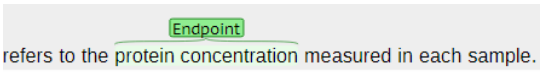
Annotation Guidelines for Experimental Methods and Endpoints Identification

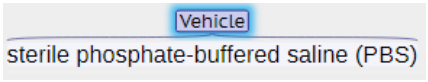
ANNOTATION INSTRUCTIONS

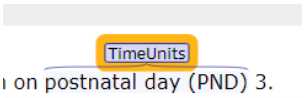
This guideline describes the specific types of information that should be annotated in a document to identify experimental methods and endpoints. Annotation of a document consists of labeling select text with a tag that indicates the type of thing or concept the text represents (as in the example below). The term mention is used throughout this document to indicate the text being labeled, e.g., the text 'protein concentration' is a mention of an experiment endpoint.



For annotation, each type of annotation listed in the section "Annotating Mentions" below should be tagged. A subsequent round of annotations will focus on indicating groups of concepts.

Unless noted in the G*uidelines for Annotation of Specific Mentions* section, annotate the longest contiguous text that describes the item of interest, including abbreviation definitions. Mentions generally should not cross sentence boundaries. In the examples below, the adjective 'sterile' and the abbreviation '(PBS)' should be included in the annotation as well as the abbreviation (PND) after postnatal day. Do not include citations at the end of relevant text; citations that appear within the relevant text should be included.

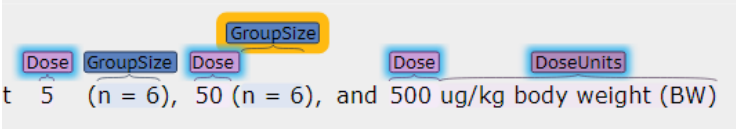




Additional general guidance:

- If an item (often the case for Species, TestArticles, and Endpoints) appears more than once in the text, annotate all instances, including the use of abbreviations, including section titles.

- For units (time and dose), annotate both abbreviations (e.g., ug/kg) and description of the units (e.g., (BW)) as in the example below.



ANNOTATING MENTIONS

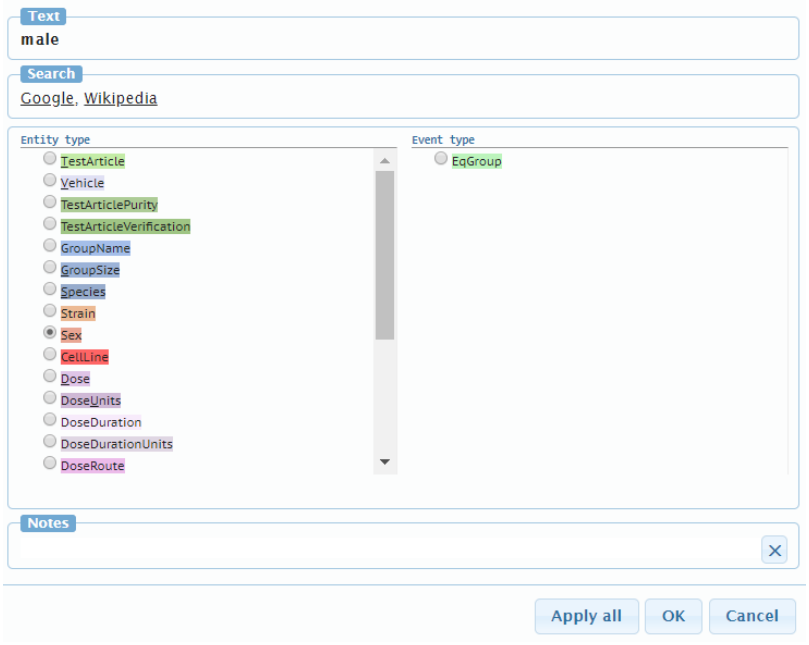
Table 1. Types of Tags

Category	Annotation Tag	Description
EXPOSURE	TestArticle	Test article or exposure evaluated
	Vehicle	The solution the test article is in
	TestArticlePurity	Purity of test article
	TestArticleVerification	Text indicating that the test article was confirmed, if present, typically just a statement saying the purity was confirmed by a third party
ANIMAL GROUP	GroupName	If reported, a name given to animal treatment groups (e.g., 'DES-10', 'treated') or control groups ('negative control', 'positive control').
	GroupSize	The number of animals in a group where a group is a set of animals given the same dosing regimen or used for an endpoint measurement.
	SampleSize	The number of animals used in an experiment
	Species	The species names
	Strain	The strain names
	Sex	Sex of the animal group(s)
	CellLine	The cell line name used in the experiment
DOSE GROUP	Dose	Dose
	DoseUnits	Units of dose
	DoseFrequency	Frequency at which doses are given
	DoseDuration	Duration of treatment (dose)
	DoseDurationUnits	Units of dose duration
	DoseRoute	Route of administration
	TimeAtDose	Time when dose is given (typically the age)
	TimeUnits	Units used for time (typically days)
	TimeAtFirstDose	Time at which first dose is given
	TimeAtLastDose	Time at which last dose is given
ENDPOINT	Endpoint	Endpoint evaluated
	EndpointUnitOfMeasure	Units of measured endpoint
	TimeEndpointAssessed	Time at which the endpoint was accessed (typically number of days after some event)

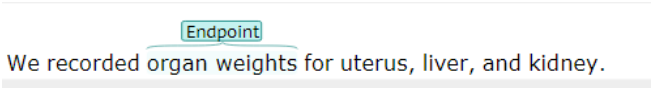
ANNOTATION FEATURES WITHIN BRAT

Apply All: Rather than individually tagging recurring information (ex. sex, or species), the “apply all” button allows you to tag each instance of a word to a category.

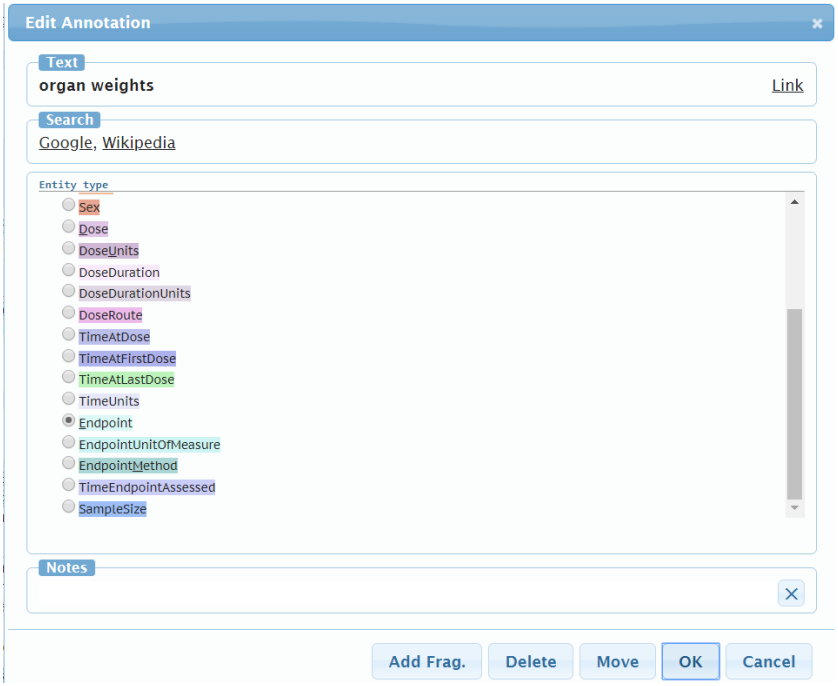
The apply all tag is listed to the left of the “OK” button. In this example each instance of the word “male” will be tagged as “sex”.



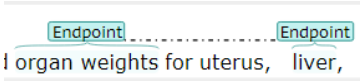
Combining Fragments: In some cases, a mention will consist of multiple words that are separated by irrelevant text. In this case, the separated text should be combined together using the “Add Frag” capability of Brat. As an example, in the figure below, there are three endpoints (uterus organ weight, liver organ weight, and kidney organ weight). To annotate liver organ weight, first, annotate “organ weight” as an endpoint.



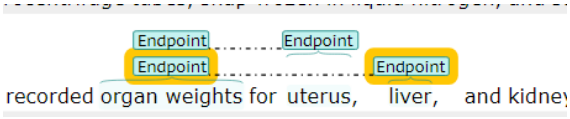
Then double click on the “organ weights” endpoint label to bring up the annotation dialog, shown below.



Click “Add Frag.”. The dialog will then disappear. If you then click “liver” the endpoint annotation will be extended to include liver, as shown below.



Note in this case, you would repeat the full process to annotate “uterus organ weight” with the results shown below.



Keyboard Shortcuts: Once you select or double click on text, the annotation dialog should appear. You can select the correct annotation by either clicking the appropriate Radio button or you can hit a keyboard shortcut. Entities with shortcuts have an underline under the character that serves as a shortcut. These are:

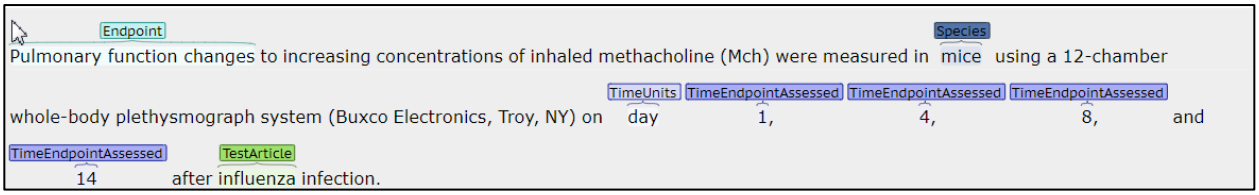
- D=Dose
- T=TestArticle
- V=Vehicle
- G=GroupSize
- S=Species
- U=DoseUnits
- E=Endpoint

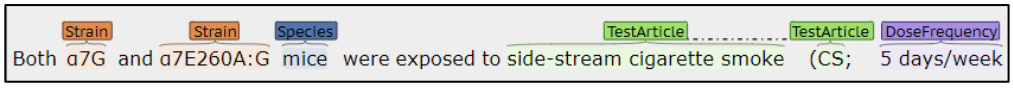
GUIDELINES FOR ANNOTATION OF SPECIFIC MENTIONS

1. Exposure Tags

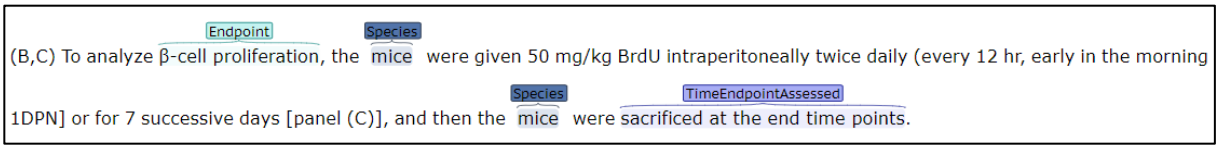
TestArticle:

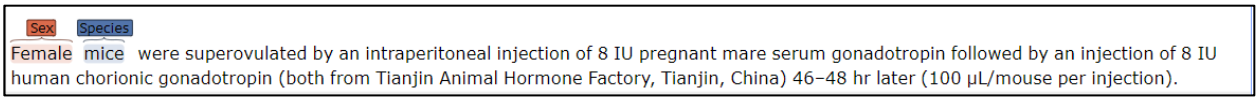
The test article is defined as the exposure (chemical or stressor) for which the experimental design is intended to evaluate. In many cases the test article will be mentioned using different forms, e.g., as a chemical name, as a CASRN number, and as a generic name for the substance. In these cases, annotate all instances. In experiments where a positive control is included, the positive control should be labeled as a TestArticle.



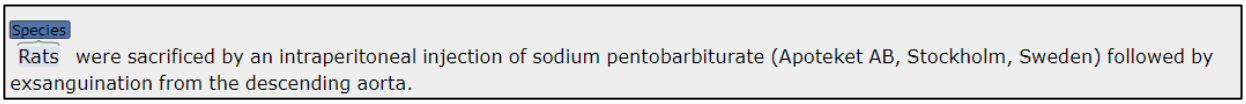


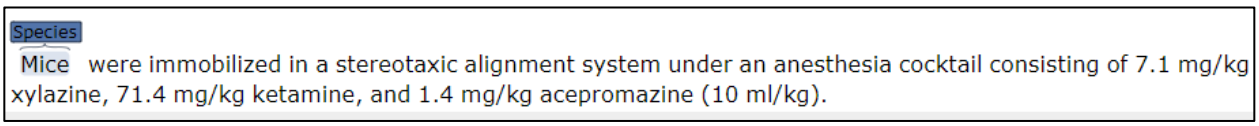
Reagents used for endpoint analysis (i.e. BrDu staining, glucose in tolerance test, insulin, methacholine challenge, superovulation etc.) are not annotated as test articles.





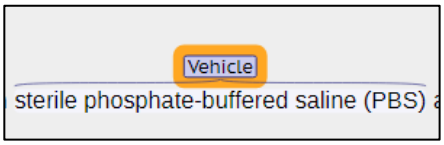
In addition, anesthesia reagents are not considered test articles, as shown in the examples below.





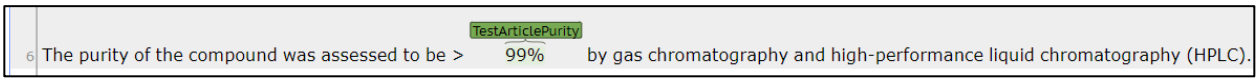
Vehicle:

Same as TestArticle. In the example below, the term 'sterile' is included as this further describes the vehicle. The abbreviation '(PBS)' is also included following the general rule of including abbreviations.



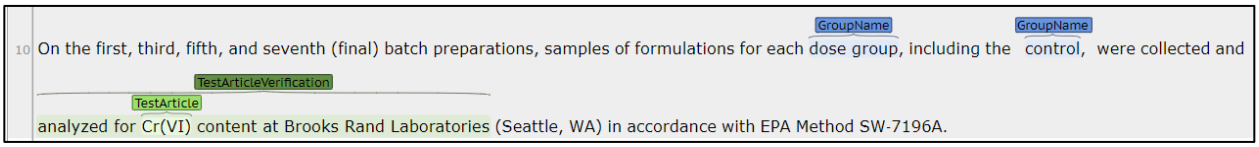
TestArticlePurity:

Annotate the purity level indicated, as well as any description of the method used and the group conducting the purity assessment.



TestArticleVerification:

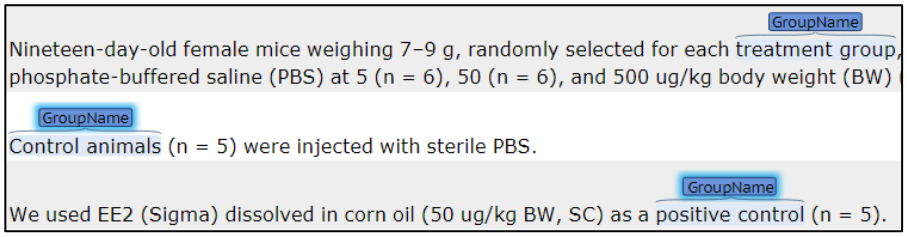
Annotate the statement which indicates that the chemical was confirmed. This may refer to a third party assessment where another company confirmed the chemical. Typically this may occur when the test article is sent off for chemical analysis before conducting the study. Only annotate the statement that the chemical was verified, no need to highlight the methods.

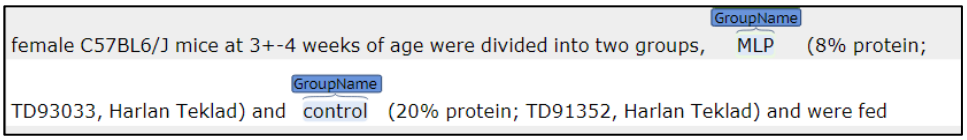


2. Animal Groups

GroupName:

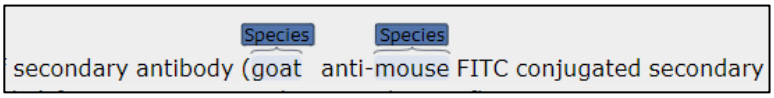
The group name is a label applied to one or more groups of animals that are given the test article or used as a control. This will often be a general name indicating one or more treatment groups as in the example below which has three treatment groups of sizes 6, 5, and 6, a control group, and a positive control group. The group name may also be a specific label describing an animal group (e.g., 'MLP' as in the second example) or a dose group (e.g., 500 ug/kg).





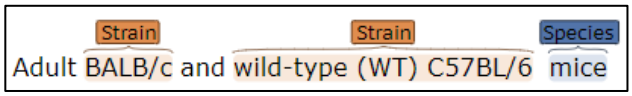
Species:

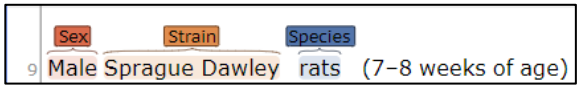
Annotate both scientific and common names for species. This includes all species listed within the text, including those not used in the experiment itself (e.g. goat anti-mouse in the example below). Do not include ambiguous names of animal classes (e.g. rodent or calf).



Strain:

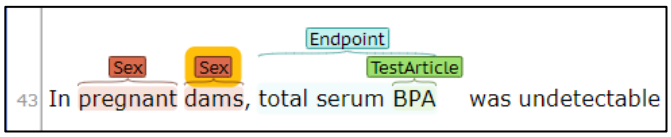
Annotate mentions of the species strain (including abbreviations) only. Annotate the full span of text relevant to determining what strain is being used. For example, in the below example the entire text describing the animal strain is highlighted.





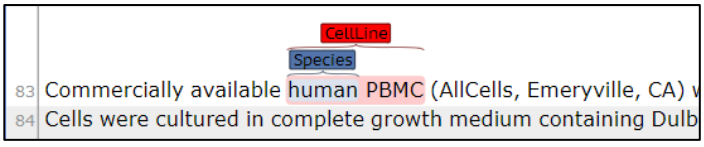
Sex:

Annotate specific mentions of the sex of animals. This includes direct mentions of male or female, related abbreviations (e.g., 'm','f') as well as sex-specific mentions of animals (e.g., dams, buck, doe, bull, cow). Also include mentions of pregnancy if the mention is directly related to an animal group(s). Pregnant females or pregnant dams should have two separate sex tags.



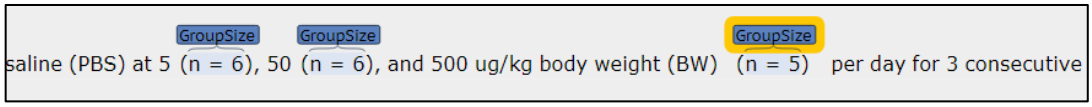
CellLine:

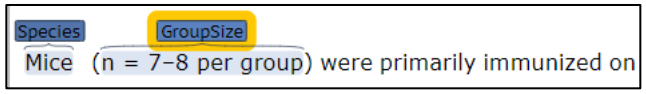
Annotate the cell line name described in the paper. A cell line would be an existing line of cells, not a primary culture. The species tag should be used to annotate the cell species. Then use the CellLine tag to annotate the cell line name.



SampleSize/GroupSize:

Annotate mentions of the number of animals being used in an experiment (sample size) and the number of animals specific to either a dose group or an endpoint measurement (group size). The first figure below gives examples of annotation of multiple group sizes, the second figure illustrates an example of sample size being used for an experiment and group sizes being used for each dose level. Include all information relevant to the group or sample size, e.g., in figures below the "n=" is kept as well as the "animals/end point", and "animals/dose group".

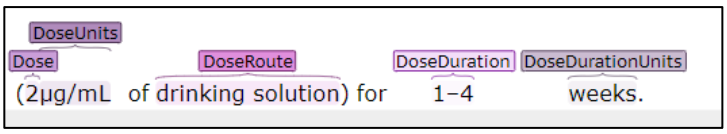




3. Dose Group Tags

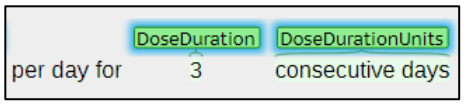
Dose/DoseUnit:

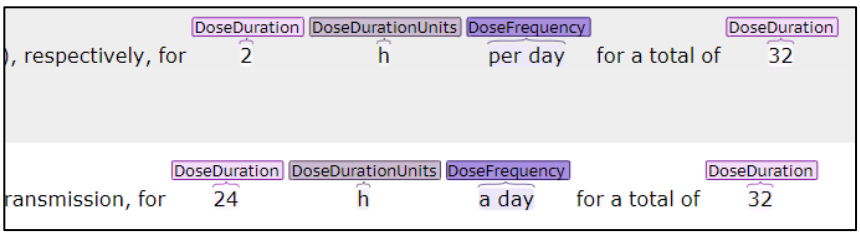
Annotate specific dose levels and dose units, example below. Only annotate doses and their respective dose units associated with a test article. Doses associated with an endpoint analysis or anesthesia should not be annotated.



DoseDuration/DoseDurationUnits:

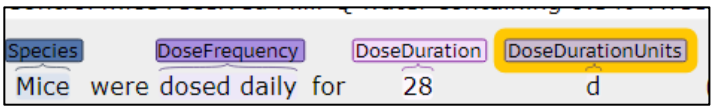
Authors often describe dosing in terms of a number of days or hours over which a treatment is given, example below. The number should be tagged as the dose duration and the explanatory text and/or units should be tagged as DoseDurationUnits.

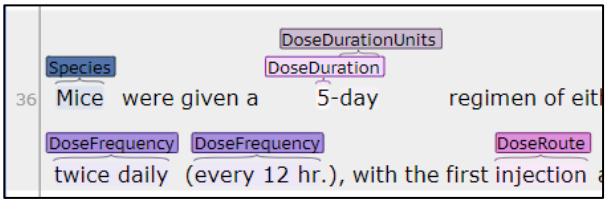




DoseFrequency:

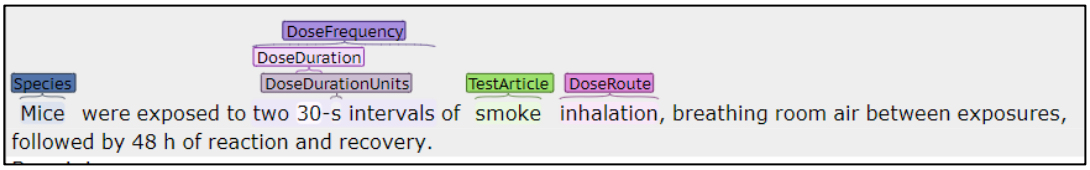
Authors often describe dosing in terms of the frequency at which doses are applied. The frequency may also include text describing the intervals between dosing. Note that dose duration and dose duration units are also frequently given and these should be separately annotated. For studies where the dose route was via diet or drinking water, “ad libitum” can be considered a dose frequency.



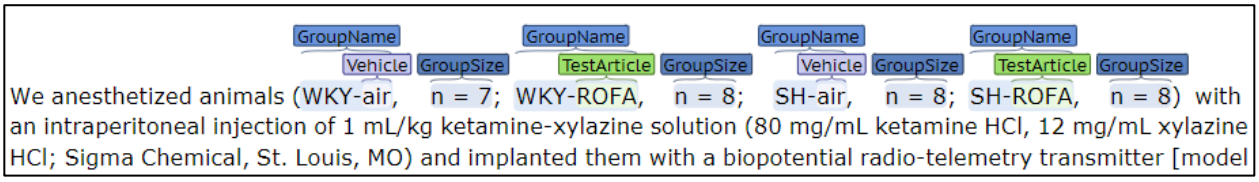


DoseRoute:

The method used to expose the animal to the test article. Examples include, inhalation, exposure chambers, oral, gavage, diet, drinking water, and injection. Routes for administering chemicals for endpoint analyses, anesthesia, or non-test article chemicals should not be tagged.

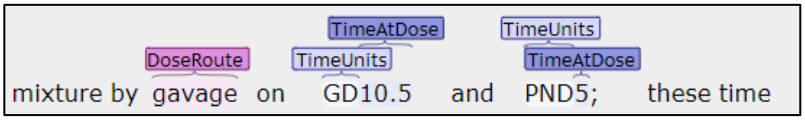


Note in the example below, “intraperitoneal injection’ is not tagged, since it was the administration route of the anesthesia, not the test article.



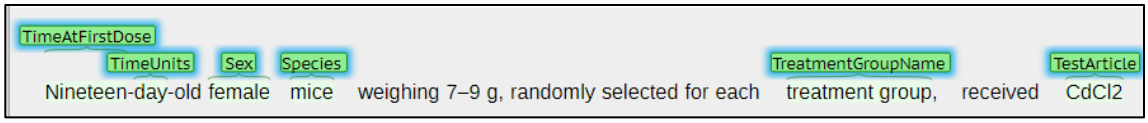
TimeAtDose/TimeUnits:

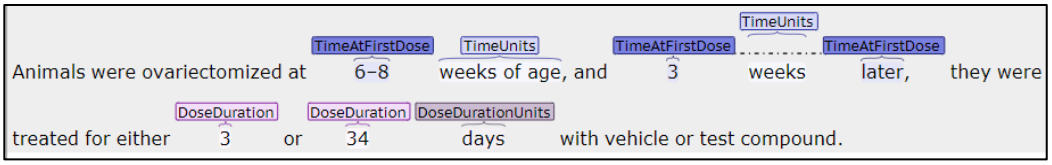
Annotate time at which doses are given and the units, example below.

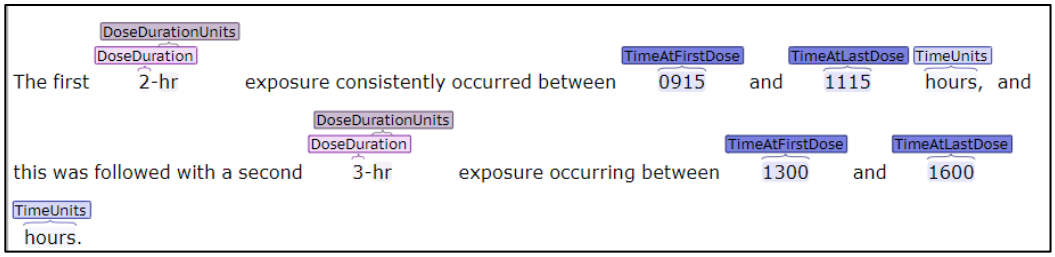


TimeAtFirstDose, TimeAtLastDose/TimeUnits:

Often the time at each dose is not given, instead the author provides the time for the first dose and information about the subsequent frequency of dosing, such as the dose duration. In this case, the value for the first dose time should be labeled with TimeAtFirstDose. In some documents, the author will also indicate the time at which the last dose is given, in that case use the TimeAtLastDose tag. Note that for TimeUnits that only 'day' was annotated as '-old' provides no further description of the time unit.



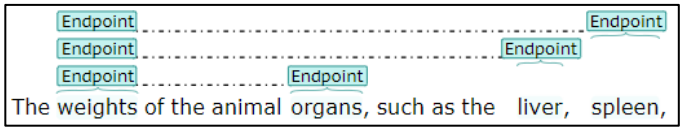


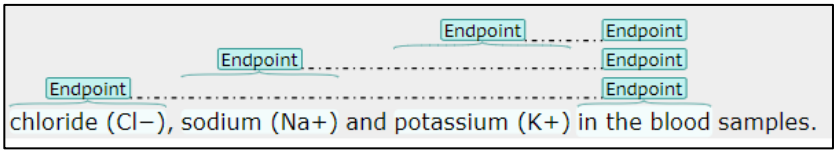


4. Endpoint Tags

Endpoint:

Annotate each endpoint being measured in a study. This includes all mentions of the endpoint throughout the methods section. General endpoints like histopathology with a list of organs should be annotated, with each organ annotated separately. Do not annotate methods describing how the endpoints were assessed. Tag sufficient context for each endpoint, for example tag "in the blood" and "chloride (Cl-)" since "chloride (Cl-)" is insufficient by itself.



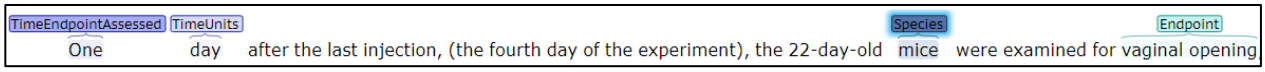


EndpointUnitOfMeasure:

Annotate the type of unit being measured for an endpoint, often this will be a value such as 'lbs', 'centimeters', etc. For some endpoints, such as a ratio, the units may be a longer phrase describing the ratio.

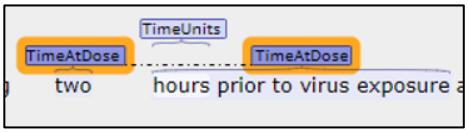
TimeEndpointAssessed:

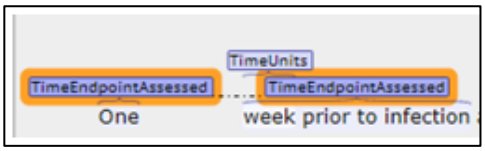
Annotate the time at which an endpoint was measured. Typically this will be in days. Note that the unit should be annotated separately using TimeUnit as in the example below.

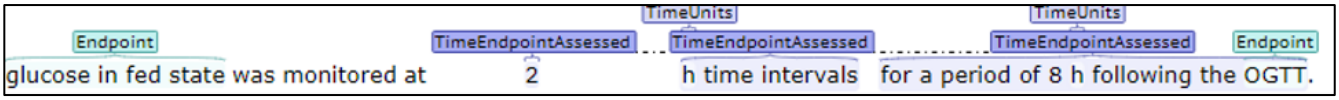


The TimeEndpointAssessed tag should provide context when the time presented is not post exposure. The default assumption is that "7" (tagged as TimeEndpointAssessed) and "hours" (tagged as timeunit) describes the time 7 hours after the test article administration or exposure.

In cases where additional context provides information that the time the endpoint was assessed was before the exposure or relative to some other time, the additional text should be tagged (see examples below)

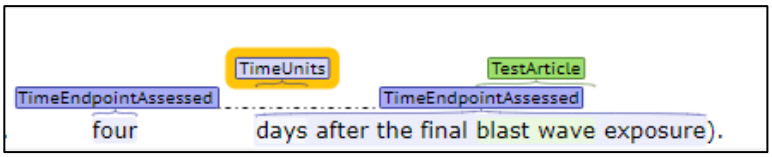


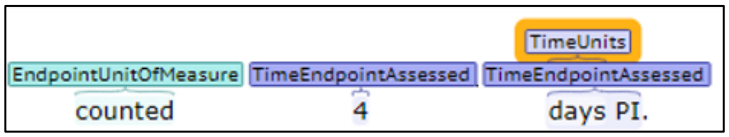




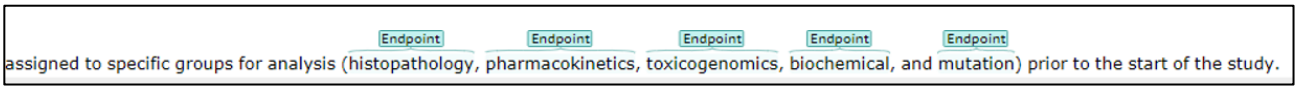
Additional guidance on how to tag context with TimeEndpointAssessed and TimeUnits

- When context is needed, the TimeEndpointAssessed tag should include the context, and TimeUnit should only include units.





- Refrain from tagging unclear time points. In the example below, "prior to the start of the study" should not be tagged as it does not provide a specific time point.



- Refrain from tagging successive time points, which potentially could have been added together to get the time endpoint assessed. In cases where a narrative is provided, and it is possible to calculate a time that the endpoint was assessed, do not tagging the multiple segments to arrive at that time endpoint assessed.

  For example, a study may say "For 4 days mice were acclimatized, and then 3 days later they were transported to their own cages. Then 2 days after that body weights were weighed". In this example "4 days"+ "3 days" + "2days" should not be tagged.

- Refrain from making assumptions between sentences. In the example below, it is not entirely clear that the cells were counted after 6-7 days. If the text said, "Then…" or "After incubation…." the 6-7 days could have been tagged as TimeEndpoint Assessed.