

Fortnightly meeting 7 May 2019

Mark

5/6/2019

Contents

1	Inhalation TTC	1
1.1	Assessment factors to extrapolate subchronic to chronic	1
1.2	Add study length info	1
1.3	Converting study results to the same units	3
1.4	Convert to mg/m ³	4
1.5	Retrieve phys-chem properties	5

1 Inhalation TTC

1.1 Assessment factors to extrapolate subchronic to chronic

I did some digging and found that both ECHA and ECETOX say to use an assessment factor of x2 to extrapolate from subchronic to chronic exposure for inhalation studies.

1.2 Add study length info

After speaking to Richard I found out he has generated a column that ID's the type of study from either the associated study information or using a long set of rules that he has come up with. I used this column (the risk_assessment_class column) as the basis for identifying the studies we are going to use for the inhalation work.

```
## [1] "subacute"           "chronic"
## [3] "acute"              "repeat dose"
## [5] "subchronic"         "cancer"
## [7] "short-term"         "subacute, subchronic"
## [9] "ecotoxicity invertebrate" "reproductive developmental"
## [11] "ecotoxicity plants"  "reproductive"
## [13] "in vitro"           "genotoxicity, repeat dose"
## [15] "developmental"       "subacute, neurotoxicity"
## [17] "multigenerational reproductive" "cancer, repeat dose"
## [19] "neurotoxicity"       "chronic, subchronic"
## [21] "air quality"         "water quality"
## [23] "clinical"           "developmental neurotoxicity"
## [25] "human"              "chemical property"
## [27] "genotoxicity"        "special toxicity study"
## [29] "immunotoxicology"
```

However, there are some potentially useful studies that have been tagged as being either short-term or repeat dose studies. To use these I wrote code that would rename them as either chronic, subchronic, or reproductive based on either the tag in the "study_type" column or using the information in the "study_duration_value" and "study_duration_units" columns.

```
ToxVal %>%
  mutate(study_length = risk_assessment_class) %>%
  # Designate if chronic study
  # If the study_type col has designation use that OR if it's not a repro/subchronic study
  # AND has study length info use that to adjust study_length column
  mutate(study_length = ifelse(risk_assessment_class %in% c("short-term", "repeat dose") &
```

```

study_type == "chronic" |
risk_assessment_class %in% c("short-term", "repeat dose") &
!study_type %in% c("reproductive", "subchronic") &
study_duration_value >=100 & study_duration_units == "day" |
risk_assessment_class %in% c("short-term", "repeat dose") &
!study_type %in% c("reproductive", "subchronic") &
study_duration_value >=14 & study_duration_units == "week" |
risk_assessment_class %in% c("short-term", "repeat dose") &
!study_type %in% c("reproductive", "subchronic") &
study_duration_value >=4 & study_duration_units == "month" |
risk_assessment_class %in% c("short-term", "repeat dose") &
!study_type %in% c("reproductive", "subchronic") &
study_duration_value >=1 & study_duration_units == "year", yes = "chronic"

# Designate if subchronic study
# If the study_type col has designation use that OR if it's not a repro/chronic study
# AND has study length info use that to adjust study_length column
mutate(study_length = ifelse(risk_assessment_class %in% c("short-term", "repeat dose") &
study_type == "subchronic" |
risk_assessment_class %in% c("short-term", "repeat dose") &
!study_type %in% c("reproductive", "chronic") &
study_duration_value >=30 & study_duration_value <100 &
study_duration_units == "day" |
risk_assessment_class %in% c("short-term", "repeat dose") &
!study_type %in% c("reproductive", "chronic") &
study_duration_value >=5 & study_duration_value <14 &
study_duration_units == "week" |
risk_assessment_class %in% c("short-term", "repeat dose") &
!study_type %in% c("reproductive", "chronic") &
study_duration_value >=1.5 & study_duration_value <4 &
study_duration_units == "month" |
risk_assessment_class == "repeat dose" &
!study_type %in% c("acute", "reproductive", "chronic") &
study_duration_value == -1 |
risk_assessment_class == "repeat dose" &
!study_type %in% c("acute", "reproductive", "chronic") &
study_duration_value == 0, yes = "subchronic", study_length)) %>%

# Designate if reproductive study
# If the study_type col has designation use that
mutate(study_length = ifelse(risk_assessment_class %in% c("short-term", "repeat dose") &
study_type == "reproductive", yes = "reproductive", study_length)) %>%

select(1:28, study_length, everything())

```

As can be seen a short-term/repeat dose study was considered to be chronic if the “study_type” column stated it was a chronic study or if the study duration was over 100 days (or week/month equivalent).

Similarly, a study was considered to be subchronic if the “study_type” column stated it was a subchronic study or if the study duration was >=35 days and < 100 days (or week/month equivalent). However, a short-term/repeat dose study was only considered to be a reproductive study if the “study_type” column stated so.

By doing this we will give ourselves approximately double the studies to be able to work with. These numbers are after filtering for:

- Test species: Rat, mouse, other rodents, or rabbit
- Route of exposure: Inhalation
- Study duration: subchronic, chronic, reproductive, developmental, and multigeneration
- POD type: NOEL, NO(A)EL, NOEC, or NO(A)EC

Number of unique chemicals:

- Based on DTXSID - 726
- With QSAR-ready SMILES - 378

The above numbers are the total number of unique chemicals and the number of unique chemicals with QSAR ready SMILES if we use Richard's designation of study type using the "risk_assessment_class" column.

Meanwhile, if we use our newly generated "study_length" column we end up with:

- 1,480 unique chemicals
- 657 unique chemicals with QSAR-ready SMILES

Additionally, we can adjust the length of study required to be considered chronic/subchronic if you don't feel comfortable with them.

Also after talking to Richard/Prachi we could probably include cancer studies in with the chronic studies, it seems like the cancer studies in ToxVal focus on the none genotoxic outcomes, should we do that ourselves?

1.3 Converting study results to the same units

Now we have a list of potential studies to draw from let's have a look at the units in which the PODs are measured.

Table 1: POD units for inhalation studies

POD units
mg/m3
ppm
mg/L
mg/kg-day
% diet
%
mg/kg
mg/kg-week
mg/m
unspecified

Above (Table 1), is the list of units used by the studies that fit our filtering criteria. What does the count of each of these units look like?

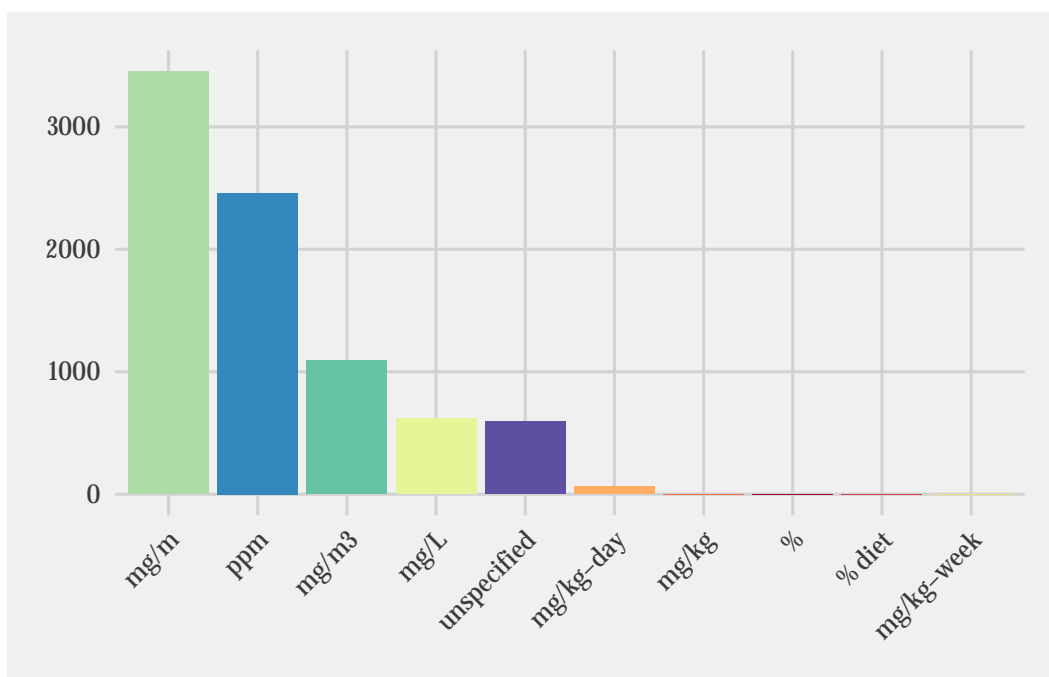


Figure 1: Count of studies each POD unit is measured in

As we can see from Figure 1, the mg/m³ is the most frequently used (I'm assuming that mg/m and mg/m³ are the same). When looking through the whole of ToxVal there were also a number of studies that had their units written as (mg/m³)-1 and (mg/kg-day)-1, because I wanted to do the conversion on all the inhalation studies I wrote something that removed the surrounding "()-1".

I'm not sure if you have a preference as to whether we convert everything to ppm or mg/m³?

1.3.1 What does UN GHS use for inhalation studies?

The GHS looks like it uses ppm for gases and mg/L for vapours and dusts & mists, there's a 1:1 ratio between these units so it's easy to switch between them.

At the minute I have the following code to convert everything to ppm:

```
ToxVal <- ToxVal %>%
  mutate(toxval_units = str_replace(toxval_units, "\\(", ""),
         toxval_units = str_replace(toxval_units, "\\)-[1]", "")) %>%
  mutate(toxval_ppm = if_else(exposure_route == "inhalation" & toxval_units == "mg/m" |
                             exposure_route == "inhalation" & toxval_units == "mg/m3", (24.45 * toxval_nu
                             if_else(exposure_route == "inhalation" & toxval_units == "g/m3", (24.45 * (to
                             if_else(exposure_route == "inhalation" & toxval_units == "ug/m3", (24
                             if_else(exposure_route == "inhalation" & toxval_units == "ppm
                             exposure_route == "inhalation" & toxval_units == "m
```

Basically, if the study is an inhalation study with either mg/m, mg/m³ the we multiply the toxval_numeric column by 24.45 (conversion made for experimental conditions assuming 1 atmosphere of pressure and temperature of 25°C) and then divided by the chemicals MW. Whilst they don't make it past the study filters, I've also added code to convert units in terms of either g/m³ or ug/m³ to ppm.

Finally, as ppm and mg/L are equivalent, I've just kept values with either of those units as is.

1.4 Convert to mg/m³

However, it seems like the EPA may favour using mg/m³ instead (pg 5 of linked PDF), as such I could convert everything into mg/m³ by doing:

```

ToxVal <- ToxVal %>%
  mutate(toxval_units = str_replace(toxval_units, "\\(", "("),
         toxval_units = str_replace(toxval_units, "\\)-[1]", "")) %>%
  mutate(toxval_mgm3 = if_else(exposure_route == "inhalation" & toxval_units == "ppm" |
                              exposure_route == "inhalation" & toxval_units == "mg/m3", toxval_mgm3,
                              if_else(exposure_route == "inhalation" & toxval_units == "g/m3", toxval_mgm3,
                              if_else(exposure_route == "inhalation" & toxval_units == "ug/m3", toxval_mgm3,
                              if_else(exposure_route == "inhalation" & toxval_units == "mg/l", toxval_mgm3,
                              exposure_route == "inhalation" & toxval_units == "mg/m3", toxval_mgm3)))

```

Do you have a preference for which units to use? It seems like we have backing either way.

1.5 Retrieve phys-chem properties

Using the DTXSIDs for the chemicals with QSAR ready SMILES I retrieved the phys-chem properties calculated by OPERA and Molecular Weight from the CompTox dashboard. These properties are:

- Molecular Weight (Average mass)
- Atmospheric hydroxylation rate
- Bioconcentration factor
- Biodegradation half-life
- Boiling point
- Henrys law (solubility of gas in liquid)
- Fish biotransformation
- Octanol-air partition coefficient (LogK_{OA})
- Soil adsorption coefficient (LogK_{OC})
- Octanol-water partition coefficient (LogP)
- Melting point
- Vapour pressure
- Water solubility

1.5.1 Are there any basic correlations we can find?

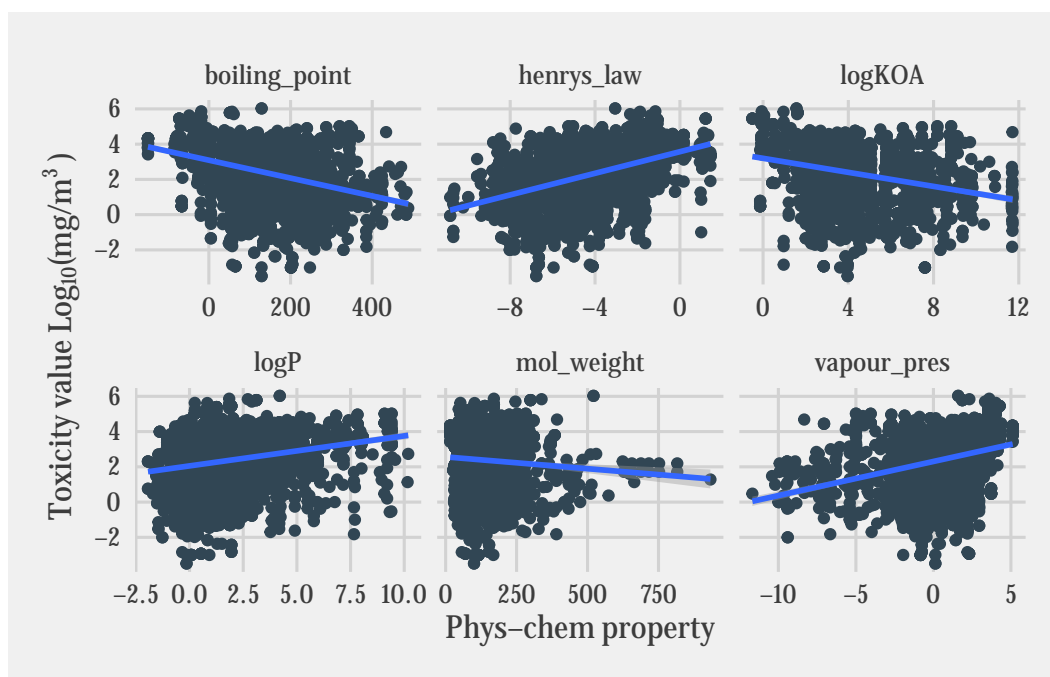


Figure 2: Comparison of phys-chem properties against log_{10} ToxVal (mg/m^3)