**Physical Exam HPO Annotation Guide**

The purpose of this exercise is to unambiguously annotate physical exam findings written by clinicians in progress / consults notes. This will allow us to train an algorithm to do the task for us automatically.

The annotation task has 4 distinct parts.

1. Process the suggestions made by the previously published system, PhenoTagger
2. Add new abnormal findings that have been missed
3. Add specifically mentioned negative findings, if any
4. Annotate the specific regions (spans) in the text that contain the findings

Graphical user interface, text, application

Description automatically generated

**Annotations Should Be Unambiguously Described in the Text**

When making annotations, it is critical that you only consider the information that is actually available **unambiguously** considering only the organ system and the text as it is written.

For example, consider the following:

HANDS FEET

A picture containing application

Description automatically generated

2,3 syndactyly, no polydactyly

Although in your clinical judgement, you realize this is *very likely* referring to toe syndactyly, that information is not unambiguously present. In this case, the abnormal finding can only be coded as “HP:0001159 syndactyly” because there is no code for 2,3 syndactyly, only HP:0004691 2,3 toe syndactyly or HP:0001233 2,3 finger syndactyly.

**Step 1: Process PhenoTagger Suggestions**

The first step is to process the suggestions presented at the top of the page. The preferred term name “Description” is presented along with all of the synonyms listed for that term.

Graphical user interface, text, application, Teams

Description automatically generated

For each suggestion, decide which of the following is correct.

* Accept – this term is most specific that can be determined unambiguously only from the information available from the organ system and text itself.
* Reject – this term is technically correct, but an unambiguous more specific term is available
* Erroneous – this term is either totally wrong or can’t be determined unambiguously from the information provided
* Negate – this term is unambiguously not present based on the text and is specifically mentioned

When you click a button, the following actions occur:

* Accept – the decision changes to “Accept” and the term is added to the correct term list below
* Reject – the decision changes to “Reject”
* Erroneous – this is the default; nothing changes unless you previously selected another action. Note that if you previously clicked accept, you will likely need to remove the term from the correct term list manually.
* Negate – the decision changes to “Negated”, the term is added to the correct term list below but with an X in the Negated column.

Graphical user interface, text, application, email

Description automatically generated

Here is an example of a negated suggestion. “No redundant neck folds” should be annotated as Negated HP:0005989.

Graphical user interface, application

Description automatically generated

In this example, the text does not support microcephaly, so the decision is Erroneous. Small anterior fontanelle is supported by the text, so the action is Accept. Note that it is not your job to determine if the small anterior fontanelle is an abnormal finding (ie because it was noted on the delivery day) or a simple normal observation. The text does not support wide anterior fontanelle, so the decision is Erroneous.

Graphical user interface, text, application

Description automatically generated

A small down red arrow appears when a term has a subclass present on the list. This can alert you that the correct action *may* be “Reject”. Carefully consider if the subclass is unambiguous.

**Step 2: Add New Abnormal Findings**

Once the suggestions are processed, think about new abnormal findings that are delineated in the text.

Graphical user interface, application

Description automatically generated

It may be challenging to find the appropriate term. For example, we need to find persistence of the fetal fingerpads. The search box only finds terms based on the preferred term name. Synonyms cannot be searched from the box. Sometimes it may be necessary to use an alternative HPO tool, like the one found on <https://hpo.jax.org>.

Graphical user interface, application

Description automatically generated

This will help you find the appropriate text to search for.

Graphical user interface, text, application, email

Description automatically generated

Note that partial text matching works in the search box:

Graphical user interface, text, application, email

Description automatically generated

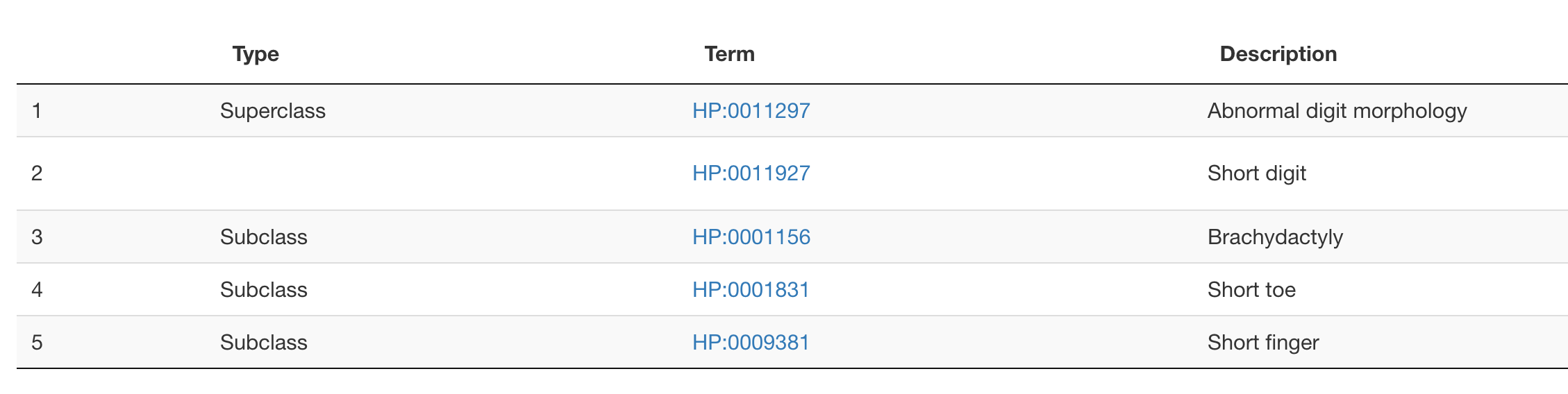
Graphical user interface, application

Description automatically generated

Sometimes it can be helpful to quickly traverse the HPO graph. To do so, click “Details” next to any term in the list.



To move up in the hierarchy, click the hyperlink on the HPO term.



Sometimes the HPO graph is not set up ideally. In this case, I would select “HP:0009381 Short finger”.

A quick approach to finding a term you aren’t sure of, is often to select a more general term. For example, if you were looking for “peaked cupid’s bow”, you might search for “abnormal upper lip” and quickly move through the hierarch:

Table

Description automatically generated

**Step 3: Add specifically mentioned negative findings**

When a clinician documents that something is not present, it not only provides information about excluded phenotypes, it may also give insight into their differential diagnosis.

Graphical user interface, application

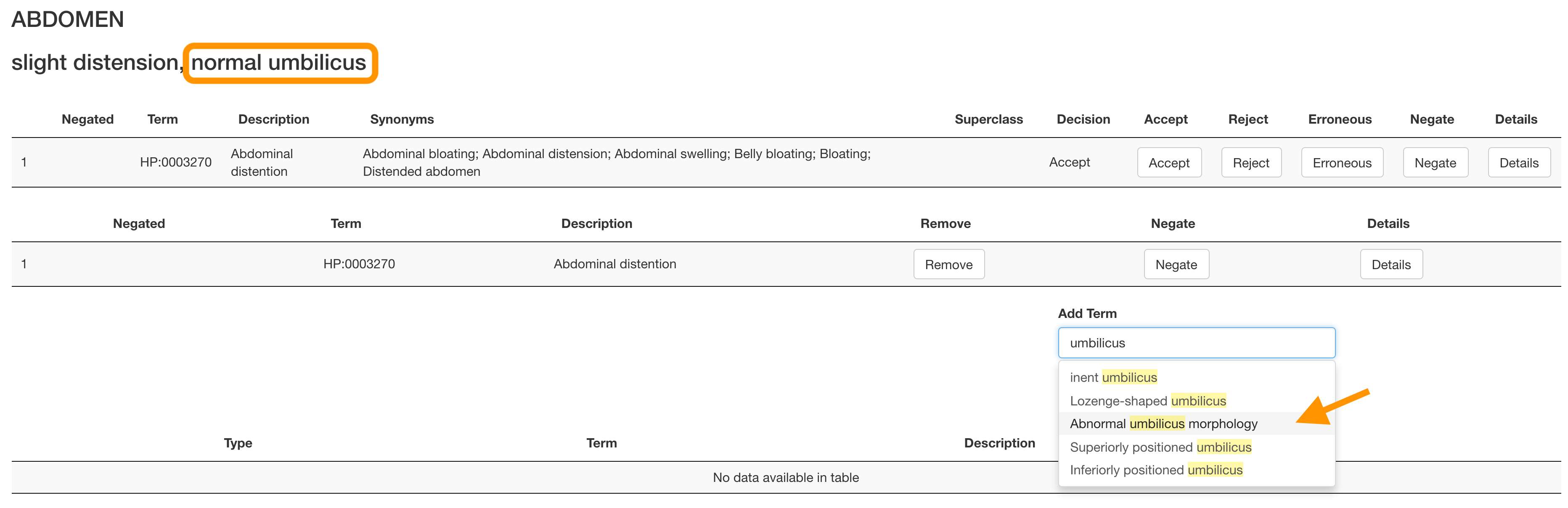
Description automatically generated

Carefully consider what negative findings are specifically mentioned. Once you have identified a concept:

1. Search for the appropriate unambiguous finding in the search box
2. Click to add the appropriate finding to the list
3. Click the Negate button
4. You will notice an X appears in the Negated column

Note that if you mistakenly negate a term, you can press the “Negate” button again.

Sometimes, clinicians specifically document a general normal finding. In these cases, try to find a general term about that anatomical structure. These are often “Abnormal \*\*\* morphology” or “Abnormality of \*\*\*”



Once you have found an appropriately scoped term, add it to the list and click “Negate”.

**Step 4: Annotate the text spans containing the terms**

Table

Description automatically generated

The purpose of this step is help our algorithm identify what part of the text contains each term.

To facilitate this task, the user interface has a set of two copies of the text. The top copy will become highlighted with the color matching the term.

1. User your cursor to highlight **the bottom copy** with the section of text you want to annotate. To avoid any bugs, **select from left to right**
2. Without deselecting the text in the bottom copy, click the Add span button for the appropriate term.
3. You will see the number indices of the characters within the text in the “Spans” column.
4. If the term is described in two (or more) discontinuous spans, highlight a second span and click the “Add Span” button. Repeat as necessary.
5. Now repeat the process for each HPO term. Note that you can use the same text for more than one term. This may bug out the highlighting, but the numeric indices should be correct.

Do not worry about qualifiers like “slightly” or “potentially”. Annotate all mentioned phenotypes are present or absent.

If a specific term is mentioned twice (IE on the left and then again on the right), add both spans.

If a term uses information mentioned both in the system (IE before the :) and in the main text, add the span that is most semantically close to the main text.

**Special Considerations**

HPO is mostly ambiguous on pluralization

For exam there are often no specific terms for only one low set ear or only one inverted nipple. In these cases, you can use the terms interchangeably.

However, occasionally, there are more specific terms. For example, cryptorchidism:

Graphical user interface, application

Description automatically generated

It can be helpful briefly check the graph if you are unsure. Remember to only choose the term that is unambiguously defined in the text.

Hooded Eyelids

Unfortunately, there are two overlapping terms for this concept. Hooded eyelid and dermatochalasis.

Graphical user interface, application

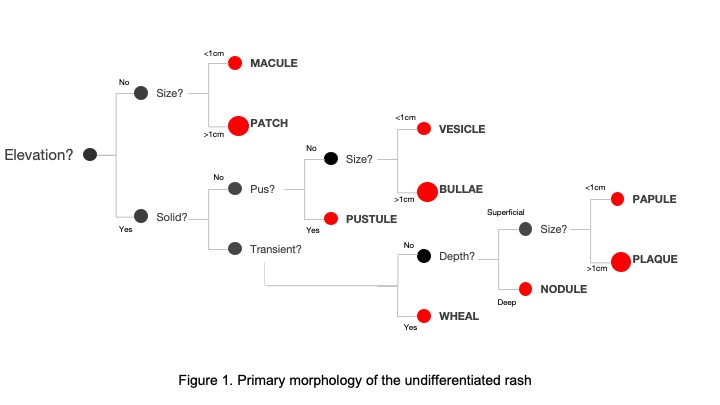
Description automatically generatedGraphical user interface, text, application

Description automatically generated

Please use HP:0030820 Hooded eyelid unless the clinician specifically uses the word dermatochalasis (or they mention *in the text* a patient characteristic that indicates this is due to aging or specifically due to skin elasticity).

Skin Findings

HPO is not well suited to skin findings. However, you can use this chart from [CanadiEM](https://canadiem.org/describing-a-rash/) to assist in lesion classification.



Unfortunately, there are not many “patch” annotations. You can use “patchy” if available or “macule” if otherwise. Clinicians often mention a color of the skin finding which cannot be matched in HPO. You can consider any dark color “hyperpigmentation”.