NRC Attributes, Potential Data Elements, Issues and Alternative			
Attribute and NRC Definition	**Potential Data Elements for this Attribute**	**Issue Identified for this Attribute**	**Alternatives**
Potency (page 95 NRC) *Indicates power or strength: amount of a contaminant required to cause an adverse health effect. How much (dose) of a contaminant does it take to cause illness (adverse effect)?* *Potency of a pathogen may refer to the # of organisms required to cause disease, while potency of a chemical refers to the dose required to cause disease.* *NRC recommended plotting as a percentile of the contaminant's potency relative to the potencies of all contaminants being considered. The percentile (0-100) scale should then be converted to a decile (1 through 10) scale.*	Non-carcinogens • RfD (chronic and/or acute) • TDI, ADI, MRL, etc • NOAEL (highest) • NOEL (highest) • LOAEL (lowest) and description of relevant endpoint • BMDL (lowest) • LD 50 oral (lowest) • RfC (if no oral) • LEL?? Carcinogens • Cancer slope factor --- • Quantitative SARs • Uncertainty factors	• Need to score non-carcinogens and carcinogens (NCAR and CAR) - for example data set - EPA used same definition as NAS and noted that the lower the critical dose/toxicity value the more toxic the contaminant. • Uncertainty factors would be applied to NOAEL, NOEL, LOAEL, BMDL, LD50 in order to make them comparable to risk-based value like the RfD. • Quality of the study should be considered • Need to determine how to select the program/programs for QSAR • Jamie noted that potency seems to run counter to scientific thinking - implies a level of exposure which is safe - while may apply to threshold chemicals - does not work for microbes and nonthreshold chemicals (genotypic carcinogens) - however, concept may be integrated with magnitude	For the example data set: • EPA established a hierarchy for data elements with risk-based values having precedence over raw data and EPA assessment having a precedence over those of other agencies • EPA found that the decile distribution of potency values was not effective because the 92 percent of a test group of 171 chemicals would have received a score of 10 based on their RfD. Instead EPA/Cadmus used a logarithmic scale to assign the scores for potency. The logarithmic distribution for the test set of chemicals was close to normal • When cancer and noncancer effects could each be scored, EPA selected the higher potency score to represent the chemical. NOTES: Inherent difference exist between using a "threshold" or a derived risk based value in comparison to a toxic effects concentration. Perhaps the threshold value, if available could be considered a toxic effect level for the purposes of scoring attributes. Units may be an issue, and if ADI's are compared with RfD's then some accounting for the derivation process may be warranted.

NRC Attributes, Potential Data Elements, Issues and Alternative			
Attribute and NRC Definition	**Potential Data Elements for this Attribute**	**Issue Identified for this Attribute**	**Alternatives**

Severity (page 94 NRC)
Degree to which a potential contaminant can cause an adverse health effect. (How bad is the effect?)
 	
Severity can be scored based on the anticipated clinical significance of the most sensitive health end point in affected individuals. NRC gave an example severity scale (page 96, Table 4.1)
	• Critical effect (associated with RfD or equivalent based on LOAEL)
	
	• Carcinogen classification
	  • EPA
	  • IARC
	  • NTP
		For example data set:
		• During scoring, tendency to consider effects at doses above the critical effect caused some bias
		• Noted that scoring requires expert judgment and depends on nature of effect (cancer, repr, dev, etc).
		• Using EPA's scoring scale, there were difficulties about how to score differences in organ and body weights.
		• If noncancer and cancer effect, used highest score to be conservative.
		• Need detailed description of critical effects.
		• Noted descriptions of critical effects data presented in IRIS were limited requiring assumptions or the desire to review IRIS files when scoring.
		• Need to develop an approach for chemicals that lack critical effect.
		• Need to expand descriptions and examples accompanying the scoring scale.
		• Note that raw data are qualitative/categorical, not numeric.
		• Inherent difference exists between severity of acute illness from pathogens and long term effects from chemicals. Temporal factors in severity will need to be discussed.
			• Definition used by EPA for the example data set - the degree of biological impact on survival and quality of life; Developed an EPA alternative severity scale to NRC's. Deleted chemical name (identity) to remove bias. (Should re-evaluate chemicals with their identity known to ensure scoring was appropriate).
			
			• Jamie suggests objective and transparent metric such as morbity/mortality and frequency of different outcomes.
			
			• Use log ($ willingness to pay to avoid the health endpoint).

NRC Attributes, Potential Data Elements, Issues and Alternative			
Attribute and NRC Definition	**Potential Data Elements for this Attribute**	**Issue Identified for this Attribute**	**Alternatives**
Prevalence (page 97 NRC) *How commonly does or would a contaminant occur in drinking water?"* *Two dimensions - temporal and spatial.* *Temporal is the fraction of time that a contaminant is found at a given locale (NRC, page 97)* *Spatial is the proportion at locales in which the contaminant is found (NRC, page 97*	• Indicators of size of water system or geographical area sampled if available • # or % of PWSs with detects • # or % sites with detects • # or % samples with detects • Water type (finished water, ambient, etc.) • Population served/% of population served with detects • Production/release/use data (quantity/year) • # Facilities using (??) • Release Medium • Temporal (spring, summer, fall winter)? • Location of sampling sites; regional or state • Location of discharge sites; regional or state	• Definition and meaning of prevalence needs to be discussed and agreed upon. • Spatial data limited; Temporal data even more limited; so may not be able to plot as NRC suggested. • Need to define temporal and spatial as far as amount of data necessary to make a temporal or spatial determination. • Type and amount of data for contaminants will vary. • Detection limit varies and changes over time so affects % prevalence • Need to define "level of concern" - appears that NRC meant for occurrence but Jamie said need to relate "level of concern" to health effects. Some folks at EPA may prefer not to - this approach may really be magnitude? • May/may not need population data • Raw data processing issues; data from various sources and data quality • Is geographical distribution of locales where the substance was found national or localized? • Inorganic contaminants have a much greater prevalence than most organic chemical contaminants, and may need a separate scale	Example data set was data rich and had no temporal data - had to only use spatial data. Most important elements may be: • # or % of sites with detections • # or % of samples at a site with detects. • Production release/use/data are secondary to finished or ambient water. • Use log (fraction of samples with positive result * median detection or reporting limit / RfD), replacing RfD by the corresponding cancer dose if a carcinogen.

NRC Attributes, Potential Data Elements, Issues and Alternative			
Attribute and NRC Definition	**Potential Data Elements for this Attribute**	**Issue Identified for this Attribute**	**Alternatives**
Magnitude (page 99 NRC) *The concentration or expected concentration (e.g., based on chemical production) of a contaminant relative to a level that causes a perceived health effect. In other words, is the level high enough to cause harm? Not absolute magnitude but magnitude relative to potency.* *NRC suggested using the median water concentration as the concentration parameter for the potency scoring.* *NRC suggested using the square root √(Potency score x 1-10 decile rank for median occurrence)*	• Median concentration detects • Mean concentration detects • Percentile concentration all data (25th, 50th, 75th, 90th, 95th, 99th) • Percentile concentration for all samples with detects (25th, 50th, 75th, 90th, 95th, 99th) • Median concentration all data • Maximum concentration all data • Mean concentration all data • Minimum concentration all data • Range of concentration • Amount released/year (to media) (TRI data) • Geographic distribution of releases • Potency score, if use NRC definition	• Need to define meaning of magnitude since some are unclear about what NRC is suggesting that we use because the Square root √(Potency score x 1-10 decile rank for occurrence) does not appear to be a measure of concentration relative to potency. It is a product of concentration and potency • Using potency data as part of magnitude score may create an interdependence of variables, which could affect model results. • Must redefine magnitude if potency is not to be used as an element in the scoring. • Minimum concentration is likely to be a "Not Detect" in almost all cases. • Should use the same measure for concentration in the scoring rather than a mixture of mean/median, 95% and minimum reporting levels.	• Example data set only used concentration to establish the magnitude scores - If use this approach what is the best method to determine and score magnitude? • Should magnitude score be derived from the distribution of the ratio of the median concentration to potency as the definition implies? • Use log (mean or median concentration among detects)

NRC Attributes, Potential Data Elements, Issues and Alternative			
Attribute and NRC Definition	**Potential Data Elements for this Attribute**	**Issue Identified for this Attribute**	**Alternatives**
Persistence/Mobility (page 100 NRC) *The likelihood that the contaminant would be found in the aquatic environment based solely on physical properties of the contaminant.* NRC suggested scoring solubility and stability into high (3), medium (2) and low (1) depending on specified length of time for persistence and specified solubility in H_2O. Then average the two scores and multiply by 10/3 to obtain score from 1-10.	• Stability (half-lives for hydrolysis, photolysis, biodegradation, aerobic soil, anaerobic soil) • Henry's Law constant • Kow, log Kow • Solubility product constant • Water solubility • Boiling point • log Koc • Melting point (?) • Vapor pressure • Transformation data - metabolites, degradation products of concern ??	• Some feel that pers/mob is really the lowest of the attributes (a default for lack of occurrence information). • Need a hierarchy for the selection and use of half life data. • When using temperature-related to physical property data, use the values for ambient temperatures. • Persistence/mobility could be important when using production/release data. Need to consider the stability of the contaminant in the environment and whether or not the contaminant moves from the site of release. • When persistence data indicate degradation of the contaminant, it is important to identify the degradates and insure that they are in the universe.	• Use log (solubility, ug/L) + stability score + log (quantity manufactured * fraction released to environment) or something along this order.