*\*\*\*these are example observations, this file is referred back to when qualitatively analyzing the data\*\*\**

**INJURY\_LEVEL**

* Implicitly mentioned in a general injury variable
* Common thing to have injured and uninjured
* Granularity of the variable is important (e.g. cervical, C5-6, C5)
* There are three ways we can find it (variable, another variable, study)
* Laterality is another similar variable mentioned in multiple datasets
* In dataset XXX there is a similar variable “Spinal Cord Section” but it is referring to the place that they got cells from.
* Another way dataset XXX did it cervical vs. not cervical; which makes sense for their own research but not data-harmonization; we can translate it into C\_ but what C\_ we do not know (we know the first component
* Dataset XXX: The level is implicitly and incompletely mentioned in another variable: “"cervical contusion sci" in a variable called operation
* Dataset XXX: There is not injury involved in this study so the variable does exist—the study was Repeated measurements of hindlimb CatWalk variables in normal rats
* Dataset XXX: It is not a homogeneous set as it is between uninjured and C5
* Dataset XXX: Another similar term is used with a different semantic meaning: injury\_ site is mentioned but refers to ipsilateral etc.
* One case where cervical was assigned 1 and thoracic was assigned 2
* Dataset XXX: Used an incomputable thing “--” that turned into “?”; also granularity
* Dataset XXX: what if injury is not made but are analyzing activity related to the level
* Dataset XXX: the variable is created but not applicable so has n/a filled out

**EXCLUSION\_REASON**

* If exclusion\_in\_origin\_study exists this usually exists as well
* There was one case (dataset XXX) where the exclusion criteria was in the dictionary but in the dataset, they just had yes for everything
* Two types of behaviors: either it is left blank or filled it with n/a (as most study participants are not excluded)

**EXCLUSION\_IN\_ORIGIN\_STUDY**

* Sometimes (e.g. dataset XXX) it says no but a reason is provided?

**GROUP**

* Meaning of the group is ambiguous multiple times
* For human studies why is it all “control”

**WEIGHT**

* Most are grams
* There are cases where there are like 8 variations of the weight variable
* There are few cases (dataset\_XXX for example) where there is a weight variable for the spleen but not for the animal
* Inputting in the format of a range is relatively common
* In a few datasets (e.g XX) they have the data filled in every few lines
* Weight change is a common variable even where there is not actual weight (e.g. dataset XXX)
* There are sometimes a lot of percentage of weights (e.g. Percent of start weight at d-14, dataset XXX)

**STRAIN**

* The definition is often just strain
* Humans it does not exist or is called NA

**INJURY\_TYPE:**

* The more recent ones have follow more of the system
* Specifically when it is not applicable people tend to copy and paste the definition from the site
* There are a few cases when people append other information on this one variable
* Laterality is often appended
* Severity is another common variable that comes
* There are some cases when it is simply not applicable

… more notes for more CoDEs