

# DSSTox Field Definition File:

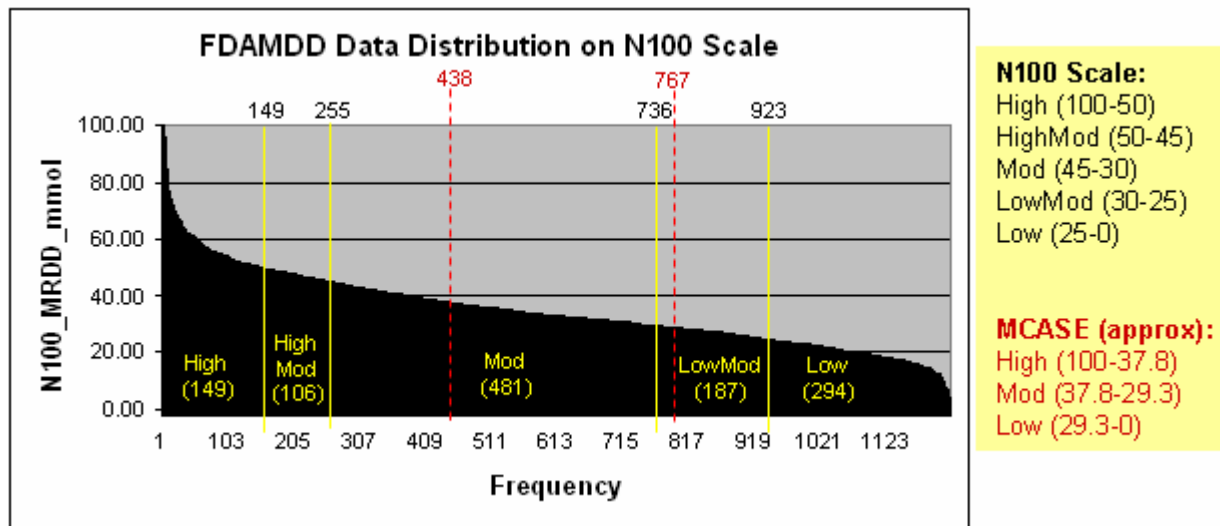
## FDA Maximum (Recommended) Daily Dose Database (FDAMDD)

(last updated 15 February 2008)

**Description:** Information in this file is intended to provide a minimum level of annotation to the DSSTox SDF (Structure Data Format) file created for the FDA Maximum (Recommended) Daily Dose (FDAMDD) database based on data obtained from the FDA Source and Main Citation listed below (Matthews et al., 2004). For further explanation of Source-specific fields, a user is encouraged to consult the listed references and information provided on the main DSSTox FDAMDD SDF Download Page [http://www.epa.gov/ncct/dsstox/sdf\\_fdamdd.html](http://www.epa.gov/ncct/dsstox/sdf_fdamdd.html). A number of modifications and additions to database content were made to facilitate use of the DSSTox SDF file in relational database applications and Structure-Activity Relationship (SAR) model development. All modifications are fully documented here and in the **Comments** section of the table below.

A significant enhancement in converting the FDA Source database to the DSSTox FDAMDD database was in the addition of CAS registry numbers (CASRN) for all main structures provided in the **STRUCTURE** field, as well as an listing of CASRN for related chemicals provided and described in the **FDAMDD\_Note** field. Many CASRN were initially provided by the FDA Source collaborators, without specific definition and with multiple CASRN provided for the majority of structure records in the FDAMDD. These multiple CASRN per record highlight a feature and potential limitation of this database in terms of the general assignment of MRDD activity values. In many cases, an MRDD value is assigned not to a single chemical structure, but rather to a family of related derivatives of a drug. Further complicating this interpretation is that in some cases multiple related derivatives of a drug are listed in separate structure records, but assigned the same MRDD activity value. Hence, the treatment of related derivatives of a drug sharing a common MRDD value is inconsistent within the database; in some cases collapsed to a single structure record, in other cases represented in multiple records. To alert a user to instances of the latter and to facilitate location of multiple records sharing the same **Dose\_MRDD\_mg** value, we include a field **ChemClass\_MRDD\_grouping** to indicate membership of multiple drug records within the same activity group (e.g., Estradiol), with the corresponding **ChemicalReplicateCount** field indicating a count of 2 or more records assigned the same **Dose\_MRDD\_mg** value within the **ChemClass\_MRDD\_grouping**. With the use of these 2 fields, future use or expansion of the FDAMDD database could list additional CASRN derivative structures as separate database records. Alternatively, a user could choose to collapse all instances of **ChemicalReplicateCount** greater than 1 to the main structure derivative record (i.e., **ChemicalReplicateCount=1**) to provide a single **Dose\_MRDD\_mg** value assignment and not artificially inflate the influence of a particular family of derivatives (and structural features) in the database.

FDAMDD includes a field **TherapeuticCategory**, obtained from two sources: Merck Index 12<sup>th</sup> Edition as a primary source and ACD Labs ChemFolder Dictionary, version 8.0, as a secondary source. Also, a number of activity fields are offered in addition to the original field (**Dose\_MRDD\_mg**) provided by the FDA Source. To facilitate future SAR analysis of the DSSTox FDAMDD, activity values were converted to millimolar equivalents, **Dose\_MRDD\_mmol**, inverted and converted to log scale, **LOGINV\_MRDD\_mmol**, and the latter mapped to an 0-100 scale, **ActivityScore\_FDAMDD** (formerly **N100\_MRDD\_mmol**), with 100 representing the most toxic chemical (most potent, smallest MRDD dose) and 1 representing the least toxic chemical (least potent, largest MRDD dose). To facilitate qualitative analysis, **ActivityScore\_FDAMDD** values were mapped to 5 gross activity categories, **ActivityCategory\_MRDD\_mmol**, consisting of *High, High-Moderate, Moderate, Low-Moderate, Low*. These activity category cutoffs were chosen by purely subjective means to provide reasonable numerical representation within each categorical range of activity (see figure below). These categories correspond closely, however, to the results obtained by clustering methods (choosing 5 clusters) using KNN or agglomerative nesting based on Ward's minimum variance (Chihae Yang, personal communication). To provide approximate correspondence of the **ActivityCategory\_MRDD\_mmol** assignment to the original Source MCASE modeling study (Main Citation), additionally included is the MRDD activity categorization, **MCASE\_Activity\_mg**, which is based on cutoff values applied to the original **Dose\_MRDD\_mg** activity values (see additional discussion below). Users should consult the Main Citation for further details.



Description of **DSSTox Standard Chemical Fields** can be found in the Central Field Definition Table located at: <http://www.epa.gov/ncct/dsstox/CentralFieldDef.html>

Previously, the **STRUCTURE\_Shown** field entry for each record in this database was listed as “*active ingredient of formulation*” and the **TestSubstance\_Description** entry was “*mixture or formulation*” since most drugs are administered as formulations in a clinical setting, and because this designation implied a higher degree of uncertainty in the association of an activity measure with a particular dose of an active drug form. However, to enable test substance comparisons across DSSTox files, these field entries have reverted to “*tested chemical*” and “*single chemical compound*”, respectively, with the following text added to the Source-specific field, **Note\_FDAMDD**: “*form of drug administered not known*”.

The first section of the Table below lists the **DSSTox Standard Toxicity Fields** employed for this database, followed by the **FDAMDD Source-Specific Fields** containing the toxicity information particular to FDAMDD. The **Field Type** indicates the type of the field, such as numeric, integer, defined text, memo, etc. All **Units** and **Descriptions** are extracted from Source reference materials unless otherwise noted. **Allowable Entries** lists allowed field entries occurring in FDAMDD, separated by slashes for exclusive entries (i.e., cannot occur with another entry) and semicolons or spaces for non-exclusive entries (i.e., can occur with other values). These are defined and explained in the **Description** section.

**Source Website:** Users may wish to consult the FDA MRDD Source Website [http://www.fda.gov/cder/Offices/OPS\\_IO/MRTD.htm](http://www.fda.gov/cder/Offices/OPS_IO/MRTD.htm). However, noting the substantial number of corrections and additions incorporated into the DSSTox version of the FDAMDD database, use of the DSSTox version is recommended.

**Source Contacts:** Edwin Matthews and R. Daniel Benz, FDA Center for Drug Evaluation Research, Rockville, MD; email: [edwin.matthews@fda.hhs.gov](mailto:edwin.matthews@fda.hhs.gov), [r.daniel.benz@fda.hhs.gov](mailto:r.daniel.benz@fda.hhs.gov)

**Main Citation:** Publications reporting use of DSSTox SDF file for the FDA Maximum Recommended Daily Dose database are asked to list the full DSSTox file name, including date stamp, and to cite as primary reference the following:

"Matthews, E.J., Kruhlak, N.L., Benz, R.D., and Contrera, J.F. Assessment of the health effects of chemicals in humans: I. QSAR estimation of the maximum recommended therapeutic dose (MRTD) and no effect level (NOEL) of organic chemicals based on clinical trial data, *Current Drug Discovery Technologies*, 2004, 1(1): 61-76.

\*pdf of Main Citation can be downloaded from the FDAMDD Database webpage located at the central DSSTox website:  
[http://www.epa.gov/ncct/dsstox/sdf\\_fdamdd.html](http://www.epa.gov/ncct/dsstox/sdf_fdamdd.html)

### SDF Usage Notes:

Each DSSTox SDF file contains a single **STRUCTURE** field. For each chemical record, the **STRUCTURE** field entry directly corresponds to the content of the **STRUCTURE\_...** fields. The **STRUCTURE\_Shown** field documents the relationship between what is displayed in the **STRUCTURE** field and the actual tested chemical substance, i.e. **TestSubstance\_...** fields, with the latter corresponding directly to the toxicity data field entries. Commercial chemical relational database (CRD) applications may automatically insert one or more structure identifier fields upon import or export of an SDF file (e.g., Formula, FW or Mol\_ID), fields that may augment or duplicate one or more of the DSSTox Standard Chemical Fields. Users are cautioned that fields containing null values in the first record of the SDF will be reordered upon import into most applications; for this reason, the word "blank" has been inserted into null fields in Record 1 of DSSTox SDF files and can be deleted after SDF import. Users are additionally cautioned that some fields (**STRUCTURE\_SMILES** and **STRUCTURE\_InChI**, in particular) may exceed the 200 character limit specified in the MDL CTfiles SDF standard (see <http://www.epa.gov/ncct/dsstox/MoreonSDF.html>), and that some CRD applications may insert a line break or truncate these fields upon SDF import or export. Finally, CRD application-specific molecular header information in the SDF file is deleted in the final DSSTox SDF files; users running CRD applications requiring a unique molecule header upon import of the SDF can specify either **DSSTox\_RID** or the **DSSTox\_FileID** be used. Upon SDF import, **DSSTox\_CID** can be used to identify and manage chemical structure duplicates and **DSSTox\_Generic\_SID** can be used to identify common Test Substances across and within DSSTox files (similar to CASRN-substance, but available for all DSSTox substances and further distinguishes among different purity/grade substances).

As an MS Word document, the following table is best viewed onscreen using either Normal or Web Layout View in Landscape page orientation.

Field Name	Field Type	Units	Allowable Entries	Description	Comments
<b>DSSTox Standard Toxicity Fields</b>					
<b>Study Type</b> (no spaces)	defined text		Clinical Reports	Field is used to label all records in the database, generally with the same entry, and is designed to facilitate record identification for cross-database structure searching. Field entry refers to the main type of toxicity study for which data is represented in the database.	Field names and content are being coordinated with the public ToxML standardization effort.
<b>Endpoint</b>	defined text		Maximum Recommended Daily Dose	Field is used to label all records in the database, generally with the same entry, and is designed to facilitate record identification for cross-database structure searching. Field entry refers to the type of toxicity measure represented within the database.	Field names and content are being coordinated with the public ToxML standardization effort.
<b>Species</b>	defined text		human	Field is used to label all records in the database, generally with the same entry, and is designed to facilitate record identification for cross-database structure searching. Field entry refers to the species of animal(s) listed in the data record and used in the toxicity study or studies.	Field names and content are being coordinated with the public ToxML standardization effort.
<b>FDAMDD Source-Specific Fields</b>					
<b>ChemicalReplicateCount</b> (no spaces)	defined text		1/ # of #total	Counter field specifying instances of replicates or related forms in the database (i.e. structurally related chemicals assigned the same activity). Entry is "1" in first case of unique substance, parent structure, or 2D	In a number of cases, FDAMDD assigns the same activity value, <b>Dose_MRDD_mg</b> , to multiple records for closely related chemical derivatives. If a chemical is a member of such a

				<p>structure. If replicates or related forms exist, entry is a counter number (1,2,3, etc) followed by "of" and the total number of replicates or related forms for that case., e.g.:</p> <p>1 of 3, 2 of 3, 3 of 3, are 3 record entries in a case of 3 replicates.</p>	<p>derivative family, the generic name is listed in the <b>ChemClass MRDD_grouping</b> field. For each member of this derivative family sharing a common <b>ChemClass MRDD_grouping</b> designation, a <b>ChemicalReplicateCount</b> numbering is provided. Only two sets of compounds have replicate parent forms (Bephenium and Homatropine) and only two sets of compounds have replicate CAS (lophendylate, m and o, and Iodinated Glycerol A and B). <i>Previously listed as a DSSTox Standard Chemical Field in v2a.</i></p>
<b>ChemClass_MRDD_grouping</b> <i>(no spaces)</i>	defined text		Astromicin Bephenium Betamethasone Cefamandole .../ <i>blank</i>	Non- <i>blank</i> entry only when closely related chemical derivatives are a member of a group assigned the same MRDD activity, <b>Dose_MRDD_mg</b> , within the database. All members of the group are assigned a common <b>ChemClass_MRDD_grouping</b> name and <b>ChemicalReplicateCount</b> values ranging from "1 of n" to "n of n", where n is the number of derivatives included in the <b>ChemClass_MRDD_grouping</b> .	See Comments for <b>ChemicalReplicateCount</b> .
<b>Therapeutic Category</b> <i>(no spaces)</i>	text		e.g. analgesic, anti inflammatory, antipyretic, carbonic anhydrase inhibitor, cardiotonic, ...	Therapeutic categories for use of drug as listed in Merck Index, 12 <sup>th</sup> Ed.	Where therapeutic category was not found in Merck Index, was obtained from ACD Labs Dictionary (version 8.0) or on-line literature. In general, therapeutic category bears no relationship to the types of adverse effects or toxicity observed at doses exceeding the MRDD values.

<b>Dose_</b> <b>MRDD_mg</b> <i>(no spaces)</i>	numeric	mg/kg- bw/day	#	Maximum recommended daily dose (or maximum recommended therapeutic dose) values were determined from pharmaceutical clinical trials that employed an oral route of exposure and daily treatments, usually for 3-12 months. Drugs were given as single or divided dose treatment regimens to achieve desired pharmacological effects. Roughly 5% of the pharmaceuticals in the FDAMDD database were antineoplastics and anesthetics and were administered intravenously and/or intramuscularly. When separate MRDDs were reported for different routes of exposure, only the oral MRDD was included in the database and only MRDD values reported for the average adult patient were used. Pharmaceuticals that are administered orally are usually tested over a limited range of doses and have MRDDs reported as mg/day. The mg/day unit was converted to mg/kg-body weight (bw)/day based upon an average adult weighing 60 kg. In contrast, the dose unit for most antineoplastic drug MRDDs is reported as mg/m <sup>2</sup> which was converted to mg/kg-bw/day using the formula mg/kg-bw/day = mg/m <sup>2</sup> /37 for an average adult. Additionally, a few drugs had MRDDs reported in parts per million (ppm) which were converted to mg/kg-bw/day on the basis that 1000 ppm equals 25 mg/kg-bw/day for an average 60 kg adult.	MRDD values were extracted from <i>Martindale: The Extra Pharmacopoeia</i> (1973, 1983, and 1993) and <i>The Physicians' Desk Reference</i> (1995 and 1999).
<b>Dose_</b> <b>MRDD_mmol</b> <i>(no spaces)</i>	numeric	mmol/ kg- bw/day	#	Daily dose measure converted to millimoles: <b>Dose_MRDD_mmol = Dose_MRDD_mg / MolecularWeight</b>	Note that this conversion of mg to mmol in FDAMDD assumes that the compound dose in mg corresponds to the active ingredient of a formulation.
<b>LOGINV_</b> <b>MRDD_mmol</b> <i>(no spaces)</i>	numeric	log(1/( mmol/k g- bw/day) )	#	<b>LOGINV_MRDD_mmol =</b> $\text{Log}_{10}(1/\text{Activity}) = \text{Log}_{10}(1/\text{Dose\_MRDD\_mmol})$	$\text{Log}_{10}(1/\text{Activity})$ measure is the standard measure of activity used in QSAR modeling studies
<b>ActivityScore_</b> <b>FDAMDD</b> <i>(no spaces)</i>	integer		INTEGER[1-100]	Mapping of <b>LOGINV_MRDD_mmol</b> spanning activity range [MIN, MAX] onto Integer 1-100 Activity range, where 100 is highest potency and 1 is lowest potency:  ActivityScore = 100 * $\text{INTEGER}[(\text{LOGINV\_MRDD\_mmol} - \text{MIN})/(\text{MAX} - \text{MIN})]$	Summary activity for use in PubChem and structure-activity relationship studies.  <i>Field modified in v3b, formerly</i> <b>N100_MRDD_mmol</b>

<b>ActivityCategory_MRDD_mmol</b> <i>(no spaces)</i>	defined text		High/ High-Moderate/ Moderate/ Low-Moderate/ Low	<p>Based on <b>ActivityScore_FDAMDD</b> 100 scale, provides qualitative estimate of activity or potency of chemical within database; <b>ActivityScore_FDAMDD</b> range listed along with <b>Dose_MRDD_mmol</b> range for each activity category:</p> <p>High (100-50) = (1.3E-8 to 3.6E-4 mmol)  High-Moderate (50-45) = (3.7E-4 to 1.0E-3 mmol)  Mod (44-30) = (1.1E-3 to 2.2E-2 mmol)  Low-Moderate (29-25) = (2.3E-2 to 6.3E-2 mmol)  Low (24-0) = (6.4E-2 to 1.1E+1 mmol)</p>	<p><b>ActivityCategory_MRDD_mmol</b> cutoffs were chosen for FDAMDD by purely subjective means to provide reasonable numerical representations of chemicals in different activity ranges (see FDAMDD Data Distribution Figure above); different cutoffs could be defined by other criteria.</p> <p>Previously listed as HighModerate and LowModerate.</p>
<b>ActivityCategory_MCASE_mg</b> <i>(no spaces)</i>	defined text		High/ High-Moderate/ Moderate/ Low	<p>The Main Citation (Matthews et al, 2004) reported two different sets of activity classification cutoffs used for separate MCASE analyses of “Inactives” and “Actives”. These cutoffs correspond reasonably closely between High and Marginal categories, and almost exactly between Marginal and Low categories in that study. To assist users in identifying both “Active” and “Inactive” classifications from that earlier study, we assigned <b>ActivityCategory_MCASE_mg</b> values to <i>High</i>, <i>High-Moderate</i>, <i>Moderate</i>, and <i>Low</i>, where the <i>High-Moderate</i> category is the only set of compounds that are categorized differently in the two MCASE analyses:</p> <p>High (0.00001-2.49 mg)  High-Moderate (2.5-2.67 mg)  Moderate (2.7-4.87 mg)  Low (5-999 mg)</p>	<p><b>ActivityCategory_MCASE_mg</b> categories are based on the <b>Dose_MRDD_mg</b> measure and, hence, only approximately correspond to the <b>ActivityCategory_MRDD_mmol</b> classifications based on <b>ActivityScore_FDAMDD</b>. We provide the <b>ActivityCategory_MCASE_mg</b> measure in the DSSTox FDAMDD only for historical comparison purposes.</p> <p>Previously listed as HighModerate.</p>
<b>Note_FDAMDD</b>	memo		Text	<p>Field used to provide supplementary Source-specific information pertaining to the chemical and toxicity fields.</p>	<p>Text inserted in this field to replace previous <b>STRUCTURE_Shown</b> entry of “active ingredient in formulation” and <b>TestSubstance_Description</b> entry of “mixture or formulation”: “form of drug administered not known”.</p> <p>FDAMDD-specific information pertaining to the chemical record previously included in <b>ChemicalNote</b> field has been moved to this field.</p> <p>Provided herein by the FDA Source are a number of additional CAS numbers that may apply to the listed MRDD activity. Wherever we were able to identify the additional CAS derivatives, we provide brief description. If a chemical is a member of a <b>ChemClass_MRDD_grouping</b>, and is not the first listed member, only the CAS of the first listed member is provided; CAS for “related forms”, assigned the same MRDD value, are provided only for the first listed member of the group.</p> <p>Wherever we were able to identify the additional CAS derivatives, we provide brief description in this field. Where a CAS provided by Source could not be verified, it has been removed. In</p>

					<p>some cases, an unidentified CAS may be "retired", i.e., an older CAS that was replaced with newer CAS number.</p> <p>Alternate chemical name to <b>TestSubstance_ChemicalName</b> provided only in cases where a different primary name was listed in Merck Index, 12th Ed, when we had difficulty locating a structure with the Source-provided name, or when a second name was particularly well known or highly informative.</p>