

```

data<-read_excel(tfile1,sheet=tsheet1)
names(data)[names(data) == 'CASRN'] <- "CAS"
if(dim(data)[1] != 0){
d2<-inner_join(tlist,data,by="CAS")

d3<-filter(d2,!grepl("pig frog",d2$SPECIES_COMMON) &
!grepl("pigeon",d2$SPECIES_COMMON) &
!grepl("pigfish",d2$SPECIES_COMMON) &
!grepl("pigmey",d2$SPECIES_COMMON) &
!grepl("pigtoe",d2$SPECIES_COMMON) &
!grepl("champignon",d2$SPECIES_COMMON) &
!grepl("pig bean",d2$SPECIES_COMMON) &
!grepl("quadruplicatum",d2$SPECIES_COMMON) &
!grepl("catesbeiana",d2$SPECIES_COMMON) &
!grepl("caterpillar",d2$SPECIES_COMMON) &
!grepl("catfish",d2$SPECIES_COMMON) & !grepl("cat
flea",d2$SPECIES_COMMON) &
!grepl("catawba",d2$SPECIES_COMMON) &
!grepl("catchweed",d2$SPECIES_COMMON) &
!grepl("cattle tick",d2$SPECIES_COMMON) &
!grepl("catjang",d2$SPECIES_COMMON) &
!grepl("catla",d2$SPECIES_COMMON) &
!grepl("catmint",d2$SPECIES_COMMON) &
!grepl("cattail",d2$SPECIES_COMMON) &
!grepl("catchfly",d2$SPECIES_COMMON) &
!grepl("subspicatus",d2$SPECIES_COMMON) &
!grepl("etunicatum",d2$SPECIES_COMMON) &
!grepl("cat's ear",d2$SPECIES_COMMON) &
!grepl("tunicate",d2$SPECIES_COMMON) &
!grepl("delicate",d2$SPECIES_COMMON) &
!grepl("yucatan",d2$SPECIES_COMMON) &
!grepl("dog tick",d2$SPECIES_COMMON) &
!grepl("dogfish",d2$SPECIES_COMMON) &
!grepl("dogwood",d2$SPECIES_COMMON) &
!grepl("dogwinkle",d2$SPECIES_COMMON) &
!grepl("dogstail",d2$SPECIES_COMMON) &
!grepl("dogbane",d2$SPECIES_COMMON) &
!grepl("dogfennel",d2$SPECIES_COMMON) & !grepl("dog
violet",d2$SPECIES_COMMON) &
!grepl("doghobble",d2$SPECIES_COMMON) & !grepl("dog
whelk",d2$SPECIES_COMMON) &
!grepl("rabbitbells",d2$SPECIES_COMMON) &
!grepl("rabbitbrush",d2$SPECIES_COMMON) &
!grepl("rat snake",d2$SPECIES_COMMON) & !grepl("rat-
tailed",d2$SPECIES_COMMON) &
!grepl("rats tail",d2$SPECIES_COMMON) & !grepl("rat
flea",d2$SPECIES_COMMON) &
!grepl("invertebrate",d2$SPECIES_COMMON) &
!grepl("migratory",d2$SPECIES_COMMON) &
!grepl("pirate",d2$SPECIES_COMMON) &
!grepl("prostrate",d2$SPECIES_COMMON) &
!grepl("firebrat",d2$SPECIES_COMMON) &
!grepl("coelenterate",d2$SPECIES_COMMON) &

```

```

!grepl("ageratum",d2$SPECIES_COMMON) &
!grepl("monkeyflower",d2$SPECIES_COMMON) &
!grepl("monkey river prawn",d2$SPECIES_COMMON) &
!grepl("mouse ear",d2$SPECIES_COMMON) &
!grepl("mouse-ear",d2$SPECIES_COMMON) &
!grepl("mouseear",d2$SPECIES_COMMON) &
!grepl("swine cress",d2$SPECIES_COMMON) & !grepl("longtailed
fruit fly parasite",d2$SPECIES_COMMON) &
!grepl("burrowing",d2$SPECIES_COMMON) & !grepl("greater
prairie-chicken",d2$SPECIES_COMMON))

if(dim(d3)[1]!=0){
for (i in 1:dim(d3)[1]) {
d3$acute[i]<-0
d3$syst[i]<-0
d3$neuro[i]<-0
d3$carc[i]<-0
d3$repro[i]<-0
d3$muta[i]<-0

d3[, 't2a']<-
ifelse(grepl("cancer",d3$STUDY_TYPE)|grepl("Cancer",d3$STUDY_TYPE)|

grepl("carcinogenicity",d3$STUDY_TYPE)|grepl("Carcinogenicity",d3$STUDY_T
YPE),"c",
ifelse(d3$STUDY_TYPE=="developmental neurotoxicity","b",

ifelse(grepl("neuro",d3$STUDY_TYPE)|grepl("Neuro",d3$STUDY_TYPE),"n",

ifelse(grepl("Development",d3$STUDY_TYPE)|grepl("development",d3$STUDY_TY
PE)|

grepl("Reproduct",d3$STUDY_TYPE)|grepl("reproduct",d3$STUDY_TYPE)|

grepl("fertility",d3$STUDY_TYPE)|grepl("uterotrophic",d3$STUDY_TYPE)|

grepl("Growth",d3$STUDY_TYPE)|grepl("Morphology",d3$STUDY_TYPE),"r",

ifelse((grepl("short",d3$STUDY_TYPE)|grepl("Short",d3$STUDY_TYPE)|

grepl("acute",d3$STUDY_TYPE)|grepl("Acute",d3$STUDY_TYPE)|
grepl("Mortality",d3$STUDY_TYPE))&
(((d3$STUDY_DURATION_UNITS %in%
c("days","Days","day","Day(s)"))&
(d3$STUDY_DURATION_VALUE<=14))|
((d3$STUDY_DURATION_UNITS %in%
c("week","Weeks"))&
(d3$STUDY_DURATION_VALUE<=2))|

(grepl("hour",d3$STUDY_DURATION_UNITS)|grepl("hr",d3$STUDY_DURATION_UNITS
))), "a",

ifelse((grepl("short",d3$STUDY_TYPE)|grepl("Short",d3$STUDY_TYPE)|

```

```

grepl("acute",d3$STUDY_TYPE)|grepl("Acute",d3$STUDY_TYPE)|
      grepl("Mortality",d3$STUDY_TYPE)) &
d3$STUDY_DURATION_UNITS=="generations","r","s")))))

if((d3$SPECIES_COMMON[i] %in% c("american bison","american
mink","bat","boar","botta's pocket gopher",
      "burro","camel","cattle","common
shrew","cow family","coyote",
      "damaliscus pygargus","deer","domestic
goat","eptesicus fuscus","ferret",
      "gambian rat","gerbil","goat","gray
fox","horse","house shrew","indian bush rat",
      "lamb","little brown bat","long eared
hedgehog","malaysian wood rat","mammal",
      "meadow vole","micromys
minutus","microtus arvalis","microtus canicaudus",
      "microtus guentheri","microtus
oregoni","microtus pennsylvanicus","microtus sp.",
      "microtus
subterraneus","mink","mongolian jird","monodelphis domestica",
      "montane shrew","montane
vole","mountain beaver","mustela erminea",
      "mustela putorius","mustela putorius
ssp. Furo","neovison vison",
      "northern short-tailed shrew","northern
treeshrew","notomys mitchellii",
      "occult little brown myotis
bat","ochotona rufescens ssp. rufescens",
      "octodon degus","odocoileus
hemionus","odocoileus hemionus ssp. hemionus",
      "ondatra zibethicus","oryzomys
palustris","ovis aries","oxen",
      "peromyscus californicus","peromyscus
gossypinus","peromyscus maniculatus",
      "peromyscus polionotus","peromyscus
sp.,"pine vole","pipistrellus pipistrellus",
      "plains pocket gopher","prairie
vole","prarie vole","polynesian rat","raccoon",
      "red fox","reithrodontomys
fulvescens","rhinopoma microphyllum","river otter",
      "root vole","sciurus
carolinensis","shaw's jird","sheep","short-tailed shrew",
      "sminthopsis
crassicaudata","spermophilus beecheyi","springbok","suncus murinus",
      "tadarida brasiliensis","tammar
wallaby","tatera indica",
      "thirteen line ground
squirrel","thomomys bottae ssp. navus","thomomys monticola",
      "thomomys talpoides","townsend's
chipmunk","trichosurus Vulpecula",
      "tristram's jird","trowbridge's
shrew","vole","vulpes fulva","water buffalo",

```

```

                                "weasel","western pocket
gopher","white-tailed deer") |
                                grepl("field mouse",d3$SPECIES_COMMON[i]) |
grepl("field rat",d3$SPECIES_COMMON[i]) |
                                grepl("mole rat",d3$SPECIES_COMMON[i]))){
                                if(d3$TOXVAL_TYPE[i] %in%
c("ATCN","BIED","EC","EC0","EC01","EC05","EC10","EC100","EC5","EC50",
"ECNA","ED0","ED01","ED05","ED1","ED10","ED100","ED5","ED50","EL0","EL10"
,
"EL50","ER01","ER05","ER10","ER50","ET0","ET100","ET50","IC01","IC05","IC
10",
"IC100","IC50","ID10","IP","IP0","IP10","IP100","IP5","IP50","L(Ct) 50","L
BR50",
"LCNA","LETC","LL0","LL10","LL100","LL50","LR10","LR100","LR50","MATC","M
ATR",
"MCIG","No_Effect_ESL","NOELR","NOER","NR-LETH","NR-
ZERO","PNEC","ALC","ALD",
                                "conc. level","conc. Level:","conc.
Level:","FEL","ID50","LC","LC0","LC01",
"LC50","LC100","LCLo","LCLow","LD","LD0","LD01","LD05","LD10","LD50","LD1
00",
"LDLo","PAC_1","PAC_2","PAC_3","other","TD05","LEL","Cancer slope
factor",
                                "cancer slope
factor","cancer_slope_factor","CSFo","cancer unit risk","IUR",
                                "inhalation UR","inhalation unit
risk","Inhalation Unit Risk","inhalation TC",
                                "inhalation slope factor","Inhalation
SF","inhalation SF","oral slope factor",
                                "Oral slope factor","oral SF","oral
unit risk","SFO","other:") |
                                grepl("BMC", d3$TOXVAL_TYPE[i]) | grepl("BMD",
d3$TOXVAL_TYPE[i]) |
                                grepl("dose level ", d3$TOXVAL_TYPE[i]) |
grepl("MLD", d3$TOXVAL_TYPE[i]) |
                                grepl("AEGL", d3$TOXVAL_TYPE[i]) | grepl("MRL",
d3$TOXVAL_TYPE[i]) |
                                grepl("OAEL", d3$TOXVAL_TYPE[i]) | grepl("Observed
effect level ", d3$TOXVAL_TYPE[i]) |
                                grepl("NEL", d3$TOXVAL_TYPE[i]) | grepl("OEC",
d3$TOXVAL_TYPE[i]) |
                                grepl("OAEC", d3$TOXVAL_TYPE[i]) | grepl("RfC",
d3$TOXVAL_TYPE[i]) |
                                grepl("RfD ", d3$TOXVAL_TYPE[i]) | grepl("TDI",
d3$TOXVAL_TYPE[i]) |
                                grepl("ADI", d3$TOXVAL_TYPE[i]) | grepl("TWI",
d3$TOXVAL_TYPE[i]) |

```

```

                                grepl("REL", d3$TOXVAL_TYPE[i]) | (grepl("OEL",
d3$TOXVAL_TYPE[i]) &
                                d3$TOXVAL_TYPE[i] %notin%
c("NOELR", "LOELR"))){d3$exp[i]<-1}}
else if(d3$TOXVAL_TYPE[i]=="AC50" &
d3$SPECIES_COMMON[i]=="nematode"){d3$exp[i]<-1}
else if(grepl("pig",d3$SPECIES_COMMON[i]) |
grepl("human",d3$SPECIES_COMMON[i]) |
        grepl("rat",d3$SPECIES_COMMON[i]) |
grepl("cat",d3$SPECIES_COMMON[i]) | grepl("dog",d3$SPECIES_COMMON[i]) |
        grepl("hamster",d3$SPECIES_COMMON[i]) |
grepl("rabbit",d3$SPECIES_COMMON[i]) |
        grepl("mouse",d3$SPECIES_COMMON[i]) |
grepl("mice",d3$SPECIES_COMMON[i]) | grepl("monkey",d3$SPECIES_COMMON[i])
|
        grepl("primate",d3$SPECIES_COMMON[i]) | grepl("mesocricetus
auratus",d3$SPECIES_COMMON[i]) |
        grepl("swine",d3$SPECIES_COMMON[i]) |
grepl("macaque",d3$SPECIES_COMMON[i]) |
        grepl("marmoset",d3$SPECIES_COMMON[i]) |
grepl("oryctolagus",d3$SPECIES_COMMON[i]) |
        grepl("sylvilagus",d3$SPECIES_COMMON[i]) | grepl("norway
rat",d3$SPECIES_COMMON[i]) |
        grepl("black rat",d3$SPECIES_COMMON[i]) | grepl("fruit
fly",d3$SPECIES_COMMON[i]) |
        grepl("fruitfly",d3$SPECIES_COMMON[i]) | grepl("leydig
cell",d3$SPECIES_COMMON[i]) |
        grepl("mammalian cell line",d3$SPECIES_COMMON[i]) |
grepl("salmonella typhimurium",d3$SPECIES_COMMON[i]) |
        grepl("bacillus subtilus",d3$SPECIES_COMMON[i]) |
grepl("escherichia coli",d3$SPECIES_COMMON[i]) |
        grepl("cell line",d3$SPECIES_COMMON[i]) | d3$SPECIES_COMMON[i] %in%
c("-", "zebrafish")){
  if (d3$TOXVAL_TYPE[i] %in% c("ALC", "ALD")){
    d3$quan[i]<-1
    d3$acute[i]<-1
    if (d3$SPECIES_COMMON[i] != "-"){d3$animal[i]<-1}}
  if (d3$TOXVAL_TYPE[i] %in% c("conc. level", "conc. Level:")){
    d3$quan[i]<-1
    if (d3$SPECIES_COMMON[i] %notin% c("human", "mammalian cell line", "-
")){d3$animal[i]<-1}}
  if (d3$TOXVAL_TYPE[i]=="FEL"){
    d3$quan[i]<-1
    d3$animal[i]<-1}
  if (grepl("BMC",d3$TOXVAL_TYPE[i]) | grepl("BMD",d3$TOXVAL_TYPE[i])) {
    d3$quan[i]<-1
    d3$animal[i]<-1}
  if (d3$TOXVAL_TYPE[i]=="ID50"){
    d3$quan[i]<-1
    d3$animal[i]<-1
    d3$neuro[i]<-1}
  if (d3$TOXVAL_TYPE[i] %in%
c("LC", "LC0", "LC01", "LC50", "LC100", "LCLo", "LCLow")){
    d3$quan[i]<-1

```

```

d3$acute[i]<-1
if (grepl("fruit fly",d3$SPECIES_COMMON[i]) |
grepl("fruitfly",d3$SPECIES_COMMON[i])){d3$exp[i]<-1}
else {d3$animal[i]<-1}}
if(d3$TOXVAL_TYPE[i] %in%
c("LD","LD0","LD01","LD05","LD10","LD50","LD100","LDLo")) {
d3$quan[i]<-1
d3$acute[i]<-1
if(d3$SPECIES_COMMON[i] %in% c("cell line","mammalian cell line") |
grepl("fruit fly",d3$SPECIES_COMMON[i]) |
grepl ("fruitfly",d3$SPECIES_COMMON[i])) {d3$exp[i]<-1}
else {d3$animal[i]<-1}}
if(d3$TOXVAL_TYPE[i] %in% c("PAC_1","PAC_2","PAC_3")) {
d3$quan[i]<-1
d3$acute[i]<-1}
if(d3$TOXVAL_TYPE[i]=="other") {
d3$quan[i]<-1
d3$animal[i]<-1}
if(d3$TOXVAL_TYPE[i]=="other:" &
d3$SPECIES_SUPERCATEGORY[i]=="mammals") {
d3$quan[i]<-1
d3$animal[i]<-1}
if(grepl("dose level",d3$TOXVAL_TYPE[i])) {
d3$quan[i]<-1
d3$animal[i]<-1}
if(grepl("MLD",d3$TOXVAL_TYPE[i]) | d3$TOXVAL_TYPE[i]=="TD05") {
d3$quan[i]<-1
d3$animal[i]<-1
d3$acute[i]<-1}
if(grepl("AEGL",d3$TOXVAL_TYPE[i])) {
d3$quan[i]<-1
d3$acute[i]<-1}
if(grepl("MRL",d3$TOXVAL_TYPE[i])) {
d3$quan[i]<-1
d3$animal[i]<-1
if(d3$STUDY_DURATION_UNITS[i]=="day" &
d3$STUDY_DURATION_VALUE[i]==14){d3$acute[i]<-1}
else {d3$syst[i]<-1}}
if(grepl("OAEL",d3$TOXVAL_TYPE[i]) | grepl("Observed effect
level",d3$TOXVAL_TYPE[i]) |
(grepl("OEL",d3$TOXVAL_TYPE[i]) & !grepl("NOELR",d3$TOXVAL_TYPE[i])
& !grepl("LOELR",d3$TOXVAL_TYPE[i]))) {
d3$quan[i]<-1
if(!grepl("fruit fly",d3$SPECIES_COMMON[i]) &
!grepl("fruitfly",d3$SPECIES_COMMON[i]) &
!grepl("cell line",d3$SPECIES_COMMON[i])) {d3$animal[i]<-1}}
if(d3$TOXVAL_TYPE[i]=="LEL") {
d3$quan[i]<-1
if(!grepl("fruit fly",d3$SPECIES_COMMON[i]) &
!grepl("fruitfly",d3$SPECIES_COMMON[i]) &
!grepl("human",d3$SPECIES_COMMON[i]) |
!grepl("salmonella",d3$SPECIES_COMMON[i]) &
!grepl("bacillus",d3$SPECIES_COMMON[i]) & !grepl("escherichia
",d3$SPECIES_COMMON[i])) {d3$animal[i]<-1}}

```

```

    if(grepl("NEL",d3$TOXVAL_TYPE[i])) {
      d3$quan[i]<-1
      if(!grepl("salmonella",d3$SPECIES_COMMON[i]) &
!grepl("escherichia",d3$SPECIES_COMMON[i]) &
      !grepl("bacillus",d3$SPECIES_COMMON[i])) {d3$animal[i]<-1}}
      if(d3$TOXVAL_TYPE[i] %in% c("Cancer slope factor","cancer slope
factor","cancer_slope_factor","CSFo",
      "cancer unit risk","IUR","inhalation UR", "inhalation unit
risk","Inhalation Unit Risk","inhalation TC",
      "inhalation slope factor","inhalation SF","Inhalation SF","oral
slope factor","Oral slope factor",
      "Oral Slope Factor","oral SF","oral unit risk","SFO")) {
        d3$quan[i]<-1
        d3$animal[i]<-1
        d3$carc[i]<-1}
      if(grepl("OEC",d3$TOXVAL_TYPE[i]) | grepl("Oaec",d3$TOXVAL_TYPE[i])) {
        d3$quan[i]<-1
        if(!grepl("fruit fly",d3$SPECIES_COMMON[i]) &
          d3$SPECIES_COMMON[i] %notin% c("human","cell line","mammalian cell
line")) {d3$animal[i]<-1}}
        if(grepl("RfC",d3$TOXVAL_TYPE[i])) {
          d3$quan[i]<-1
          d3$animal[i]<-1}
        if(grepl("RfD",d3$TOXVAL_TYPE[i])) {
          d3$quan[i]<-1
          d3$animal[i]<-1
          if(d3$SUB_SOURCE[i]=="OPP" & d3$STUDY_TYPE[i]=="acute"){d3$acute[i]<-
1}}
        if(grepl("TDI",d3$TOXVAL_TYPE[i])) {
          d3$quan[i]<-1
          d3$syst[i]<-1}
        if(grepl("TWI",d3$TOXVAL_TYPE[i]) | grepl("ADI",d3$TOXVAL_TYPE[i])) {
          d3$quan[i]<-1
          d3$animal[i]<-1
          d3$syst[i]<-1}
        if(grepl("REL",d3$TOXVAL_TYPE[i])) {
          d3$quan[i]<-1
          d3$animal[i]<-1
          if(d3$STUDY_TYPE[i]=="acute"){d3$acute[i]<-1}
          if(d3$STUDY_TYPE[i] %in% c("chronic","-")) {d3$syst[i]<-1}}
        if(d3$TOXVAL_TYPE[i] %in% c("ABS","GIABS")) {
          d3$tox[i]<-1
          d3$t_aabs[i]<-1}}
      if(d3$SPECIES_COMMON[i] %in% ("zebrafish")){
        d3$animal[i]=0
        if(d3$quan[i]==1){d3$exp[i]=1}}

if(grepl("Genetic toxicity",d3$STUDY_TYPE[i])) {
  d3$muta[i]<-1
  if(d3$SPECIES_COMMON[i] != "human") {d3$animal[i]<-1}}
if(d3$SPECIES_COMMON[i] %in% c("human") & d3$animal[i]==0) {d3$human[i]<-
1}
if(d3$SPECIES_COMMON[i] %in% c("hen","chicken") & d3$STUDY_TYPE[i] %in%
c("neurotoxicity","Neurotoxicity")) {

```

```

    d3$animal[i]<-1
    d3$neuro[i]<-1}
if(grepl("cell line",d3$SPECIES_COMMON[i]) & d3$STUDY_TYPE[i] !=
"genotoxicity") {d3$exp[i]<-1}

if(d3$acute[i] != 1 & d3$syst[i] != 1 & d3$neuro[i] != 1 & d3$carc[i] !=
1 & d3$muta[i] != 1 &
    d3$repro[i] != 1) {
  if(d3$t2a[i]=="c"){d3$carc[i]<-1}
  if(d3$t2a[i]=="n"){d3$neuro[i]<-1}
  if(d3$t2a[i]=="r"){d3$repro[i]<-1}
  if(d3$t2a[i]=="a"){d3$acute[i]<-1}
  if(d3$t2a[i]=="s"){d3$syst[i]<-1}
  if(d3$t2a[i]=="b"){
    d3$neuro[i]<-1
    d3$repro[i]<-1}}

if(d3$animal[i]==1) {
  if(d3$acute[i]==1) {d3$a_acute[i]<-1}
  if(d3$syst[i]==1) {d3$a_syst[i]<-1}
  if(d3$neuro[i]==1) {d3$a_neuro[i]<-1}
  if(d3$carc[i]==1) {d3$a_carc[i]<-1}
  if(d3$muta[i]==1) {d3$a_muta[i]<-1}
  if(d3$repro[i]==1) {d3$a_repro[i]<-1}}

if(d3$human[i]==1) {
  if(d3$acute[i]==1) {d3$h_acute[i]<-1}
  if(d3$syst[i]==1) {d3$h_syst[i]<-1}
  if(d3$neuro[i]==1) {d3$h_neuro[i]<-1}
  if(d3$carc[i]==1) {d3$h_carc[i]<-1}
  if(d3$muta[i]==1) {d3$h_muta[i]<-1}
  if(d3$repro[i]==1) {d3$h_repro[i]<-1}}

if(d3$quan[i]==1) {
  if(d3$acute[i]==1) {d3$qn_acute[i]<-1}
  if(d3$syst[i]==1) {d3$qn_syst[i]<-1}
  if(d3$neuro[i]==1) {d3$qn_neuro[i]<-1}
  if(d3$carc[i]==1) {d3$qn_carc[i]<-1}
  if(d3$repro[i]==1) {d3$qn_repro[i]<-1}}
}}

d4<-d3[,c(2,5:ncolt)]
n<- d4 %>% group_by(CAS) %>%
  summarise_all(sum)

all<-merge(n,tlist,all.y=TRUE,by="CAS")
commonNames <- names(n)[which(colnames(n) %in% colnames(tlist))]
commonNames <- commonNames[commonNames != "CAS"]
for(i in commonNames){
  left <- paste(i, ".x", sep="")
  right <- paste(i, ".y", sep="")
  all[is.na(all[left]),left] <- all[is.na(all[left]),right]
  all[right]<- NULL
  colnames(all)[colnames(all) == left] <- i

```

```
}  
  
all<-all[order(all$CID,all$CAS,na.last=FALSE),]  
  
all<-all[,c(ncolt-2,1,(ncolt-1):ncolt,2:(ncolt-3))]}  
if(dim(data)[1] == 0){all<-tlist}  
  
wb <- loadWorkbook(outfile)  
writeData(wb, sheet = "Toxval", all[1:ncolt], colNames = F,startRow=3)  
saveWorkbook(wb,outfile,overwrite = T)
```