Support Document to the 2022 Substance Priority List (Candidates for Toxicological Profiles)

Agency for Toxic Substances and Disease Registry
Division of Toxicology and Human Health Sciences
Atlanta, GA 30333

December 2022

CONTENTS

SUPPORT DOCUMENT TO THE
2022 SUBSTANCE PRIORITY LIST
(Candidates for Toxicological Profiles)

Background Statement

The Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA or Superfund), as amended by the Superfund Amendments and Reauthorization Act of 1986 (SARA), requires that ATSDR and the U.S. Environmental Protection Agency (EPA) prepare a Priority List of Hazardous Substances commonly found at facilities on the CERCLA National Priorities List (NPL). The Priority List of Hazardous Substances (also called the Substance Priority List) includes substances that have been determined to be of greatest public health concern to persons at or near NPL sites. CERCLA, as amended, also requires that the Substance Priority List be revised periodically. Each substance on the list is a candidate to become the subject of a toxicological profile prepared by ATSDR and subsequently a candidate for the identification of priority data needs. This Support Document describes the methodology ATSDR used to generate the Substance Priority List. The priority list is posted at www.atsdr.cdc.gov/SPL. Substance Priority List resources, including a spreadsheet containing expanded data on the candidate substances, are available at www.atsdr.cdc.gov/SPL/resources.

METHODOLOGY USED TO GENERATE THE 2022 SUBSTANCE PRIORITY LIST

1. OVERVIEW

The ranking of hazardous substances on the Substance Priority List is based on three scored criteria, which are combined for a total score. The three criteria are:

- FREQUENCY OF OCCURRENCE AT NPL SITES – A substance must be identified at 3 or more NPL hazardous waste sites or facilities to be considered a candidate for this list. ATSDR's science database is the source of this data. Presence in at least one environmental medium per NPL site constitutes one occurrence. See Section 2.

- TOXICITY - If available, final reportable quantities (RQs) are used to assess the toxicity of candidate substances during the listing activity. If a final RQ is not available, the RQ methodology is applied to candidate substances to establish a Toxicity/Environmental Score (TES). This process is only used in scoring the substances with respect to their toxicity and the results should not be interpreted as regulatory values. See Section 3.

- POTENTIAL FOR HUMAN EXPOSURE - The exposure component is based on two parts: (1) the concentration of the substances in environmental media at sites, and (2) the exposure status of populations at sites. ATSDR's science database is the source of this data, with concentration and exposure data obtained from ATSDR public health assessments and health consultations. See Section 4.

Using these three criteria, the hazard potential of each candidate substance is ranked according to the following algorithm:

TOTAL SCORE	=	*NPL FREQUENCY*	+	*TOXICITY*	+	*POTENTIAL FOR HUMAN EXPOSURE*
(1,800 *max. points*)		(600 *points*)		(600 *points*)		(300 *concentration* + 300 *exposure points*)

The algorithm generates a candidate list of substances ordinally ranked based on their total score. The top 275 scoring substances becomes the Substance Priority List. Currently, there are 862 candidate substances for the priority list.

1.1 Criteria for Inclusion on the Substance Priority List

Substances considered for the 2022 Substance Priority List came from the array of substances present at NPL sites, as indicated in ATSDR's science database from either health assessment or site file information. Currently, over 3,400 uniquely identifiable substances are found at hazardous waste sites according to the ATSDR database. Only those substances found at three or more NPL sites were considered for the priority list; 862 substances were found at three or more sites.

The list of candidate substances was reviewed to identify petroleum-related substances. Substances of petroleum origin are regulated by legislation other than CERCLA [see CERCLA Section 101(14)]. These substances were assigned total point scores of -1 to place them at the bottom of the list of candidate substances.

2. DETERMINATION OF THE FREQUENCY OF OCCURRENCE CRITERION

2.1 Overview

ATSDR's science database is the source of data on the frequency of occurrence of substances at NPL hazardous waste sites or facilities. The sources of this site-specific information include ATSDR public health assessments and health consultations and other site-specific documents submitted to ATSDR by EPA, state agencies, and other parties. ATSDR has information on 1,868 sites that have been proposed for, listed on, or delisted from the NPL.

ATSDR's database contains information on substances found in various environmental media. The frequency of occurrence is based on the number of NPL sites at which the substance was found in any environmental medium. Included are substances that were positively identified at the site as a result of chemical analyses (i.e., at concentrations above the limits of detection), inventories, or other documentation collected during the ATSDR health assessment process. Substances identified in documents as "Tentatively Identified Compounds" (TICs) are not included in the Agency's database and, therefore, were not considered. Presence of a substance in at least one environmental medium per NPL site constitutes one occurrence.

2.2 Frequency of Occurrence Scoring

The frequency of occurrence component of the algorithm was assigned a maximum score of 600 points. These points were distributed between the maximum and minimum frequencies, with the maximum frequency receiving 600 points. Lead had the highest frequency of 1,299 NPL sites and therefore received 600 frequency points. The assignment of points for the remainder of substances was calculated using the following formula:

$$\frac{\textit{Current substance's frequency}}{\textit{Maximum frequency}} \times 600$$

For example, if a substance's NPL frequency = 845, then its frequency points = (845/1,299) x 600 = 390.

This method of point assignment was used in an effort to scale the measured frequency values into the allotted point range of 1-600, while maintaining their proportional relationship. As mentioned in Section 1.1, only those substances found at three or more NPL sites were considered for the priority list.

3. DETERMINATION OF THE TOXICITY COMPONENT

3.1 Overview

The reportable quantity (RQ) approach is used as the toxicity hazard scoring system for several reasons. This approach provides the most complete characterization of toxicity of all hazard scoring systems reviewed; other schemes were more limited in either the consideration of different types of toxic effects, severity of effects, or potency. In addition, toxicity data used in the RQ approach are derived from primary peer-reviewed literature, and RQs have already been established for the majority of substances that are frequently detected at hazardous waste sites. Moreover, the determination of RQ health effect values uses weight-of-evidence considerations in evaluating data.

The reportable quantity ranking scheme was developed by EPA to set RQs for hazardous substances as required by CERCLA. Section 103(a) of CERCLA requires that the National Response Center and state and local response authorities be notified immediately if a hazardous substance has been released in a quantity that equals or exceeds its RQ. EPA develops RQs for individual chemicals and for waste streams that have already been designated as hazardous substances under CERCLA, Section 101(14).

EPA assigns each CERCLA hazardous substance to one of five tiered RQ categories (1, 10, 100, 1,000, and 5,000 pounds) based on acute toxicity, chronic toxicity, carcinogenicity, aquatic toxicity, and ignitability and reactivity. RQs are determined separately for each criterion; the lowest of these is selected as the RQ for the substance, subject to adjustment for potential biodegradation, hydrolysis, or photolysis (BHP) in the environment. The RQ scoring scheme is described in the following four *Federal Register* notices: 50 FR 13456, April 4, 1985; 51 FR 34534, September 29, 1986; 52 FR 8140, March 16, 1987; 54 FR 35988, August 30, 1989. A brief overview is discussed at www.epa.gov/epcra/reportable-quantity-rq-adjustment-methodology.

ATSDR applied the RQ methodology to those candidate substances without final CERCLA RQs in order to establish a toxicity/environmental score (TES). These scores were developed for use only in the ranking methodology and do not represent regulatory values. TESs have been assigned to more than 450 candidate substances. Substances that received a TES greater than 5,000 (using the RQ methodology) were dropped to the bottom of the candidate list because of their lack of known toxicity and received a total score of zero points. A breakdown of the TESs developed for candidate substances and an overview of the toxicity scoring methodology is provided on the substance priority list resource webpage at www.atsdr.cdc.gov/SPL/resources.

3.2 Sources of Information Used to Determine the Toxicity/Environmental Score (TES)

Several sources of information on toxicity, reactivity/ignitability, and environmental fate were used to determine the TESs for substances lacking RQs. The National Library of Medicine (NLM) databases were the main sources of this information. These databases include the Hazardous Substances Data Bank (HSDB), Chemical Carcinogenesis Research Information

System (CCRIS) and Toxicology Literature Online (Toxline) database. In addition, the Registry of Toxic Effects of Chemical Substances (RTECS) database and EPA's Integrated Risk Information System (IRIS) and Ecotoxicology databases (ECOTOX) were also used.

3.3 Assumptions Used in Determining the Toxicity/Environmental Score

3.3.1 Ignitability/Reactivity. Where no specific values were found to express potential for ignitability/reactivity, professional judgment was applied. For example, if a substance was classified as extremely flammable, but no flash point was given, a score of 10 was assigned for the ignitability/reactivity component. Similarly, if no information was found to indicate the substance was ignitable or reactive, the substance was assigned a score of >5,000 for this component of the TES.

3.3.2 Aquatic Toxicity. Specific aquatic toxicity data were not found in the above information sources for many substances. In some of these cases, the reference text of Sax's Dangerous Properties of Industrial Materials was used to assess aquatic toxicity. The standard method of reporting aquatic toxicity in this text provides a range of toxicity without identifying the test species. Seventy-five percent of the maximum value was used for the aquatic toxicity component for substances that lacked any other source of aquatic toxicity information (for example, if the LC_{50} range was 100-1,000, a value of 750 was used).

3.3.3 Chronic Toxicity. Some substances lacked chronic toxicity data in the NLM online databases but were mentioned as having developmental or reproductive effects at a specified dose. For these substances, the developmental or reproductive effects were used to assess the chronic toxicity component because these effects are given the highest effect ranking (R_e in the RQ methodology) and potentially occur, regardless of duration of exposure.

3.3.4 Carcinogenicity. Substances classified by EPA or the International Agency for Research on Cancer (IARC) in cancer classification groups A, B, or C were assigned TES scores of 1, 10, or 100, respectively. Substances with limited evidence of carcinogenicity in animals, but not classified by IARC or EPA for carcinogenicity, were assigned a TES score of 100. Substances with evidence of carcinogenicity in animals but noted in the data source as "lacking sufficient evidence for carcinogenicity" by EPA or IARC were not evaluated for carcinogenicity (group D - insufficient evidence). Substances for which no information on carcinogenicity could be located were not evaluated for carcinogenicity.

3.3.5 Radionuclides. The RQs for radionuclides are expressed in curies (seven tiered categories), whereas other RQs are expressed in pounds. To provide comparative values and consistency in this activity, the 7-tiered categories of radionuclide RQs (in curies) are distributed into the toxicity point scale (see Section 3.4) so that the most harmful radionuclides receive the highest number of toxicity points and the less harmful radionuclides receive a lower number of toxicity points. Radionuclides with an RQ equal to 0.001 curie, 0.01 curie, or 0.1 curie receive a TES of 1 and receive the highest number of 600 points for the toxicity component. Radionuclides with an RQ equal to 1 curie receive a TES of 10 (400 toxicity points); 10 curies receive a TES of 100 (178 toxicity points); 100 curies receive a TES of 1,000 (53 toxicity points); and 1,000 curies

receive a TES of 5,000 (10 toxicity points). This method of point assignment allows the list to distinguish between the more harmful radionuclides (such as plutonium-238) and less harmful radionuclides (such as krypton-85).

3.3.6 Substances Lacking Data. For several substances, essentially no relevant information was located. In these cases, TESs were assigned based on the RQs for structurally related substances.

3.4 Toxicity Component Scoring

A scoring system using a 2/3 cumulative exponential decay is the scoring method for the toxicity component of the priority list. Using this scoring system, the toxicity points value is equal to 2/3 raised to the exponent of the cumulative ordinal rank, multiplied by 600 (the highest value for the toxicity points). The point assignments are presented in Table 4.

Table 4. Toxicity Component Scoring

RQ or TES	Ordinal Rank	Cumulative Ordinal Rank (COR)	2/3 Raised to Exponent of COR	Toxicity Points ($2/3^{COR}$ x 600)
1	0	0	1.0000	600
10	1	1	0.6667	400
100	2	3	0.2963	178
1,000	3	6	0.0878	53
5,000	4	10	0.0173	10

4. DETERMINATION OF THE POTENTIAL FOR HUMAN EXPOSURE COMPONENT

The exposure component of the algorithm is based on two factors: (1) concentrations of the substances in environmental media at sites and (2) exposure status of populations at sites as described in ATSDR health assessments or consultations. These two parts of the potential-for-human-exposure portion of the algorithm were assigned a maximum of 300 points each. If no concentration or exposure data were available for the substance, no points were assigned.

4.1 Concentrations of the Substances in Environmental Media

4.1.1 Overview. To provide a means of ranking substances based on concentration data, the following formula for calculating a relative source contribution (SC) was used.

$$SC = \frac{(\overline{C}_a A_a) + (\overline{C}_w A_w) + (\overline{C}_s A_s)}{RQ \text{ or } TES}$$

Where $\overline{C}_x$ = geometric mean of maximum concentrations of the substance in a particular environmental medium (a = air, w = water, s = soil); A_x = standard exposure assumption for the

particular environmental medium to approximate a theoretical daily dose to humans (e.g., 1 liter of drinking water consumed per day - see Section 4.1.4); and RQ or TES = the Reportable Quantity or Toxicity/Environmental Score for the substance.

The calculation of the source contribution was included in the methodology to distinguish between those substances that occur at low concentrations but are highly toxic and those substances that occur at higher concentrations but are relatively less toxic.

Note: Because of the complexity and uncertainty associated with calculating a daily dose for radioactive substances and asbestos compounds, source contribution values were not calculated for these substances.

4.1.2 Source of Concentration Data. ATSDR's science database served as the source of concentration data for NPL site contaminants. The database contains concentration data for hazardous substances that are documented in ATSDR health assessments and health consultations for NPL (as well as non-NPL) hazardous waste sites. This concentration data represents the maximum concentration found in a particular environmental medium at a specific site. Concentrations were converted to standard units for calculating the estimated daily dose. The media and submedia used as sources of concentration data are presented in Table 5.

Table 5. Types of Media Used as Sources of Concentration Data

Media Type	Submedia
Ground Water	Private Well
	Public Well
	Public Supply
	Tap
Surface Water	Public Supply
	Sample Station
	Tap
	Other
Soil	Surface
	Subsurface
	Unspecified
Air	Indoor
	Outdoor

4.1.3 Calculation of the Geometric Mean of Maximum Concentrations. Since the majority of concentration data in the Agency's database represent the maximum concentration found per environmental medium, the geometric mean calculated in this listing process represents the geometric mean of the maximum concentrations found per medium. Substances were evaluated

per environmental medium and the geometric mean for these maximum concentrations was calculated for all water, soil, or air data across all sites.

The geometric mean was chosen over other methods to calculate average concentrations because environmental concentrations often vary over many orders of magnitude and approximate a normal distribution when in log form. Units for geometric mean concentrations were converted to milligrams per kilogram (mg/kg) for soil concentrations, milligrams per liter (mg/L) for water concentrations, and milligrams per cubic meter (mg/m^3) for air concentrations. Air particulates were converted from parts per million (ppm) using the molecular weight of the substance in the calculation. Conversion to standard units per medium allowed a comparison of all substances under consideration for the substance priority list.

4.1.4 Calculation of Theoretical Daily Dose. Standard exposure assumptions for children (1 liter of water consumed per day, 200 milligrams of soil ingested per day, and 15 cubic meters of air breathed per day) were used to assist in the determination of a theoretical daily dose. These exposure assumptions were multiplied by the geometric mean concentration for their respective media, and then added together to determine the theoretical daily dose. The theoretical daily dose is equal to the numerator of the source contribution formula (see Section 4.1.1).

4.1.5 Source Contribution Scoring. This component received 300 maximum points. Source contributions (SC) for all substances were scored based on the logarithm of their SCs, because the SCs (which cover many orders of magnitude) approximate a normal distribution when in log form. A "cutoff" of two geometric standard deviations (GSDs) from the geometric mean (GM) was also used (see Table 6). This allows for better discrimination of the individual data points because: 95% of the data within two GSDs of the GM is more widely distributed across the 300 SC points that are available, average (i.e., GM) values fall in the center of the distribution, and particularly high or low outliers neither draw the average away from the center, nor reduce the spread of points available for the majority of substances. Points were assigned using the following formula:

$$\frac{(\text{log of substance's SC} - \text{log of SC low cutoff})}{(\text{log of high SC cutoff} - \text{log of low SC cutoff})} \times 300$$

SCs below the minimum cutoff (GM - 2 GSD) received 0 SC Points, and SCs above the maximum cutoff (GM + 2 GSD) received 300 points:

Table 6. SC Mean and Cutoffs

Minimum SC Cutoff	Source Contribution Geometric Mean	Maximum SC Cutoff
3.2E-8	2E-4	1.2E0

4.2 Exposure Status of Populations

4.2.1 Overview. Information concerning documented exposure or potential exposure to a particular substance, or to environmental media in which a substance was found, was also used in the exposure component. In this component, the number of reported occurrences of exposure to a substance, or exposure or potential exposure to any media containing a substance, were counted. The Agency database provides information obtained from ATSDR health assessments and health consultations on exposure or potential exposure to specific substances and to media, such as drinking water, in which substances have been reported. Substances were scored differentially with respect to identification of exposure to a particular substance, or of exposure or potential exposure to an environmental medium (see Table 7).

Exposure to contaminant (Category 1) counts the number of times that a substance was explicitly mentioned as being in a Completed Exposure Pathway (CEP) in an ATSDR site document. This column is similar to the CEP Site Count Report, except that this is the number of pathways in which the substance appeared, not the number of sites. (A substance can appear in more than one pathway at a site.)

The two exposure to medium values (Categories 2 and 3) count the number of times that a substance might implicitly have led to exposure at a site. "Implicit" means that an ATSDR site document might have stated that people drank groundwater via wells, and it also stated that the substance had been found in groundwater – but the document did not explicitly state that the substance was found in a groundwater completed exposure pathway, per se. (For example, it was in groundwater far from a well that was used.) The implicit exposure columns are essentially artificial constructions that lend more discriminatory data points particularly to any substances with little or no exposure data. Potential exposure to medium (Category 3) means that the document mentioned potential exposure to the medium pathway, whereas there was clear exposure to the medium for Category 2.

4.2.2 Exposure Status Scoring. The assignment of points to each of these three categories is presented in Table 7. Information on all the exposure categories was assessed. If there were positive occurrences in Category 1 (exposure to contaminant, also called a completed exposure pathway), then that category was considered the prevailing exposure and the substance was scored based on that exposure status. If there were no occurrences in Category 1, then Category 2 (exposure to medium) was used to assign exposure points; if there were no occurrences in Category 1 or 2, then Category 3 was used.

A maximum of 300 points was possible for this part of the algorithm. Points within each category were distributed from the highest to the lowest exposure count instances, with the maximum exposure receiving 300 points. Lead had the highest exposure count in Category 1 of 638, and therefore received 300 exposure points. The assignment of points for the remainder of substances was calculated using the following formula:

$$Exposure\ points = \left[\frac{Substance's\ exposure\ count}{Max.\ exposure\ count}\ X\ (Max.\ allowed\ points - Min.\ allowed\ points)\right] + Min.\ allowed\ points$$

The Max. and Min. allowed points correspond to the specific prevailing category for the substance (see Table 7). For example, if a substance's prevailing exposure count (from Category 1) equals 197, then its exposure points = [(197/638) x 100] + 200 = 230.

Table 7. Exposure Status Scoring

Exposure Status	Point Range Assignment
Category 1, Exposure to Contaminant	300 - 200
Category 2, Exposure to Medium	200 - 100
Category 3, Potential Exposure to Medium	100 - 1

Note that the design of the algorithm effectively makes it unlikely for high scores to appear in Category 2 or 3. This is because a substance that is found in numerous media pathways at numerous sites is also likely to have occurrences of exposure to the substance (Category 1). Thus, its Category 1 score prevails over its Category 2 or Category 3 score, as discussed. Due to this “masking” effect, only substances with exposure via media at a few sites have Category 2 or 3 scores that are not masked by Category 1 occurrences. Therefore, exposure point scores based on Category 2 or 3 data alone are on the low end of the range of points available for those two categories. This effect on the point score is appropriate because the documented existence of exposure to a substance (Category 1) is a considerably more reliable measure of exposure than indicators based solely on the inferred possibility of exposure via media (Categories 2 and 3).