Documentation of the Fields in the Carcinogenic Potency Database (CPDB)

Organization

There are 12 Excel datasets:

Dataset Name	Brief Description
Section 1: Main datasets	
cpdb.cpdb.ncintp.xls	All data for NCI/NTP except doses and incidence
cpdb.cpdb.ncintpdose.xls	Dose and incidence data for NCI/NTP
cpdb.lit.xls	All data for literature except: doses and incidence
cpdb.litdose.xls	Dose and incidence data for literature
cpdb.chemname.xls	Chemical names, three-letter identification codes and CAS numbers
cpdb.cit.xls	Brief citation to published paper in literature
Section 2: Datasets of code definitions	
cpdb.species.xls	Species code definitions
cpdb.route.xls	Route code definitions
cpdb.strain.xls	Strain code definitions
cpdb.tissue.xls	Tissue code definitions
cpdb.tumor.xls	Tumor histopathology code definitions
cpdb.journal.xls	Journal code definitions

Section 1: Main datasets

cpdb.lit.xls and cpdb.ncintp.xls

Structure of the data. Each row in cpdb.ncintp.xls or cpdb.lit.xls represents a tissue-tumor combination for an experiment with a corresponding TD_{50} value.

There is a one-to-many mapping between rows in cpdb.ncintp.xls and rows in cpdb.ncintpdose.xls. The same idea applies to cpdb.lit.xls and to cpdb.litdose.xls. The "idnum" field is the key that maps the tissue-tumor combination for an experiment to its associated doses and incidence. Each tissue-tumor combination in an experiment has a unique idnum.

The cpdb.ncintp.xls and cpdb.lit.xls datasets are sorted on "chemcode", "papernum", "species", and "sex".

For cpdb.lit.xls and cpdb.ncintp.xls datasets, an experiment is defined as a unique combination of the following fields (defined below): chemcode, papernum, species, strain, sex, route, xprtime, and xpotime.

Fields in the cpdb.ncintpdose.xls and cpdb.litdose.xls datasets have identical meanings.

Differences between cpdb.ncintp.xls and cpdb.lit.xls datasets. The fields of the cpdb.ncintp.xls and cpdb.lit.xls datasets have identical meanings in almost all cases. Exceptions are:

tissue	Always length 3 for literature, varies widely for NCI/NTP.
tumor	Always length 3 for literature, varies widely for combinations of tumors in NCI/NTP. Tissue length always equals tumor length since tissue is repeated for combinations of tumor types.
inad	Field exists in cpdb.ncintp.xls only.
mandtry	Field exists in cpdb.ncintp.xls only.
mixberk	Field exists in cpdb.ncintp.xls only.
poundsgn	Field exists in cpdb.ncintp.xls only.
step	Field exists in cpdb.ncintp.xls only.

Definitions

The cpdb.ncintp.xls and cpdb.lit.xls datasets

The order of the fields and definitions below is based on protocol information, results and incidence data. To facilitate locating the fields when the field name is given, the following is the alphabetic list of fields showing the number in the order presented.

chemcode (2); ctotal (29); ctumors (31); curve (27); datanum (32); historic (11); idnum (1); inad (12); lc (22); lifetbl (20); mandtry (14); mixberk (13); ndoses (26); ndsig (28); notes (17); opinion (10); papernum (3); plotsym (24); pool (30);

poundsgn (25); pval (21); route (7); sex (6); species (4); step (18); strain (5); td50 (19); tissue (8); tumor (9); uc (23); xpotime (15); xprtime (16)

1 idnum — A unique number assigned to every row in the cpdb.ncintp.xls and cpdb.lit.xls datasets. It is used to link the tissue-tumor combinations for an experiment in cpdb.ncintp.xls or cpdb.lit.xls to its associated doses in cpdb.ncintpdose.xls or cpdb.litdose.xls, respectively.

2 chemcode — A three-character-code that identifies the test compound. See cpdb.chemname.xls for definitions and CAS numbers.

3 papernum — For literature, the unique identification number assigned to each paper. Can contain alphabetic characters when more than one experiment is reported in a paper. For NCI/NTP this is the Technical Report number. For NCI/NTP, there is only one chemcode per paper number, i.e. one chemical name.

4 species — "R" for rat, "M" for mouse, "H" for hamsters, "D" for dogs, "P" for monkeys, "N" for prosimians.

5 strain — For NCI/NTP mouse is always "b6c" for $B6C3F_1$; rat is either "f34" for Fischer F344/N, "sda" for Sprague-Dawley, or "osm" for Osborne-Mendel. Nomenclature reflects that used by the literature author. For monkeys and prosimians this field describes the species, e.g. "rhe" for Rhesus. See appendix 1 and dataset cpdb.strain.xls for definitions.

6 sex — "M" for male, "F" for female. Additionally in literature "B" is used for both sexes combined when the data in the published paper are reported only for both sexes combined.

7 route — route of administration of the compound.

Route code	Full name
cap	capsule (used for some dog experiments)
eat	diet
gav	gavage
inh	inhalation
ipj	intraperitoneal injection
ivj	intravenous injection
mix	multiple routes
orl	gavage preweanling, then diet, used only for the Innes series of experiments (1968/1969)
wat	water

8 tissue — Single tissue or group of tissues. Nomenclature reflects that used by NCI/NTP or by the literature author. See Appendix 2 of tissue codes and definitions below and also the dataset cpdb.tissue.xls. Each tissue code is 3 characters long, so a combination of 2 tissues in cpdb.ncintp.xls will be 6 characters long. See the mixberk and mandtry fields, above for cpdb.ncintp.xls.

The dataset cpdb.lit.xls does not have explicit mixes of tissues as cpdb.ncintp.xls does (e.g., cpdb.lit.xls will report "mix" rather than an explicit list of tissues). The tissue field length for cpdb.lit.xls is exactly 3 characters long. For a given row, the number of tissues equals the number of tumors.

9 tumor — Single tumor type or group of tumors. Nomenclature reflects that used by NCI/NTP or by the literature author. See Appendix 3 of tumor codes and definitions below and also the dataset cpdb.tumor.xls. Each tumor code is 3 characters long, so a combination of 2 tumors in cpdb.ncintp.xls will be 6 characters long. See the mixberk and mandtry fields, above for cpdb.ncintp.xls.

The dataset cpdb.lit.xls does not have explicit mixes of tumors as cpdb.ncintp.xls does (e.g., cpdb.lit.xls will report "mix" rather than an explicit list of tumors if tumors are reported in the paper or "tum" if the tumor types are not reported). The tumor field length for cpdb.lit.xls is exactly 3 characters long. For a given row, the number of tissues equals the number of tumors.

10 opinion — The author's opinion.

cpdb.lit.xls: the original author's opinion as to carcinogenicity of test agent at the tissue and tumor combination. Determined from the published paper and sometimes by personal communication in addition.

- \+ Author in literature evaluated the tissue-tumor combination as induced by the test agent. Every tissue-tumor combination that the author stated was induced is included with a "+". Occasionally an author evaluated a test agent as "carcinogenic" without reporting a target site; a "+" opinion is given for "all tumor-bearing animals" (tba) in this case.
- − Author evaluated the test agent as negative for carcinogenicity. Expressly indicated that the test agent did not induce the tumors at this site, and a minus opinion is used.
- 0 No opinion or ambiguous opinion

cpdb.ncintp.xls: Every tissue-tumor combination that NCI/NTP gave an opinion to has a value in this field indicating the evaluation.

- c "Carcinogenic" in the NCI/NTP Technical Report evaluation; "clear evidence" evaluation in NTP reports since 1986.
- p "Some evidence of carcinogenicity" in Technical Report evaluation; used by NTP since 1986.
- a Tumors are "associated" with carcinogenicity or the evidence was suggestive. Used in Technical Reports published through 1986. These evaluations are consistent with Haseman *et al.* (*Environ. Health Perspect.* 74: 229-235, 1987).
- e "Equivocal evidence of carcinogenicity" in Technical Report evaluation; used by NTP only since 1986.
- 0 NCI/NTP did not give an evaluation for this tissue-tumor combination or evaluated the experiment as inadequate. The site is one of the following: 1) a statistically significant site (likelihood ratio test); 2) "all tumor-bearing animals" (tba); 3) mandatory site; 4) Berkeley mix.
- – For NCI/NTP experiments that do not have a "c", "p", "a" or "e" opinion, one site in the experiment will be given a "–" opinion unless the experiment is inadquate (see "inad" field).

For negative NCI/NTP tests, the "–" opinion is given for "all tumor bearing animals" unless there is a statistically significant ($p<0.05$) site, in which cases the "–" is given to that site (see field "poundsgn").

11 historic — The literature author or NCI/NTP based a positive opinion for the tissue-tumor combination on historical control information. Value is "h" for historical, otherwise value is "0".

12 inad — A few NCI/NTP experiments were evaluated by NCI/NTP as inadequate. These have the value "i", others have the value "0".

13 mixberk — Only used for cpdb.ncintp.xls dataset. Mixes created for the CPDB (Berkeley Mixes) by combining target sites that are evaluated individually by NCI/NTP. This field is "0" for all sites that are not Berkeley mixes. The opinion field is "0" for these cases.

- c a mix of tissues and tumors with "c" opinions, i.e. clear evidence.
- m a mix of tissues and tumors with "c" or "p" opinions, i.e. clear or some evidence.
- p a mix of tissues and tumors with "p" opinions, i.e. some evidence.
- s a site or mix which has no "c", "a", "p" or "e" in the author's opinion field, and has pval<0.05, and is not a mandatory site from the NCI/NTP Technical Report. The author's opinion field is "0".

14 mandtry — Only used for cpdb.ncintp.xls dataset. Indicates mandatory sites calculated as Berkeley Mixes for all NCI/NTP experiments. When the row represents "all tumor bearing animals", this field has the value "t". For other mandatory sites, this field has the value "m" and the tissue and tumor fields are one of the following:

1) rats or mice: tissue=liver and tumor=hpa (hepatocellular adenoma), hpc (hepatocellular carcinoma), nnd (neoplastic nodule)
2) rats or mice: tissue=liver and tumor=hpa, hpb (hepatoblastoma), hpc
3) mice: tissue=lung and tumor=a/a (alveolar bronchiolar adenoma), a/c (alveolar bronchiolar carcinoma).

All sites except these mandatory sites have the value "0" for this field.

15 xpotime — The length of time in weeks that the animals were administered the test agent. If for example, dosing was once a week for 40 weeks, then xpotime is 40 weeks. Within a single experiment, all rows have one xpotime and one xprtime.

16 xprtime — The length of time in weeks the animals were on test from first day to terminal sacrifice or time of death of last dosed animal. This value is not the age of the animals

17 notes — Supplementary information that is helpful in evaluating the experimental data. For example, the note code "s" is used to denote that *survival* was poor due to toxicity or disease, and the note code "v" denotes that dosing was *variable* or irregular, e.g., dose level changed during the course of the experiment. Other note codes indicate such factors as: the experiment was a serial sacrifice in a longer study (note code "k"), or that histopathological examination was restricted to only a few tissues (note code "r"). See the file "Note codes.rtf" for note code definitions.

18 step — Only used for the dataset cpdb.ncintp.xls. In some recent NTP bioassays, results for the kidney were reported in the Technical Reports for the standard histopathology protocol and separately for results including additional sections of the kidney. The value is "s" for step incidence data including step sections and standard histopathology; otherwise value is "0" .

19 td50 — value, in mg/kg/day, of potency calculation. TD_{50} may be defined as follows: for a given target site(s), if there are no tumors in control animals, then TD_{50} is that chronic dose-rate in mg/kg body wt/day which would induce tumors in half the test animals at the end of a standard lifespan for the species.

20 lifetbl — An "l" indicates that the TD_{50} was calculated using lifetable data, and an "s" indicates summary data. In the literature, only a few series of experiments had lifetable data available. In NCI/NTP all are lifetable TD_{50}s except for some of the kidney sites with step sections.

21 pval — The likelihood ratio statistic tests the hypothesis that the test agent has no carcinogenic effect, i.e., the statistical significance (2-tailed) associated with testing whether the slope of the dose-response is different from zero. When pval=0, this implies that $p \leq 0.0005$.

22 lc — lower 99% confidence limit of TD_{50}, given in mg/kg/day. lc≥1e8 indicatest that no lower confidence could be estimated. See "Methods.rtf" for details.

23 uc — upper 99% confidence limit of TD_{50}, given in mg/kg/day. If uc≥1e8 then $p>0.01$ and the 99% confidence limit could not be calculated.

24 plotsym — the designation for whether this TD_{50} is the most potent TD_{50} estimated in the experiment and therefore the plotted symbol on the TD_{50} graph in the plot. "%" indicates most potent, "$" is all other.

25 poundsgn — For NCI/NTP only. When the most potent TD_{50} is the only evidence for a treatment-related effect and pval<0.05, this field has the value "#", otherwise it is "0".

26 ndoses — Number of dose groups in the experiment in addition to controls.

27 curve — The shape of the dose-response; based on the χ^2 goodness-of-fit statistic to test the validity of a linear relationship between dose and tumor incidence.

- \ Experiment has 2 dose groups in addition to controls. Goodness-of-fit test indicated significant departure from linearity ($p<0.05$), departure was downward, and TD_{50} calculated for one dose group only.
- * Experiment has 2 or more dose groups in addition to controls, and consistent with linearity.
- / The experiment has 2 dose groups in addition to controls, and the goodness-of-fit test indicated significant departure from linearity and departure was upward. All dose-groups are used for the pval field.
- Z Experiment has more than 2 dose groups in addition to controls. Goodness-of-fit test indicated significant departure from linearity and departure was either upward or downward. The field ndsig indicates the number of doses used in the TD_{50} calculation and the *p*-value calculation. If ndsig is less than ndoses, then the analysis was repeated without the highest dose group.
- 0 Either no dose-related effect ($p=1$), or no curve shape could be determined because experiment had only one dose group in addition to controls.

28 ndsig — Number of dose-groups used for TD_{50} and statistical significance in cpdb.ncintp.xls or cpdb.lit.xls. If the dose-response curve is non-linear curving downwards, the TD_{50} and *p*-value are estimated without the highest dose, and therefore ndsig will be lower than ndoses.

29 ctotal — For NCI/NTP, the number of control animals at the start of the experiment. For literature, ctotal is either the starting number of control animals or else the effective number. Effective number is defined as either: (1) the number of animals alive at the time of the first tumor, or if that is not reported, then (2) the number of animals examined histopathologically.

30 pool — The incidence is based on pooled control data (value is "p" for pool, otherwise value is "0").

31 ctumors — Number of tumors in control group.

32	datanum	Corresponds to the publication of the CPDB in which the data were first plotted. Numbers 1 through 6 appeared in *Environmental Health Perspectives*: 1 is volume 58 (1984), 2 is volume 67 (1986), 3 is volume 74 (1987), 4 is volume 84 (1990), 5 is volume 100 (1993), 6 is volume 103 (Supplement 8) (1995). Number 7 is for data appearing for the first time in the combined plot (1 through 7) in *Handbook of Carcinogenic Potency and Genotoxicity Databases*, L. S. Gold and E. Zeiger, eds. Boca Raton, FL: CRC Press (1997). Number 8 is *Environmental Health Perspectives* volume 107 (Suppl. 4) (1999).

The cpdb.ncintpdose.xls and cpdb.litdose.xls datasets. A row in these datasets is a dose-group within an experiment. Control data are reported in cpdb.lit.xls and cpdb.ncintp.xls, not in this dataset.

idnum	Links a dose record to a unique number assigned to every tissue-tumor combination in the datasets. This number can be used to join the doses in the cpdb.ncintpdose.xls and cpdb.litdose.xls datasets with their corresponding tissue-tumor combinations in cpdb.ncintp.xls or cpdb.lit.xls. For an idnum, there can be 1 or more doses having that idnum.
dose	The value of the dose-rate in mg/kg/day. If exposure time is less than experiment time then the daily dose-rate is an average rate over the length of the experiment.
order	For all but 5 chemicals in cpdb.ncintpdose.xls, dose-rates (mg/kg/day) are ordered as they were administered. Due to variable or discontinued dosing schedules, the order is non-monotonic for some experiments in: kepone, 1-amino-2-methylanthraquinone, methyl bromide, 5-nitro-*o*-anisisine, and 2,3,5,6-tetrachloro-4-nitroanisole.
tumors	The number of animals in this dose group with tumors of the type in the tissue-tumor combination.
total	For NCI/NTP the number of animals in the group at the start of the experiment, whether or not all were examined histologically at the site. For literature, the starting number or effective number.

The cpdb.chemname.xls dataset

chemcode	Three-character-code. This is the key for merging the full chemical names into the cpdb.lit.xls and cpdb.ncintp.xls datasets.
name	Full chemical name; can be up to 150 characters long.
sortordr	After you have merged the names into a dataset, if you want to sort the names "chemo-alphabetically". The chemo-alphabetical sort first looks at names by word, e.g., "1-allyl-1-nitrosourea" is 4 words. Names are sorted by their first word, then second word, etc. Numbers, short words (≤3 letters), punctuation and certain keywords (e.g., "food") are ignored for sorting. In the example, the sort is by "allyl" and then by "nitrosourea".
cas	Chemical-Abstract-Service registry number, when one is given. If there is no CAS number, this field is "___".

The cpdb.cit.xls dataset

papernum	Literature paper number. This field is used to merge with the cpdb.lit.xls datasets to retrieve brief citation information.
citation	The brief citation. May include personal communication as well as a journal or book citation.

Section 2: Datasets of code definitions

Datasets of code definitions

The cpdb.journal.xls dataset

jcode	the journal code. This field is used to merge with the cpdb.cit.xls dataset. The field "citation" in cpdb.cit.xls contains the journal code embedded in it.
jname	the name of the journal or book

The cpdb.route.xls dataset

route	the route code. These codes are also given above. This field is used to merge with the route field in the cpdb.ncintp.xls and cpdb.lit.xls datasets.
rtename	the name of the route

The cpdb.species.xls dataset

species	the species code. These codes are also given above. This field is used to merge with the species field in the cpdb.ncintp.xls and cpdb.lit.xls datasets.
spname	the name of the species

The cpdb.strain.xls dataset

strain	the strain code. These codes are also given in Appendix 1. This field is used to merge with the strain field in the cpdb.ncintp.xls and cpdb.lit.xls datasets.
strname	the name of the strain

The cpdb.tissue.xls dataset

tissue	the tissue code. These codes are also given in Appendix 2. This field is used to merge with the tissue field in the cpdb.ncintp.xls and cpdb.lit.xls datasets to obtain definitions. In the case of cpdb.lit.xls, a merge can be made directly. In the case of cpdb.ncintp.xls, multiple tissues may appear in the field, so a single tissue will have to be extracted before merging.
tisname	the name of the tissue

The cpdb.tumor.xls dataset

tumor	the tumor code. These codes are also given in Appendix 3. This field is used to merge with the tumor field in the cpdb.ncintp.xls and cpdb.lit.xls datasets. In the case of cpdb.lit.xls, a merge can be made directly. In the case of cpdb.ncintp.xls, multiple tissues may appear in the field, so a single tumor will have to be extracted before merging.
tumname	the name of the tumor

Section 3: Appendices

Appendix 1: Strains

Code	Strain
aaa	analbuminaemic (Sprague-Dawley derived)
aah	A/He
aap	Alpk/Ap
abi	Ab x IF
aci	ACI
agu	AGUS
aif	A x IF
ain	ACI/n
ajj	A/JJms
akr	AKR
aks	AKR/J
alb	albino
amm	A
aps	Alderly Park
asd	Sprague-Dawley albino
asp	ASH-CS1
asw	Swiss-Webster albino
aug	August
ays	AE/WffC3Hf/Nctr x YS/WffC3Hf/Nctr
b46	BR 46
b62	monohybrid cross offspring of $B6CF_1$ (C57BL/6 x BALB/c)
b6a	$B6AKF_1$
b6b	($B6C3F_1$ x B6C3 background, brachymorphic) inter se= $B6C3F_2$ brachymorphic
b6c	$B6C3F_1$
b6n	($B6C3F_1$ x B6C3 background, brachymorphic) inter se= $B6C3F_2$ phenotypically normal
baa	Black a/a (YS x VY)F_1
baj	BALB/cJ
bal	BALB/c
bbb	Bush babies [*Galago crassicaudatus*]
bbl	Bethesda black
bce	BALB/cHe
bcn	BALB/cStCrlfC3Hf/Nctr
bd1	BDF_1
bd2	BD II
bd9	BD IX
bdf	BD VI
bdr	BD
beg	beagle
bfm	Buffalo-Mai
bld	BALB/cLacDp
buf	Buffalo
c17	C17
c3c	C3H/AnCum
c3d	C3Hf/Dp
c3e	C3HeB/Fe
c3h	C3H
c3j	C3H/HeJ
c3l	C3H (C3H/Anl) (Anl 70)
c3p	C3HeB
c3s	C3H/St
c3v	C3H/HeN–MTV–/Nctr

c56	C57BL/6J
c5c	C57BL/10ScSn
c5j	C57BL/10J
c5l	C57BL
c5n	C57BL/6N
c5v	C57BL/BVI
c6s	C57BL/6CrSlc
c7b	(C57BL/6 x BALB/c)F_1
c7l	C57L
cb6	C57BL/6
cba	CBA
cbc	CBA/Cb/Se
cbh	CBA/H-T6
cbj	C3HeB/FeJ
cbl	C57BL
cbn	C57BL/6JfC3Hf/Nctr x BALB/cStCrlfC3Hf/Nctr inter se
cbo	C.B. hooded
cbr	CB
cbs	Cb/Sc
cbt	Chester Beatty albino
cd1	Charles River CD1
cdf	CDF_1
cdr	Charles River CD
cen	C3H/HeN
cf1	CF-1
cfe	CFE
cff	C57BL/6JfC3Hf/Nctr x BALB/cStCrlfC3Hf/Nctr
cfi	C3H/FIB
cfl	CFLP
cfn	CFN
cfr	CF
che	C57BL/He
chf	C3HfB
chg	C3H/He germfree
chh	C3H/He
chi	CD-1 HaM/ICR
chj	C3HeB/Jax
chm	Charles River
cif	(C57 x IF)F_1
clw	Colworth (Wistar derived)
crf	(C3H x RIII)F_1
crw	Charles River Crl:COBS(WI)BR
csa	Charles River albino
csb	CSb
csc	C57L/He x 129/Rr x C3HeB/De x SWR/Ly
ctn	CTM
cva	BALB/cStCrlfC3Hf/Nctr x VY/WffC3Hf/Nctr-(A/A)
cvy	BALB/cStCrlfC3Hf/Nctr x VY/WffC3Hf/Nctr-(A^{vy}/A)
cwf	Carworth Farms
cws	CFW
cym	Cynomolgus [*Macaca fascicularis*]
dba	DBA/2
dbx	DBA
ddd	DDD
ddn	ddNi
ddx	dd

ddy	DDY
don	Donryu
esd	Eastern Sprague-Dawley
f34	Fischer 344
f3d	F344/DuCrj
f3l	Fischer 344/LATI
fdr	FDRL
fds	Food and Drug Research Laboratories stock rats
fis	Fischer
fmf	Fischer 344/Mai fBR
hew	Hebrew University
hic	Ha/ICR
hra	HRA/Skh (hairless)
hrl	Harlan
hza	Holtzman albino (Sprague-Dawley derived)
ic3	ICRC x C3h (Jax)
ici	ICI
icm	ICR
icr	ICR/Jcl
ifc	IF x C57
ifm	IF
jic	JCL: ICR
leb	Long-Evans BLU: (LE)
lee	Leeds albino
lev	Long-Evans
mar	Marshall
mgr	mongrel
mrc	MRC
mrw	MRC-Wistar
nbr	NBR
nbw	NZBW (hooded black and white strain)
nmb	Bor:NMRI, SPF-bred NMRI
nmh	Han: NMRI
nmr	NMRI
non	non-inbred
nra	Norwegian albino
nss	not specified
nzb	NZO/BlGd
nzd	NZR/Gd
of1	OF1
ofs	OFA (Sprague-Dawley derived)
osm	Osborne-Mendel
por	MRC Porton (Wistar derived)
pva	Lean pseudoagouti Avy/a
r3m	RIII
rfm	RF
rhe	Rhesus [*Macaca mulatta*]
scd	Swiss CD-1
scp	Cpb:Swiss random
sda	Sprague-Dawley
sdz	Sandoz
shc	Sherman COBS
she	Sherman
shr	Swiss/H/Riop
sic	Swiss/ICR
sjs	SJL/J

sls	Slc-Wistar
smw	Sas: MRC(WI)BR
ssa	S strain albino
sss	Sprague-Dawley Spartan
stm	ST/a
swa	Swiss albino
swi	Swiss
swr	SWR
sww	Swiss Webster
syg	Syrian Golden
tf1	Tuck
the	Theiller's Original
tmm	TM
tst	Tree shrew [*Tupaia glis*]
wag	WAG
wal	Wistar albino
wi2	Wistar II
wid	Wistar/FDRL
win	Wistary/NIN
wio	Wistar-OSU
wis	Wistar
wmf	Wistar-Mai-Furth
wsh	Han: WIST
wsr	Wistar-random
wsw	Wilmslow Wistar
wws	Wistar W.74
xvi	XVII/G
yva	Obese yellow Avy/a

Appendix 2: Site codes

Code	Site
---	all target sites
abc	abdominal cavity
abd	abdomen
adr	adrenal gland
adu	acoustic duct
amd	adrenal medulla
aol	aorta and large arteries
arp	adrenal capsule
asc	colon, ascending
auc	external auditory canal
aur	auricular region
b/l	lung, bronchiole
bil	bile duct
blv	blood vessels
bmd	brain, medulla
bod	body cavities
bom	bone marrow
bon	bone
bra	brain
brf	brown fat, dorsal
brm	brain, meninges
brs	brain stem
ccx	cerebral cortex
cec	cecum

chp	cheek pouch
clb	cerebellum, cerebrum
cli	clitoral gland
clm	cerebellum, meninges
clr	colorectum
cns	central nervous system
col	colon
crb	cerebrum
crl	cerebellum
cst	cardiac stomach
cvu	cervix uteri
cvx	cervix
cyx	coccyx
der	dermis
dgt	digestive tract
dsc	colon, descending
duo	duodenum
eac	ear canal
ear	ear
edu	ear duct
ehp	extrahepatic tissue
eld	eyelid
epg	epiglottis
epi	epidermis
epy	epididymis
eso	esophagus
eye	eye
fat	fat
fgr	forestomach, greater curvature
fhd	forehead
fls	forestomach, lesser curvature
for	forestomach
frb	forebrain
gab	gall bladder/bile duct
gal	gall bladder
gam	gastric mucosa
git	gastrointestinal tract
gnv	gingiva
hag	Harderian gland
hea	heart
hnt	hard palate/nasal turbinates
hpl	hypophysis
hum	humerus
ilm	ileum
isp	interscapulum
itl	intestinal tract
itn	intestine
jej	jejunum
k/p	kidney/pelvis
kcx	kidney cortex
kid	kidney
kpp	kidney papilla
ktu	kidney tubule
kur	kidney/ureter
l/b	lung, bronchus
lar	larynx

lgi	large intestine
liv	liver
lmr	lymphoreticular system
lpp	lip
lun	lung
lyd	lymph node
mam	mammary tissue (other than or including more than mammary gland)
mds	mediastinum
mei	mesenteric intestine
meo	mesovarium
mey	mesentery
mgl	mammary gland
mix	more than one site; sites specified in published paper
mln	mesenteric lymph node
mls	multiple sites
mth	mouth
mul	multiple organs
mus	muscle
MXA	more than one site, combined by NCI/NTP
MXB	more than one site, combined by Berkeley
myc	myocardium
nac	nasal mucosa
nap	nasal passageway
nas	nasal cavity
ncp	nasal cavity, posterior region
ner	nervous system
nof	nasal cavity, olfactory epithelium
nol	n. olfactorius
npc	nasal and paranasal cavity
npl	nipple
nre	nasal cavity, respiratory epithelium
nse	nose
nsm	nasal septum
nsp	nasopharynx
ntu	nasal turbinate
olb	olfactory bulb
omt	omentum
opx	oropharynx
orc	oral cavity
orm	oral mucosa
ova	ovary
pae	pancreas, exocrine
pal	palate
pan	pancreas
pdu	pancreatic duct
pec	peritoneal cavity
pel	pelvis
pep	paraepididymal tissue
per	peritoneum
phr	pharynx
pit	pituitary gland
pls	palate, soft
pnd	pancreas/pancreatic duct
pni	pancreatic islets
pnl	paranasal sinus
pnr	peripheral nerves

pns	peripheral nervous system
pre	preputial gland
prn	pararenal tissue
pro	prostate
pta	pituitary gland, anterior
pty	parathyroid
rec	rectum
rel	reticuloendothelium
rep	reproductive tract
res	respiratory system
sbg	sebaceous gland
sev	seminal vesicle
sft	skin of foot and toe
skb	skin of back
skf	skin of flank
ski	skin
sku	skull
slg	salivary gland
smi	small intestine
spc	splenic capsule
spd	spinal cord
spl	spleen
spn	spinal nerves
srp	splenic red pulp
ssq	stomach, squamous
ssu	skin and subcutis
stg	stomach, glandular
stn	stomach, nonglandular
sto	stomach
sub	subcutaneous tissue
tba	all tumor bearing animals; for NCI/NTP interstitial-cell tumors of the testis are excluded for male rats
tes	testis
thi	thigh
thm	thymus gland
thx	thorax
thy	thyroid gland
tna	tunica albuginea
tnv	tunica vaginalis
ton	tongue
trh	trachea
tyf	thyroid follicle
ubl	urinary bladder
ugi	upper gastrointestinal tract
unt	urinary tract
ure	ureter
urt	urethra
ute	uterus
utm	uterus, endometrium
vag	vagina
ver	vertebra
vse	vascular epithelium
zym	Zymbal’s gland

Appendix 3: Histopathology

Code	Histopathology

---	all tumors
a/a	alveolar/bronchiolar adenoma
a/c	alveolar/bronchiolar carcinoma
abt	alveolar/bronchiolar tumor
aca	adenocarcinoma in adenomatous polyp
acb	alveolar/bronchiolar adenocarcinoma
acc	acinar-cell carcinoma
acn	adenocarcinoma, NOS
act	alveolar-cell tumor
ada	adenocarcinoma, type A
adb	adenocarcinoma, type B
adc	adenocarcinoma
ade	adenoma
adf	adenofibroma
adi	adenocarcinoma, bilateral
adm	adenomatous polyp, NOS or adenocarcinoma in adenomatous polyp
adn	adenoma, NOS
ado	adenoacanthoma
adp	adenomatous polyp
adq	adenosquamous carcinoma
aep	adenomatous endometrial polyp
agc	alveogenic adenocarcinoma
agm	angioma
agt	alveogenic tumor
ahs	axillary histiocytic sarcoma
akt	adenoma-like tumor
ala	alveolar-cell adenoma
alc	alveolar-cell carcinoma
ald	alveolar adenoma
amy	adenomyoma
ana	acinar-cell adenoma
anb	adenoma, bilateral
ane	angio-endothelioma, malignant
ang	angiosarcoma
aoc	acinar-cell adenocarcinoma
aod	adenocarcinoma, acinar or ductal
apc	anaplastic carcinoma
apn	adenomatous polyp, NOS
asl	astrocytoma, malignant
asm	adenocarcinoma with squamous metaplasia
ast	astrocytoma
ata	atypic adenoma
bca	basal-cell adenoma
bcc	basal-cell carcinoma
bcd	bronchiolar adenoma
bcp	basal-cell papilloma
bct	basal-cell tumor
bda	bile duct adenoma
bdc	bile duct carcinoma
bde	bronchiolar adenocarcinoma
bdt	bile duct tumor
ben	benign tumor
bhp	hepatoma, benign
bht	hepatocellular tumor, benign
blc	biliary cystadenoma
bly	B-cell lymphoma

bro	bronchogenic carcinoma
bsa	basophil adenoma
bsb	basosquamous tumor benign
bsn	basophilic nodule
caa	cholangioadenoma/carcinoma
cab	cholangiocellular tumor, benign
cac	cholangioadenocarcinoma
cad	cholangioadenoma
can	carcinoma, NOS
car	carcinoma
cas	carcinosarcoma
cca	c-cell adenoma
ccb	c-cell carcinoma, bilateral
ccn	cystadenocarcinoma, NOS
ccr	c-cell carcinoma
ccy	cholangioma, cystic
cdb	c-cell adenoma, bilateral
cgd	cholangiocarcinoma, ductular
cgf	cholangiofibroma
chc	cholangiosarcoma
cho	cholangioma
cic	carcinoma, in situ
cla	clear-cell adenoma
clc	cholangiocarcinoma
cnb	carcinoma, bilateral
cnd	carcinoid tumor, malignant
coa	cortical adenoma
coc	cortical carcinoma
con	cortical adenoma, NOS
cra	chromophobe adenoma
crc	chromophobe carcinoma
crn	cortical adenocarcinoma, NOS
crt	carcinoma, combined glandular and squamous type
csa	cortical subcapsular adenoma
cuc	ceruminous carcinoma
cvh	cavernous hemangioma
cyc	cystadenocarcinoma
cye	cystadenoma
cyn	cystadenoma, NOS
dhs	deep cervical, histiocytic sarcoma
ead	endometrium, adenoma
edc	endometrium, adenocarcinoma
emp	endometrial polyp
ena	endometrial adenocarcinoma
ene	esthesioneuroepithelioma
ens	endocardial sarcoma
epc	epidermoid carcinoma
epd	ependymoblastoma
epn	epithelial neoplasm
epo	epithelioma
ept	epidermoid tumor
esa	eosinophilic adenoma
esn	eosinophilic nodule
esp	endometrial stromal polyp
ess	endometrial stromal sarcoma
exa	exocrine adenoma

exp	exophytic papilloma
fab	follicular-cell adenoma, bilateral
fba	fibroadenoma
fbs	fibrosarcoma
fca	follicular-cell adenoma
fcc	follicular-cell carcinoma
fct	follicular-cell tumor
fcy	follicular-cell adenocarcinoma, bilateral
fdc	follicular adenocarcinoma
fep	fibroepithelial tumor
fib	fibroma
fih	fibrous histiocytoma
gcb	granulosa-cell tumor, benign
gcc	granulosa-cell carcinoma
gcl	granulosa-cell tumor, NOS
gcm	granulosa-cell tumor, malignant
gct	granulosa-cell tumor
ghc	hepatocellular carcinoma, glandular
glb	granulosa-cell tumor, bilateral
gli	glioma
gln	glioma, NOS
gmf	glioma malignant, focal, mild
grb	granular-cell tumor, benign
grl	granulocytic leukemia
gsa	granulocytic sarcoma
hae	hemangioendothelioma
hca	hepatocellular carcinoma/adenoma
hcs	histiocytic sarcoma
hct	hepatocellular tumor
hem	hemangioma
hes	hemangiosarcoma
het	hemorrhagic tumor
hga	hemangiosarcoma anaplastic
hmb	hemangioendothelioma, benign
hmm	hemangioendothelioma, malignant
hms	hemangioendothelial sarcoma
hmt	hamartoma
hnd	hyperplastic nodule
hpa	hepatocellular adenoma
hpb	hepatoblastoma
hpc	hepatocellular carcinoma
hpd	hepatocellular adenocarcinoma
hph	hepatocellular hyperplastic nodule
hpm	hemangiopericytoma, malignant
hpn	hepatocellular neoplastic nodule
hps	hepatocellular carcinoma, solid
hpt	hepatoma
iab	interstitial-cell adenoma, bilateral
ica	interstitial-cell adenoma
icb	interstitial-cell tumor, benign
ict	interstitial-cell tumor
ihs	iliac histiocytic sarcoma
ile	leukemia, indeterminate type
isa	islet-cell adenoma
isc	islet-cell carcinoma
ism	insuloma

itm	interstitial-cell tumor, malignant
ivc	carcinoma, invasive
ivt	transitional-cell carcinoma, invasive
kcs	Kupffer-cell sarcoma
ker	keratoacanthoma
lbl	lymphoblastic lymphoma
lca	liver-cell adenoma
lcb	liver-cell tumor, benign
lcc	liver-cell carcinoma
lcl	lymphocytic lymphoma
lcm	liver-cell tumor, malignant
lct	liver-cell tumor
ldc	Leydig-cell tumor
lei	leiomyosarcoma
leu	leukemia
ley	leiomyoma
lhc	lymphoma, histiocytic type
lip	lipoma
lkm	lymphoma leukemia
lkn	leukemia, NOS
lle	lymphocytic leukemia
lls	lymphoblastic leukemia-lymphosarcoma
lna	nonlymphocytic leukemia, acute
lpb	liver-cell tumor, type B
lps	liposarcoma
lsl	systemic and localized lymphoma
lut	luteoma
lyk	lymphatic leukemia
lym	lymphoma
lyp	lymphangioma
lys	lymphosarcoma
lyt	lymphoid tumor
mag	malignant glioma
mal	malignant tumor
mcc	mucinous carcinoma
mda	medullary adenoma
mdt	medullary tumor
mec	muco-epidermoid carcinoma
mem	mixed cell mucoepidermoid papilloma
men	mesothelioma, NOS
mfh	fibrous histiocytoma, malignant
mhb	hibernoma, malignant
mhc	mixed hepato/cholangio carcinoma
mhp	malignant hepatoma
mhs	histiocytoma, malignant
mht	hepatocellular tumor, malignant
mix	more than one tumor type; tumor types specified in published paper
mlc	melanocytoma
mle	monocytic leukemia
mlh	malignant lymphoma, histiocytic type
mlk	myelogenous leukemia
mlm	malignant lymphoma, mixed type
mlp	malignant lymphoma, lymphocytic type
mlt	melanotic tumor
mlu	malignant lymphoma, undifferentiated type
mly	malignant lymphoma

mng	meningioma
mnl	mononuclear-cell leukemia
mnm	meningioma, malignant
mno	malignant lymphoma, NOS
mnp	mesenchymal neoplasm
msb	mesothelioma, benign
msm	mesothelioma, malignant
mso	mesothelioma
mtb	mixed tumor, benign
mtm	mixed tumor, malignant
mua	mucinous adenocarcinoma
muc	mucinous cystadenocarcinoma
MXA	more than one tumor type, combined by NCI/NTP
MXB	more than one tumor type, combined by Berkeley
mye	myelocytic leukemia
myl	myeloid leukemia
myo	myelomonocytic leukemia
nen	neoplasm, NOS
neo	neoplasm
nep	nephroblastoma
neu	neuroblastoma
nfm	neurofibroma
nfs	neurofibrosarcoma
ngs	neurogenic sarcoma
nhs	inguinal histiocytic sarcoma
nim	neurinoma
nlm	neurilemoma, malignant
nnd	neoplastic nodule
nod	nodular hyperplasia
npm	neoplasm, NOS, malignant
nsc	neurosarcoma
nvc	carcinoma, noninvasive
nvt	transitional-cell carcinoma, noninvasive
oec	olfactory epithelial carcinoma
ogm	olfactory lobe, glioma malignant
olc	olfactory carcinoma
oli	oligodendroglioma
oln	olfactory neuroblastoma
olp	olfactory neuroepithelioma
onm	olfactory lobe, neuroblastoma malignant
ost	osteosarcoma
otm	osteoma
pac	papillary adenocarcinoma
pam	papilloma
pas	papillomatosis
pbb	pheochromocytoma benign, bilateral
pbm	pheochromocytoma, benign/malignant
pca	parenchymal adenoma
pcn	papillary cystadenocarcinoma, NOS
pcy	papillary cystadenoma, NOS
pda	pars distalis adenoma
pdc	pars distalis carcinoma
pfa	parafollicular-cell adenoma
phc	pheochromocytoma, complex
phe	pheochromocytoma
phm	pheochromocytoma, malignant

pla	polypoid adenoma
plc	plasmacytoma
pmb	pheochromocytoma malignant, bilateral
pms	papillary mesothelioma
pob	pheochromocytoma, benign
pol	polyp
ppa	papillary adenoma
ppc	papillary carcinoma
ppn	papilloma, NOS
ppp	papillary polyp
ptc	papillary transitional-cell carcinoma
ptm	papillary tumor
pvc	carcinoma, preinvasive
rab	renal tubule adenoma, bilateral
rac	renal tubule adenocarcinoma
rca	renal-cell adenoma
rcc	renal-cell carcinoma
rcs	round-cell sarcoma
rct	renal-cell tumor
ret	reticulum-cell tumor
rhb	rhabdomyosarcoma
rhm	rhabdomyoblastoma
rhs	renal, histiocytic sarcoma
rna	reticulum-cell neoplasm, type A
rsc	respiratory epithelial carcinoma
rta	reticulum-cell sarcoma, type A
rtb	reticulum-cell sarcoma, type B
rts	reticulum-cell sarcoma
rua	tubule adenoma
ruc	tubule carcinoma
rue	tubule epithelium adenoma
sad	scirrhous adenocarcinoma
sar	sarcoma
sbr	sebaceous gland carcinoma
sca	solid-cell adenoma
scc	spindle-cell carcinoma
scs	spindle-cell sarcoma
sct	Sertoli-cell tumor
sea	sebaceous adenoma
seb	sebaceous adenoma and adenocarcinoma
sec	sebaceous adenocarcinoma
sgc	sweat gland carcinoma
shs	mesenteric histiocytic sarcoma
spm	sarcoma, NOS
spt	spindle-cell tumor
sqa	squamous-cell tumor
sqc	squamous-cell carcinoma
sqi	squamous-cell carcinoma, invasive
sqk	squamous-cell carcinoma, keratinized
sqn	squamous-cell carcinoma, in situ
sqp	squamous-cell papilloma
sqs	squamous-cell carcinoma, stratified
squ	squamous-cell carcinoma, unclassified
srn	sarcoma, NOS
ssc	squamous-cell carcinoma, sebaceous
tcb	tubular-cell carcinoma, bilateral

tcc	transitional-cell carcinoma
tcm	thecoma
thc	hepatocellular carcinoma, trabecular
tla	tubular-cell adenoma
tma	thymoma
tpp	transitional-cell papilloma
tri	trichoepithelioma
tua	tubular adenoma
tuc	tubular carcinoma
tum	tumor or more than one tumor type; tumor types not specified in paper
uac	tubular-cell adenocarcinoma
ulc	undifferentiated carcinoma
ule	undifferentiated leukemia
utc	urothelial carcinoma
utp	urothelial papilloma
vlp	villous polyp
vsc	all vascular tumors