified” |
| “melanoma in situ” | “Skin lesion right shoulder, wide excision for melanoma. Sample shows no residual malignant melanoma.”  Mark lesion as 1\_Cancer diagnosis- status negative; result- negative for residual cancer. Mark ‘melanoma’ instances under 2\_Melanoma histology. This captures melanoma was not present. |

* 1. **Attribute: Recommendations**

**Possible values:** Needs further review; No mention; Other; Reexcision recommended; Unable to determine

1. **3\_HISTOLOGY\_ALL\_OTHER\_CANCERS:** Annotate a relationship from histology other to cancer diagnosis.
2. **Attribute:**  **Type**

**Possible values:** Adenocarcinoma; Basal cell (BCC); Cancer\_NOS; Carcinoma other; Leukemia; Lymphoma; Merkel cell; Myeloma; Residual cancer noted; Other; Sarcoma; Squamous cell carcinoma (SCC); Unable to determine

Ex: “sarcomatoid malignant neoplasm, completely excised.”

Ex: “Basal cell (BCC) noted on left ear.

Ex: This specimen shows Bowen’s disease. Bowen’s is early squamous cell carcinoma in situ. Either SCC or other could be correct. There may be other synonyms we run across, and we need to think about a rule.

1. **Attribute: Recommendations**

**Possible values:** Needs further review; No mention; Other; Reexcision recommended; Unable to determine

1. **4\_NODE STATUS** Annotate a relationship from the node to the corresponding cancer diagnosis.
   1. **Attribute:** Number examined

**Possible values:** 0; 1; 2; 3; 4; Measurement (varies-free text); Not performed; Unable to determine or missing

* 1. **Attribute:** Number positive

**Possible values:** 0; 1; 2; 3; 4; Measurement (varies-free text); Not performed; Unable to determine or missing

**Attribute:** Sentinel node evaluation

**Possible values:** Not performed; NOT sentinel; Unable to determine or missing; YES Sentinel

Ex: “Lymph node status: N/A”

|  |  |
| --- | --- |
| Number examined: | Not performed |
| Number positive: | Not performed |
| Sentinel node evaluation: | Not performed |

Ex: “Lymph nodes: number of Sentinel nodes examined: 1

Number of nodes with metastasis: 0”

|  |  |
| --- | --- |
| Number examined: | 1 |
| Number positive: | 0 |
| Sentinel node evaluation: | YES Sentinel |

1. **MELANOMA\_ULCERATION-** A relationship should be annotated from ulceration to cancer diagnosis.
   1. **Attribute: Presence**

**Possible values:** Absent; Yes present; Unable to determine or missing

|  |  |
| --- | --- |
| “Ulceration: not identified” |  |

1. **MELANOMA\_CLARK\_LEVEL\_OF\_INVASION:** Definition- morphologic finding indicating the cutaneous melanoma (**only applies to melanoma**).

Annotate a relationship from Clark level to cancer diagnosis.

a. **Attribute: Clark level of invasion**

**Possible values:** Level 1; Level 2; Level 3; Level 4; Level 5; Unable to determine or missing

|  |  |
| --- | --- |
| **“**Anatomic Level: V**”** | “Clark level: 4 of 5” |

1. **MELANOMA\_BRESLOW\_THICKNESS:** Definition: Maximum tumor thickness as measured from the deepest tumor edge to the overlying granular layer. Annotate a relationship from Breslow to cancer diagnosis.
2. **Attribute:** **Thickness (Breslow)**

**Possible values:** Measurement (free text); Unable to determine or missing

|  |  |
| --- | --- |
| “Tumor thickness: at least 3.5mm in the specimen”mit | “Breslow depth at .82mm” |

1. **TNM\_AND\_STAGE:** Annotate a relationship from TNM to cancer diagnosis. Mentions of the tumor stage, as well as mentions like “early” stage; should be annotated here, for example, “Stage IV’. If the stage is mentioned as a range, for example ‘Stage I-IIA’, then annotate entire span: ‘Stage I-IIA’
2. **Attribute: Stage**

**Possible values:** Stage 0; Stage 1; Stage 2; Stage 3; Stage 4; Not specified; Unable to determine

1. **Attribute: Stage type (+):**

**Possible values:** Pathologic (Default); Clinical; Not specified; Unable to determine

1. **Attribute: Tumor (T)**

**Possible values:** T0; T1; T2; T3; T4; TX; Not specified; Unable to determine

1. **Attribute: Tumor (T) invasion detail**

**Possible values:** A; B; C; Not specified; Unable to determine

1. **Attribute: Nodes (N)**

**Possible values:** NX; N0; N1; N2; N3; N4; Not specified; Unable to determine

1. **Attribute: Nodes (N) detail**

**Possible values:** A; B; C; Not specified; Unable to determine

1. **Attribute:** Metastasis (M)

**Possible values:** MX; M1; Not specified; Unable to determine

1. **Attribute: Recurrent (R)**

**Possible values:** Yes; Not recurrent (R0); Not specified; Unable to determine

1. **Attribute:** stage secondary ABC

**Possible values:** A; B; C; Not specified; Unable to determine

Example: pT1a, pNX, pMX, R0 (see below)

|  |  |
| --- | --- |
| Stage | Not specified |
| Stage type | Pathologic |
| Tumor (T) | T1 |
| Tumor (T) invasion detail | A |
| Nodes (N) | NX |
| Nodes (N) detail | Not specified |
| Metastasis (M) | MX |
| Metastasis stage secondary ABC | Not specified |
| Recurrent (R) | Not recurrent (R0) |

\*\*\* or you may see a slightly different format without mention of mets (i.e. rpT4a, rpN0)

1. **GRADE:** Definition: Describes how abnormal the cancer cells look under a microscope when compared to healthy cells. This includes both numeric values such as ‘Grade 1’, as well as textual mentions such as ‘low’ in ‘low grade’, ‘intermediate’ in ‘intermediate grade’, etc. Annotate a relationship from grade to cancer diagnosis.
   1. **Attribute: Grade**  **Possible values:** High grade; Low grade; Moderately differentiated; Other; Poorly differentiated; No mention; Unable to determine; Well differentiated

1. **MARGINS- Annotate a relationship from margin to corresponding diagnosis. Capture all margins (whether melanoma, benign, or another type of cancer from melanoma. A report may have specific margins indicating melanoma and/or melanoma-in-situ. We want to capture them separately.** 
   1. **Attribute: Presence**

**Possible values:** Clear margins; Involved margins; Unable to determine or missing; Very close to margin, Transected margins involved

* 1. **Attribute: Distance**

**Possible values:** Measurement (free text); Not specified

|  |  |
| --- | --- |
| “Margins (peripheral): Uninvolved by invasive melanoma with closest margin 1.4cm” | Surgical margins: involved |
| “Surgical margins free of involvement; closest deep margin 1.5mm.” | Peripheral and deep margins from excision are negative |
| “Melanoma in-situ closest margin 2cm, malignant melanoma closest margin 1cm” Note the 2 separate annotations with relationships made to corresponding cancer diagnoses. | Lesion was completely excised from the scalp. Per SME: “the ‘COMPLETELY EXCISED’ phrase indicates the margins are clear.” |

1. **SURGICAL\_PROCEDURE\_BX-** Annotate a relationship from surgical procedure or bx type to cancer diagnosis.
   1. **Attribute: Type**

**Possible values:** Core biopsy; Excisional biopsy; Incisional biopsy; Lymph node dissection (Complete, elective/ therapeutic); Needle biopsy/fine needle biopsy (FNA); Other; Punch biopsy; Re-excision; or Shave biopsy

* 1. **Attribute: Timing**

**Possible values:** Current; Past; Unable to determine or missing

**Ex: An adjacent scar from a prior biopsy is noted.** (mark prior biopsy and **past**)

**Ex: Prior fine needle biopsy tract identified.** (mark prior biopsy and past)

**Ex:** **Lesion was completely excised from the chest.** \*\*\*Per SME- “annotate the entire span ‘COMPLETELY EXCISED’ because some notes might say partially excised or have other modifiers.”

**Ex: Tissue is labeled to be from wide excision of scalp.**

**Ex:** Note: **Overlap of 1\_cancer diagnosis and procedure/bx** may result in double annotation: ex: ‘tissue sample is skin; wide excision; left ear’ . Annotate entire phrase for 1\_cancer diagnosis and ‘wide excision’ for procedure/bx.

**Ex:** ‘complete excision’- overlap also exists in this example: double annotate as surgical procedure/bx and margins clear per SME.

1. **METASTASIS** Annotate a relationship from metastasis to cancer diagnosis.
   1. **Attribute: Presence**

**Possible values:**  Unable to determine or missing; Not present; Yes present

* 1. **Attribute: Suggested primary per pathologist- (use this attribute when primary is unknown, and specimen tested represents metastasis of an unknown primary).**

**Possible values:** Adrenals; Bone; Brain; Breast; Colorectal; GYN; Liver; Lung; Melanoma; Multiple possibilities; No mention; Other; Other lung; Peritoneum; Prostate; Unable to determine or missing

* 1. **Attribute: Site #1 of metastasis (use this attribute when primary is known, and specimen tested represents metastasis of this known primary).**

**Possible values:** Bone mets; Brain mets; Liver mets; Lung mets; Lymph node mets; Other mets; Not specified.

**Ex: “One lymph node with metastatic melanoma.”** Mark present, lymph node mets, and melanoma histology.

* 1. **Attribute: Site #2 of metastasis**

**Possible values:** Bone mets; Brain mets; Liver mets; Lung mets; Lymph node mets; Not specified; Other mets

**Ex: “Focal microscopic metastatic melanoma identified.”-separate ‘metastatic’ from ‘melanoma’ in this example.**

1. **MITOTIC\_INDEX:** Definition: Number of mitoses per square mm in the tumor. Annotate a relationship from mitotic index to cancer diagnosis.
   1. **Attribute:** Number of mitoses per mm2 tumor

**Possible values:** 0 None; Less than 1 per mm2; 1-4 per mm2; Measurement; Other; Unable to determine or missing

|  |  |
| --- | --- |
| “Mitotic index: Less than 1/mm2” |  |

1. **PERINEURAL\_INVASION**  Drawa relationship from perineural to the corresponding cancer diagnosis.
   1. **Attribute:**  Presence

**Possible values:** No; Unable to determine or missing; Yes

1. **BIOMARKERS\_TESTED-** Definition: A measurable substance in an organism whose presence is indicative of some phenomenon, disease, infection, or environmental exposure. BRAF mutations occur in up to 50% of melanomas. Of these mutations, 95% occur at amino acid 600, most commonly as Val600Glu (V600E) or sometimes Val600Lys (V600K). Any biomarker should be annotated. **For the sake of generalization, our definition of a biomarker is very loose**. For example, we would include ‘ER’, ‘PR’ as well as ‘HER-2’ as biomarker names for breast cancer. Annotate a relationship from the biomarker to the corresponding cancer diagnosis.
   1. **Attribute: Test**

**Possible values:** APC; ATM; BRAF; BRCA1; BRCA2; EGFR; HER2; KIT; KRAS; Estrogen receptor; Progesterone receptor; OTHER; PIK3CA; PTEN; TP 53 (p53)

* 1. **Attribute: Result**

**Possible values:** Negative; Pending; Positive; Unable to determine or missing

Ex: “BRAF mutation testing is pending”

|  |  |  |
| --- | --- | --- |
| Biomarkers Tested | Test | BRAF |
|  | Result | Pending |

1. **TUMOR\_INFILTRATING\_LYMPHOCYTES-** Definition- Presence of inflammation in the primary. Annotate a relationship from TIL to the corresponding cancer diagnosis.
   1. **Attribute: Presence**

**Possible values:** Absent; Infiltrative pattern noted; No mention; Present and brisk; Present NOT brisk; Present\_NOS; Unable to determine

Ex: “Tumor infiltrating lymphocytes: Present, non-brisk”

Ex: “lymphoid response at base is moderate” (per SME).

1. **TUMOR\_SIZE-** Definition: Surface area of the cancer, usually described in the examination section of the clinical note or in the biopsy report (clinical description or gross description sections). Typically, size in square cm (or mm). Only annotate specific mention of cancer. The size of the tumor, along with the corresponding units of measurement, should be annotated. A relationship should be annotated from tumor size to corresponding cancer diagnosis.

Ex: ‘7x9x5 mm’ should be annotated as a single entity.

Ex: **Do not annotate “**no obvious masses or tumor” since this is not specific enough**.**

* 1. **Attribute: Tumor or lesion size**

**Possible values:** Measurement (Free text); Cannot be determined

Ex: “Tumor size: 2.4 X 2.2 cm” grossly; “greatest dimension: 1.9 cm” (annotate as 2 separate measurements since measurements don’t repeat.

Ex: Size of tumor is 2.5 X 3.2 cm with greatest dimension 3.2 cm (annotate as 1 phrase since measurement is repeated.

Ex: Macroscopic Tumor and Tumor Size: 4cm length scar with a vague 0.7 X 0.6cm brown region adjacent to scar

1. **AT\_NEVUS-** Annotate mentions of nevus **with cancer** **only (or negated-see example)**. Per SME: “we don't care about path reports that don't have malignancy.” Mark a relationship from nevus to corresponding diagnosis.
   1. **Attribute:** Presence

**Possible values:** Yes- at nevus; No; Unable to determine or missing

Ex: “A precursor nevus is not identified.”

\*\*\*Additional nevus guidance from SME with example of nevus mention in section B2.

Question: “If we interpret our rules strictly, we will not annotate this nevus because it is negative.  But it’s a little unclear since the SC node is positive and seems to be located close to this nevus from the pathologist’s dictation.  Would you recommend annotating the nevus using the ‘at nevus’ class vs. ignoring the nevus?

Per SME: “Overall I agree with your analysis to interpret the rules strictly. Here are the rules I would use for annotation:

1. If the nevus is in the same specimen as a melanoma, annotate it. We would be particularly interested in this if there were phrases like "melanoma associated with a nevus" or "melanoma directly adjacent to a nevus" etc. I would probably annotate the whole phrase so that we can capture things like "associated with" so that we can distinguish that nevus from others that are just totally separate in later analysis.
2. If the nevus is not in the same specimen as a melanoma, but the specimen is a lymph node biopsy then annotate it. This would be something like "lymph node negative for melanoma, but with a melanocytic nevus"
3. If the nevus is not in the same specimen as a melanoma and it is not a lymph node biopsy, then don't annotate it.

The goal would be to catch two things: a.) a nevus on the skin that is directly associated with a melanoma on the skin and b.) a nevus in the lymph node from a melanoma sentinel node biopsy. In this case, there isn't really a need to annotate this nevus because it is not associated with a melanoma on the skin, and it is not actually in the lymph node.

1. **Lymphovascular\_invasion-** Lymphovascular invasion (LVI) indicates whether microscopic lymphatic or vascular invasion is identified and includes lymphatic invasion, blood vessel invasion, or invasion of a vessel for which it cannot be determined whether it is a blood or lymphatic vessel. A relationship should be annotated from lymphovascular invasion to corresponding cancer diagnosis.
   1. **Attribute: Presence**

**Possible values:** Yes; No; Unable to determine or missing

1. **Regression-**a phenomenon that occurs because of patient’s host immune response attacks the primary tumor. Annotate mentions of regression and draw relationship to diagnosis sample.
   1. **Attribute: Presence**

**Possible values:** Present; Absent; Unable to determine or missing