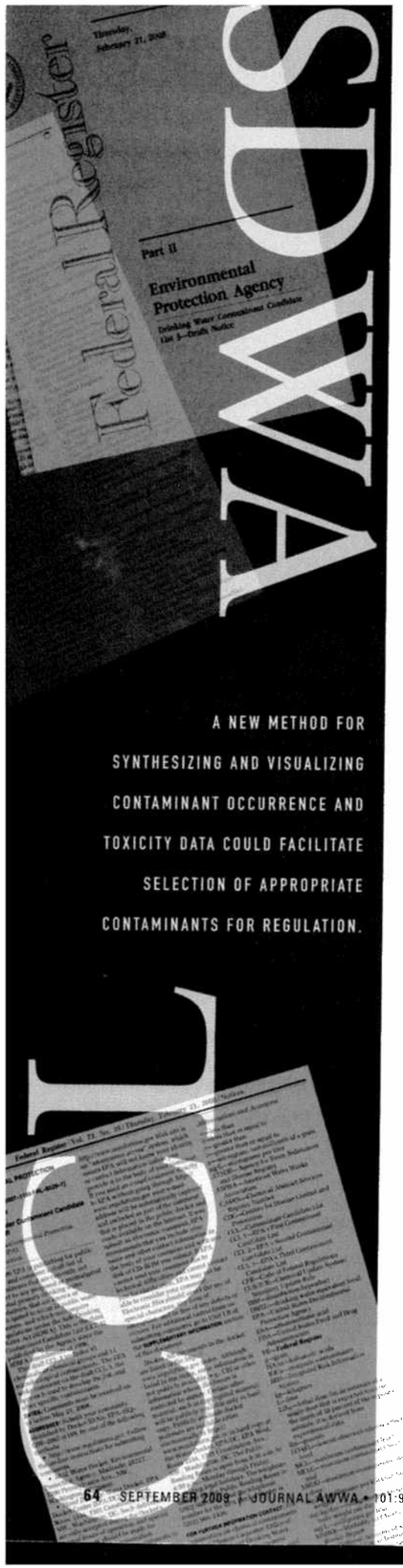


BY J. ALAN ROBERSON, MATTHEW LUEDERS, CRAIG ADAMS, AND JEFFREY S. ROSEN

Risk indexes for draft CCL3 chemicals

Under the 1996 Amendments to the Safe Drinking Water Act (SDWA, PL104-182), the US Environmental Protection Agency (USEPA) is required to publish a list of microbial and chemical contaminants that can be considered for potential regulation every five years (USEPA, 1996). This list of contaminants, known as the Contaminant Candidate List (CCL), is the foundation of USEPA's regulatory development process for national primary drinking water regulations. The evolution of contaminant selection processes since the initial 1974 SDWA (PL 93-253) has been summarized elsewhere (Rosen & Roberson, 2007).

Following development of the CCL, the second step in the regulatory development process is the regulatory determination by which USEPA decides whether to regulate a specific contaminant. Under the 1996 SDWA amendments, USEPA is required to make at least five regulatory determinations every five years. In its regulatory determination, USEPA can choose to regulate a specific contaminant, develop a health advisory, conduct more research on the contaminant, or not regulate the contaminant in drinking water.

The first CCL (CCL1) listed 60 contaminants (50 chemical and 10 microbial contaminants) that were selected primarily based on expert opinion (USEPA, 1998). In its first round of regulatory determinations from CCL1, USEPA decided not to regulate nine contaminants from CCL1 because a national drinking water regulation would not provide an opportunity for significant risk reduction as mandated by Section 1412 (b)(1)(A) of the 1996 SDWA amendments (USEPA, 2003). The second CCL (CCL2) carried forward the remaining 51 contaminants from CCL1 (USEPA, 2005). For its second round of regulatory determinations, USEPA decided to not regulate 11 contaminants from CCL2 (USEPA, 2008a) for the same reason the nine contaminants from CCL1 were not regulated.

During the CCL1/CCL2 time frame, USEPA solicited input on future CCLs from both the National Research Council (NRC) and the National Drinking Water Advisory Council (NDWAC). The NRC committee developed the conceptual approach of a "universe" of a large number of contaminants being narrowed down using simple screening criteria to a preliminary CCL (PCCL) and then further narrowed down using a more complex screening tool/algorithm to the CCL (NRC, 2001). The NRC and NDWAC input for improving the contaminant selection process has been previously summarized, and USEPA incorporated much of this input into the draft third CCL (CCL3; Rosen & Roberson, 2007). In 2006, following one of the NDWAC recommendations, USEPA solicited CCL3 nominations (USEPA, 2006). The draft CCL3 was published in the *Federal Register* Feb. 21, 2008, and contained 104 contaminants (93 chemical and 11 microbial; USEPA, 2008b). The CCL3 is anticipated to be finalized in mid to late 2009, and the third round of regulatory determinations is scheduled for 2013, five years after the second round.

The CCL process should be used as the starting point for the regulatory development process. The *Federal Register* notice for CCL1 contained a table listing each contaminant's research needs categorized into health effects, occurrence, analytical method, and treatment research (USEPA, 1998). CCL2 did not include such a list. When occurrence data are needed, USEPA can require utilities to collect such data through the Unregulated Contaminant Monitoring Rule (UCMR) and, if these contaminants are not found, make a subsequent determination that they should not be regulated. Five CCL1 contaminants were not detected at all in the first UCMR (UCMR1), which included monitoring by more than 3,300 public water systems (Roberson, 2005). On the basis of the lack of occurrence and no opportunity for meaningful health risk reduction, as required by SDWA Section 1412 (b)(1)(A) that addresses standards, USEPA appropriately decided to not regulate these five contaminants in its second round of regulatory determinations (USEPA, 2008a).

DATA VISUALIZATION FACILITATES UNDERSTANDING AND EVALUATION

A conceptual approach was used to aggregate and organize data and assess data value. The objective of the current project was to implement the previously summarized concept of a Cartesian coordinate system graphic presentation (space diagram) of health effects occurrence indexes for more than 300 chemical contaminants using data from several databases (Rosen & Roberson, 2007). This approach was selected to start the process of data evaluation to inform AWWA's comments on the draft CCL3 and to concisely aggregate the available occurrence and toxicity data for more detailed evaluation, given the anticipated 90-day comment period for the draft CCL3. The space diagrams were initially generated for 234 chemical contaminants to develop a better understanding of how to extract the appropriate information from a variety of databases and how to normalize the various numerical values obtained from them. Space diagrams for the 93 draft CCL3 chemical contaminants were subsequently generated during the public comment period (AWWA, 2008).

Data from numerous sources were ranked and rated. The initial data used for this project were gathered from five publicly available databases and from 38 published studies, with the intention of testing the extraction and graphing of data for compounds for which both occurrence and toxicity data were available. Following their release, the data used by USEPA for the draft CCL3 were included in a companion database, allowing the CCL3 data to be plotted along with the project data.

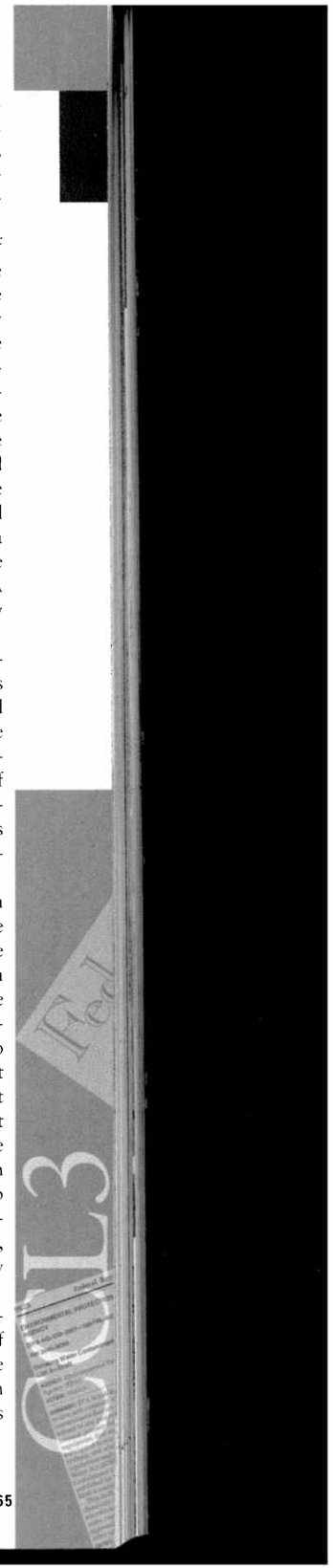
Normalization was necessary to translate the data into scores (1–10) on a common scale for plotting. USEPA's method of normalization through binning was used on both the data collected for the project and the USEPA data (USEPA, 2008c). Normalization bins for occurrence and toxicity are shown in Tables 1 and 2, respectively. The USEPA normalization method was extended to accommodate a few additional data types, including 70th percentile concentrations with median concentration, allowing the inclusion of more project data. The use of USEPA's normalization method provided a common framework within which the two data sets could be compared, and it eliminated the need to independently develop a defensible normalization protocol.

Table 3 shows an example of data normalizations for occurrence data. The project data for this one example compound comprised only two types of environmental release data (1) onsite surface water discharge and (2) total onsite and offsite disposal or other releases. These categories were also used with the USEPA data, with similar values and the same normalization results. The predominance of high normalized scores in both data sets suggested a high likelihood of occurrence. The larger amount of data in the USEPA data set may increase the certainty of this evaluation.

Table 4 shows an example of toxicity normalization. Again, there was some overlap among data types and data sources, but in this instance both data sets brought extra information into view. The similarity of normalized scores (4 or 5) demonstrated consistency among data types and sources, which increased confidence in the evaluation process.

Using USEPA's normalization method, an *x*/*y*-axes graph (space diagram) was generated with the occurrence score on the abscissa (*x*-axis) and the toxicity score on the ordinate (*y*-axis). Initially, two diagonal dashed lines were generated to aid in this visual analysis. The first diagonal line connected the highest toxicity score (10) with the highest occurrence score (10), whereas the second line connected the median scores (5/5). The logic was to develop three general compound classifications (based on the graphed areas), as previously conceptualized by Rosen and Roberson (2007).

- The high-risk category (indicating compounds to list) consisted of the area to the right and/or above the 10/10 diagonal line (a line drawn from the 10 on the occurrence axis to the 10 on the toxicity axis).



- The potential-risk category (possible compounds to be listed) took in the area between the 10/10 and 5/5 diagonal lines.
- The low-risk category (compounds not to be listed) consisted of the area to the left and/or below the 5/5 diagonal line.

The project team subsequently decided to "kink" the 10/10 line so that it ran diagonally from 10/10 to 5.5/5.5 and then horizontally from 5.5/5 to 10/5.5 (Figure 1). As a result of this kink, compounds that fell within a triangular area with frequent occurrence but relatively low toxicity moved from the "compounds to list" area into the "possible compounds to list" area. The use of this kinking procedure followed the underlying logic of the three SDWA risk-reduction criteria.

As part of the documentation for the contaminant selection process, USEPA generated a contaminant information sheet for each draft CCL3 contaminant. The sheets listed detailed occurrence and toxicity information for each compound, and from this information USEPA chose four particular data points, depending on the assortment of data types available for a specific compound (USEPA, 2008c; 2008d). USEPA then normalized these four chosen points to define the risk of the compound (USEPA, 2008c). Represented as risk indexes, all available data points from the contaminant information sheets were plotted on a space diagram (after normalization in order to provide a method for viewing all of the same data at one time, as opposed to having to go through each data point on a contaminant information sheet).

Uncertainty was estimated through combined ranking of the data types and data source. By assigning a ranking value to each data type and data source (according to their appropriateness for use for risk analysis), different combinations of available data could be easily sorted, and judgments could be made about assigning weights in decision-making. This allowed inclusion of less desirable data, an important factor given that one of the goals of this project was to use all available data to give the most robust image possible of all existing data. The ranking values assigned to the data type, termed Data Importance, ranged from 1 (most desirable) to 4 (least desirable). The ranking value assigned to the data source, termed Source Quality, ranged from 1 (most desirable) to 3 (least desirable). Specific characterizations for both Data Importance and Source Quality are shown in the sidebar on page 71.

For each value for each chemical, an individual data score was calculated using Eq 1:

$$\text{Data score} = 1 + [1/(\text{Data Importance} \times \text{Source Quality})] \quad (1)$$

which resulted in a range of final values between 2 (for the highest quality points) and 1.083 (for the lowest quality points). When the risk indexes were plotted on the space diagrams, the occurrence points were plotted against the toxicity points. The product of the two data scores was used to characterize data quality for the pair, with values between 2.75 and 4 assigned high-quality designators, those between 1.75 and 2.75 assigned medium-quality designators, and those between 1 and 1.75 assigned low-quality designators. This assignment pattern was chosen to show the appropriate variation in the spread of data quality in the space diagrams.

Within the risk indexes on the space diagram, the weighted mean value for the compound's occurrence and toxicity data was also displayed to show the center of gravity of the data cluster. This was necessary because points may lie beneath others with no evidence other than the location of the weighted mean value and the listing of the data values in the data list. The weighted mean values were calculated for both the occurrence data and the toxicity data by Eq 2:

$$\text{Weighted mean} = \frac{\Sigma(F_{H,i} \times H_i) + \Sigma(F_{M,j} \times M_j) + \Sigma(F_{L,k} \times L_k)}{\Sigma(F_{H,i}) + \Sigma(F_{M,j}) + \Sigma(D_{L,k})} \quad (2)$$

in which H_i is an individual normalized occurrence or toxicity value with a data score in the range of $1.6 < \text{data score} \leq 2$, M_j is an individual normalized occurrence or toxicity value with a data score in the range of $1.3 < \text{data score} \leq 1.6$, L_k is an individual normalized occurrence or toxicity value with a data score in the range of $1 < \text{data score} \leq 1.3$, $F_{H,i} = 1$ for each value H_i, $F_{M,j} = 0.5$ for each value M_j, and $F_{L,k} = 0.25$ for each value L_k.

Using the product of the occurrence and toxicity data score, symbology was developed to portray certainty for each individual compound data point on the space diagrams. Open circles were used to show data points developed from project data, and open triangles were used to show data points developed from USEPA data. Symbol size indicated certainty (data quality). Large circles and triangles were used to portray data points with data score products in the high-quality category, medium-sized circles or triangles were used to portray data points with data score products within the medium-quality category, and small circles or triangles were used to portray data points with data score products in the low-quality category. The data set's weighted mean values were plotted as a solid circle or triangle for the project data and USEPA data, respectively. An example plot is shown in Figure 1.

Individual compounds were assigned to groups on the basis of their chemical class. This enabled comparisons to be made among groups as well as among data sets within the same group. The weighted mean values for each compound were used to plot the single location of each compound on a group space diagram. The standard deviations of the normalized values for both the occurrence data and toxicity data were used to define the *x*-axis and *y*-axis error bars for each compound to facilitate visualization of the uncertainty present in the data because of consolidation of the data into a single point. As with the individual compound space diagrams, the circles represented project data and the

TABLE 1 Occurrence normalization bins

Normalized Value*	Bin Number 1	2	3	4	5	6	7	8	9	10
Surface water or finished water frequency detection—%	0	0.11	0.17	0.26	0.45	0.62	1.01	1.31	2.51	10
Surface water or finished water mean, median, or 70th-percentile concentration—*µg/L*	0	0.003	0.01	0.03	0.1	0.3	1	3	10	30
Surface water or finished water 90th-percentile, 95th-percentile,or maximum concentration—*µg/L*†	0	0.009	0.03	0.09	0.3	0.9	3	9	30	90
Environmental release data—*lb*	0	301	1,001	3,001	10,001	30,001	100,001	300,001	1,000,001	3,000,001

Source: USEPA, 2008c

USEPA—US Environmental Protection Agency

*Normalized value determined using USEPA's normalization method through binning.
†To accommodate the upper fraction of concentrations, the bins for the midrange were increased by a factor of three.

TABLE 2 Toxicity normalization bins

Normalized Value mg/kg/d*	Bin Number 1	2	3	4	5	6	7	8	9	10
Risk concentration based on slope factors	317	31.7	3.17	0.317	0.0317	0.00317	0.000317	0.0000317	0.00000317	0
LOAEL or NOAEL	1,700	3,170	317	31.7	3.17	0.317	0.0317	0.00317	0.000317	0
RfD	31.7	3.17	0.317	0.0317	0.00317	0.000317	0.0000317	0.00000317	3.17E-07	0

Source: USEPA, 2008c

LOAEL—lowest observed adverse effect level, NOAEL—no observed adverse effect level, RfD—reference dose, USEPA—US Environmental Protection Agency

*Normalized value determined using USEPA's normalization method through binning.

TABLE 3 Examples of normalized occurrence data for project and USEPA data sets

Parameter	Source	Value	Normalized*	*n*	Data Score†
Project data					
Onsite surface water discharges	TRI	322,487—*lb*	8.00	227	1.33
Total onsite and offsite disposal or other releases	TRI	2E+07—*lb*	10.00	318	1.33
USEPA data					
Frequency detection	DBP/ICR	55.5—%	10.00	2	2.00
Maximum detection value	DBP/ICR	30.6—*µg/L*	9.00	2	2.00
99th percentile detection value	DBP/ICR	29.7—*µg/L*	8.00	2	2.00
Surface water release	TRI	326,298—*lb/year*	8.00	57	1.33
Total release	TRI	3E+07—*lb/year*	10.00	57	1.33
Median detection value	DBP/ICR	7.6—*µg/L*	8.00	2	1.33

DBP/ICR—Disinfection Byproducts Information Collection Rule, *n*—number, TRI—Toxics Release Inventory, USEPA—US Environmental Protection Agency

*An assessment of risk for a given data value on a low (1) to high (10) scale used for direct comparison of various data types.
†An assessment of the quality of a data point based on the source and the applicability of the data type in question.

TABLE 4 Examples of normalized toxicity data for project and USEPA data sets

Parameter	Source	Value—*mg/kg/d*	Normalized*	*n*	Data Score†
Project data					
USEPA RfD	IRIS	0.2	4.00	268	2.00
Oral slope factor	California OEHHA	0.021	5.00	97	2.00
NOAEL	IRIS	15	5.00	99	1.50
LOAEL	IRIS	82	4.00	109	1.33
USEPA data					
RfD	USEPA IRIS (ITER)	0.2	4.00	33	2.00
RfD	USEPA HA	0.2	4.00	16	2.00
RfD	RAISHE	0.2	4.00	41	2.00
Minimum risk level	ATSDR (ITER)	0.2	4.00	17	2.00
Lowest oral chronic LOAEL	RTECS	12.5	5.00	50	1.33

ATSDR—Agency for Toxic Substances and Disease Registry, HA—health advisory, IRIS—Integrated Risk Information System, ITER—International Toxicity Estimates for Risk, LOAEL—lowest observed adverse effect level, *n*—number, NOAEL—no observed adverse effect level, OEHHA—Office of Environmental Health Hazard Assessment, RAISHE—Risk Assessment Information System Health Effects, RfD—reference dose, RTECS—Registry of Toxic Effects of Chemical Substances, USEPA—US Environmental Protection Agency

*An assessment of risk for a given data value on a low (1) to high (10) scale used for direct comparison of various data types.
†An assessment of the quality of a data point based on the source and the applicability of the data type in question.

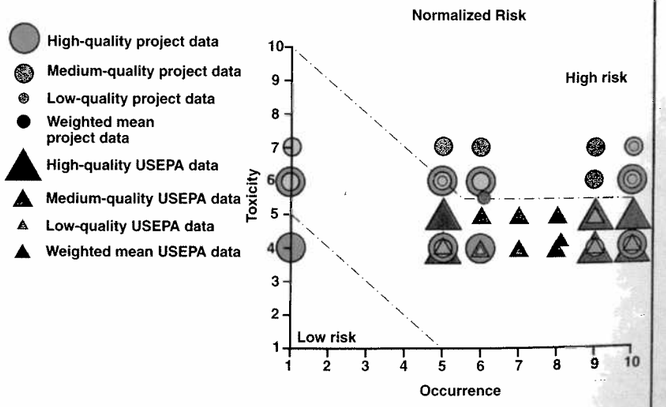
triangles represented USEPA data. Figure 2 shows an example of a group space diagram. The numeric designators next to the plotted points indicate different compounds.

For some compounds, a wide array of data may be available. For other compounds, few or no data are available. Because of this disparity in data availability, any given space diagram may be sparsely or heavily populated. Figures 3, 4, and 5 provide examples of individual space diagrams with small, moderate, and large amounts of data, respectively.

The space diagram in Figure 3 shows a wide spread of project data points (circles). Because all of the points represented medium-quality project data and the weighted mean value is shown as high at the center within the frame of the points, multiple points were overlain on top of one another. A limitation of the space diagrams is that it is not possible to see how many points are superimposed. The USEPA data, represented by triangles, were defined by fewer points and show the weighted mean value just left of center between the two outside points. This makes sense, given the location of the third point, which is farther to the left of the center. With the USEPA data characterizing the compound as low-to-moderate toxicity and the project data spanning a region well into the high-toxicity range, the specific data resulting in the points should be analyzed by an expert during any decision-making process.

Figure 4 shows a space diagram with a moderate amount of available data. More USEPA data were available than project data, as shown by the dominance of triangles on the plot. Several of the high-quality points were plotted to the right side, denoting very high occurrence, whereas several more high-quality points characterized the compound as having moderate-to-high occurrence scores. As with Figure 3, the specific data that resulted in these plot points should be evaluated by experts to resolve any discrepancies in the data. However, the tight region in which all of the data cluster is evidence that some certainty exists regarding the risk posed by this particular compound. Although the project data for the compound in Figure 4 were much more limited than the USEPA data, they fell within the same region as the USEPA data, and the weighted mean values of the data sets were consistent with one another.

FIGURE 1 Example plot with symbology legend



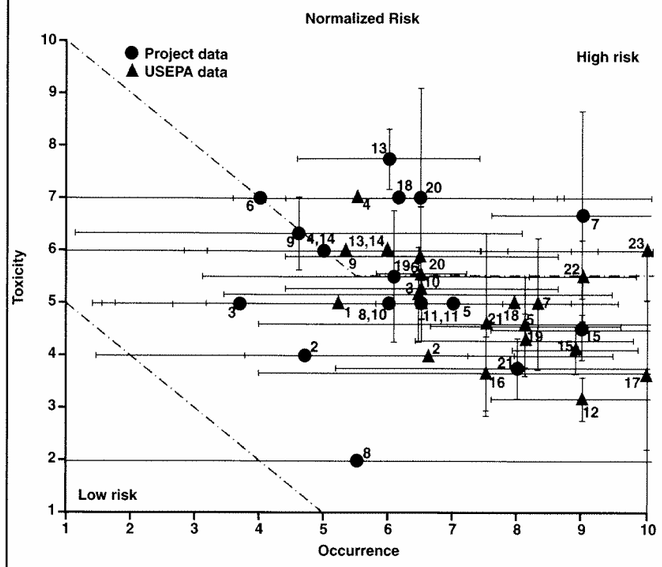
USEPA—US Environmental Protection Agency

Normalized risk determined using USEPA's normalization method through binning.

The space diagram in Figure 5 is well populated with a large amount of data. The risk indexes of both the USEPA and project data implied that a wide variety of data was available, and these data returned values over a range of qualities and values. The particular compound graphed in this figure was associated with a significant level of uncertainty. Project data showed high-quality data points at each extreme with regard to occurrence as well as an array of toxicity values, although the high-quality points did trend toward the compound being less toxic. The weighted mean value for the project data was located near the center of the risk indexes, and with the high-quality data spread across the broad populated space, provided evidence that uncertainty was high. The available USEPA data spanned smaller ranges of occurrence and toxicity values, but the high-quality data points were at each extreme of the risk indexes. This added to the uncertainty related to the USEPA data; again, the available data should be analyzed by experts in order to make meaningful decisions. In this space diagram, the weighted mean values for the USEPA and project data sets were somewhat inconsistent with one another, but the sources of each of the data sets may be compared to ascertain which one is more credible. The benefit of this approach is a concise aggregation and visualization of all of the available information so that experts can easily integrate and evaluate the amount of data, the location within the space diagrams, the spread, and the data quality.

Space diagrams are only as effective as the underlying data sets. Screening all of the sources to be included in a project is essential to ensure that comparisons among data types will be valid. Likewise, all available data sources should be included to gain the most accurate possible picture of the current data macrocosm. This approach ensures better understanding of all of the data and helps bridge the gap between using only the highest-quality data and using all available inputs.

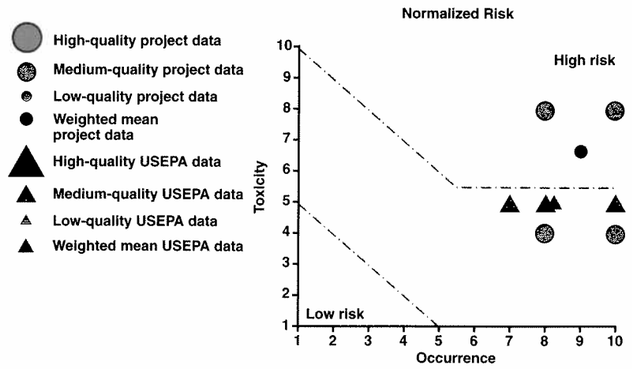
FIGURE 2 Example of a group space diagram with graphic designators



USEPA—US Environmental Protection Agency

The numeric designators next to the plotted points indicate different compounds. Multiple compounds with overlapping risk values are separated on the plot by a comma. USEPA numeric designators are shown in red to differentiate them from project numeric designators.

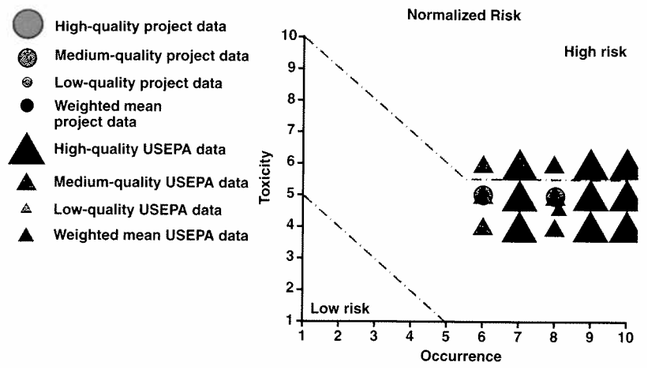
FIGURE 3 Example of a space diagram with a small amount of available data



USEPA—US Environmental Protection Agency

Normalized risk determined using USEPA's normalization method through binning.

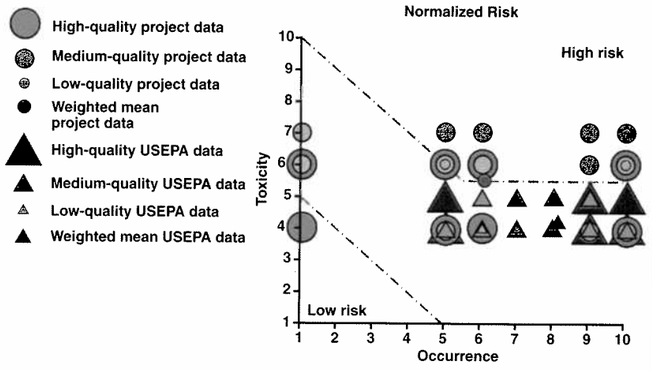
FIGURE 4 Example of a space diagram with a moderate amount of available data



USEPA—US Environmental Protection Agency

Normalized risk determined using USEPA's normalization method through binning.

FIGURE 5 Example of a space diagram with a large amount of available data



USEPA—US Environmental Protection Agency

Normalized risk determined using USEPA's normalization method through binning.

APPROACH HAS IMPLICATIONS FOR POLICY AND RECOMMENDATIONS

From a high-level policy perspective, the manner in which the CCL is generated and ultimately used profoundly affects the allocation and use of limited resources, both inside USEPA and within the drinking water community. In effect, even before any regulatory action on the list of contaminants is taken, the CCL has become a guiding set of priorities for research and for drinking water treatment plant design (given that treatment plants are designed to last 50–100 years). It is essential, therefore, that the conceptual framework surrounding data being collected to support CCL development and the organization and analyses of those data be consistent and well-defined so that the process meets its objective.

Clarification of CCL objective is needed. The primary objective of the CCL3 and the analyses underlying the overall CCL process is not completely clear in the *Federal Register* notice. It is not obvious from the draft CCL3 whether the primary purpose of the CCL is to provide a starting point for the regulatory development process or to serve as a mechanism to prioritize research needs. The 1996 SDWA amendments detail the CCL process in Section 1412(b)(1)(B)(i), Listing of Contaminants for Consideration, a title that implies that the CCL is the starting point for the regulatory development process. Section 1412(b)(1)(B)(i) addresses the schedule for the CCL and Section 1412(b)(1)(B)(ii) addresses regulatory determinations; the section preceding these sections discusses factors to consider in regulating a specific contaminant. Thus, through the SDWA statutory language, Congress gave specific guidance to USEPA for the listing and determination processes for contaminants that could potentially be regulated and indicated that the primary purpose of the CCL is to serve as a starting point for regulatory development.

The project described here led to several additional drinking water policy questions that informed the development of AWWA's comments on the draft CCL3 and ultimately should affect USEPA in its development of the final CCL3 as well as future CCLs.

- What is the appropriate number of contaminants for the CCL3?
- How can USEPA best conduct all of the research necessary for CCL3? Does the current process identify research gaps, and are these data gaps captured? Can the information collected in the process be used to identify CCL3 research priorities?
- How can the efficacy of the CCL process be determined?
- Does the CCL process work for all contaminants? For example, does the CCL process work for disinfection by-products (DBPs) as well as pharmaceuticals?

Some environmental advocates might argue that the CCL's purpose is to increase the number of regulated contaminants and not merely to decide not to regulate contaminants (though that is, in fact, what the USEPA opted to do in its first two rounds of regulatory determinations). However, neither the USEPA nor the drinking water community wants to jump back on the regulatory treadmill of the late 1980s and early 1990s when the agency was mandated by the 1986 SDWA amendments to regulate a specific list of 83 contaminants, plus 25 additional contaminants every three years. That mandate led to development of standards for 13 contaminants with zero violations within the first few years after the standards were put into place. Given that they had zero violations, these contaminants may be assumed to have minimal, if any, occurrence in source waters (Roberson, 2003). Furthermore, although these contaminants do not appear to occur to any significant extent, systems nevertheless must continue to monitor for them, probably in perpetuity, unless these standards are removed as part of USEPA's review of existing drinking water regulations (conducted every six years), which is not likely. For more than half of the contaminants on the list (42 of 83), systems had fewer than 10 violations over a multiyear time frame, leading to questions about whether these 42 contaminants should have been regulated at all. Developing new standards for the sole purpose of increasing the number of regulated contaminants is not a prudent policy decision because it does not lead to risk reduction. Resources should be directed toward development of regulations that provide meaningful health-risk reduction as mandated by the SDWA, taking into account both cost–benefit analysis and feasibility.

The appropriate number of contaminants for CCL3 is a complex question. Too small a CCL might indicate that USEPA's regulatory development process is too narrowly focused. Too large a list results in one of two potential outcomes.

First, USEPA could make determinations to not regulate contaminants that do not meet the three criteria in SDWA Section 1412(b)(1)(A) for establishing national standards: the contaminant may have an adverse health effect; the contaminant is known to occur or is likely to occur with a frequency and at levels of public health concern; and a national regulation will provide an opportunity for a meaningful health-risk reduction. As previously discussed, five CCL1 contaminants were not detected at all in the UCMR1, so USEPA appropriately decided to not regulate these five in its second round of regulatory determinations.

The second potential outcome of too large a list is that the USEPA, given its ongoing budgetary challenges, will never have sufficient funding to conduct all the research required to make the appropriate regulatory determinations from a large list. Tracking down the exact amount of the agency's drinking

Data Importance and Source Quality Metric Rankings

In this project, each data type and data source was assigned a ranking value so that different combinations of available data could be sorted and judgments made about assigning weights in decision-making. The ranking system allowed inclusion of data that were less desirable but still valuable to creating a robust picture of all existing data.

The type of data, termed Data Importance, was ranked from 1 (most desirable) to 4 (least desirable).

- A rank of 1 was assigned to the following occurrence data: finished water frequency detection, 95th-percentile concentration, and maximum concentration and to the following toxicity data: reference dose, adjusted reference dose, cancer oral slope factor, adjusted cancer oral slope factor, public health goals, and California notification level for drinking water.
- A rank of 2 was assigned to the following occurrence data: surface water frequency detection, 95th-percentile concentration, and maximum concentration and to the following toxicity data: no observed adverse effect level, no observed effect level, no significant risk level, maximum allowable dose level, and suggested no adverse response level.
- A rank of 3 was assigned to the following occurrence and application data: finished water median, 75th-percentile concentrations, and 90th-percentile concentrations and to the following toxicity data: lowest observed adverse effect level, lowest observed effect level, and lowest effect level.
- A rank of 4 was assigned to the following occurrence data: surface water median, 75th-percentile concentrations, and 90th-percentile concentrations, and to the following toxicity data: benchmark doses and LD_{50} values.

The source of the data, termed Source Quality, was ranked from 1 (most desirable) to 3 (least desirable).

- A rank of 1 was assigned to government sources and studies within the United States.
- The rank of 2 was not assigned.
- A rank of 3 was assigned to studies conducted outside the United States.

water research budget can be difficult because typically that number is buried at the back of budget documents. USEPA's drinking water research budget for fiscal year 2008 was $48.5 million (USEPA, 2008e).

In the following hypothetical example, it is assumed that each CCL3 contaminant can be placed in one of the following categories on the basis of the probable amount of its research needs: limited ($2 million for research), moderate ($5 million for research), and significant ($10 million for research). Assuming that the research needs costs are split equally across the 104 draft CCL3 contaminants—with one third of funding going to the limited-research category, one third to the moderate-research category, and one third to the significant-research category—total CCL3 research needs approach $600 million. With an assumption of the agency's current drinking water research budget at $50 million per year, with half (which may or may not be close to the actual percentage) of that budget dedicated to the CCL, the necessary research would take more than 20 years to complete, a time frame that does not support the five-year cycles for the CCLs and subsequent regulatory determinations.

Prioritization of CCL research is essential. No matter how many contaminants are listed in the CCL, the most important research must be conducted first. Research needs to be appropriately planned and executed to maximize its effectiveness. Information gaps for specific contaminants should be captured in a deliberate manner as contaminants proceed to the PCCL to the CCL. The graphing process outlined in this article provides a framework for facilitating data summaries through visual aggregation and for identifying and recording data and information gaps.

Development of a research plan is only a starting point. Executing the research in a timely manner for regulatory development can be challenging, and current research funding is stagnant. Through the Water Industry Technical Action Fund, AWWA conducted a retrospective analysis of the research plans developed by the drinking water community and by USEPA for arsenic regulation and for the microbial/DBP cluster. This analysis determined how much of the identified research was completed and used in the final regulations. Based on best professional judgment (which is somewhat subjective), the analysis found that about half of the high-priority research tasks were completed and incorporated or partially incorporated in the final regulations (Seidel et al, in press).

In light of the number of contaminants on the CCL1, CCL2, and CCL3, what would constitute an appropriate drinking water research budget? Under the assumptions and calculations presented previously, 24 CCL contaminants would require $120 million in research dollars over a five-year period, a figure that is close to the $125 million that would be available (assuming half of the drinking water research budget is dedicated to the CCL). Political pressure and priorities for drinking water research come from many different directions, and competing priorities will only increase in the future as other issues—e.g., climate change, infrastructure rehabilitation, and emerging contaminants such as pharmaceuticals and endocrine disrupting compounds—vie for funding. The Total Coliform Rule Distribution System Advisory Committee has also developed a long list of distribution research issues as part of its Agreement in Principle for the revised Total Coliform Rule (TCRDSAC, 2008). These kinds of policy questions, however, cannot be answered with simple math. Other policy considerations must be taken into account when making complex CCL decisions.

Assessing the effectiveness of the CCL is difficult. As previously discussed, the success of the CCL cannot be measured by the increase in the number of regulated contaminants. One measure could be the degree of public availability and the openness of the process of winnowing the CCL from the universe of chemicals. The 90-day public comment on the draft CCL3 provided minimal time to review the extensive documentation and background materials. USEPA could not provide open public access to the background documents for the draft CCL3 during its development because to do so would have essentially constituted earlier publication of the list. However, open access to the PCCL background documents would have allowed the general public to gain a better understanding of these documents and the underlying technical and policy issues.

Another measure of the effectiveness of the CCL (and of the SDWA regulatory program in general) could potentially be obtained through retrospective studies that attempt to link specific regulations to health outcomes. For example, a potential retrospective study could assess the relative risk reductions achieved by existing drinking water regulations, e.g., theoretical cases of gastrointestinal illnesses avoided, theoretical cancer cases avoided, and statistical lives saved versus actual outcomes. Another possible retrospective study could compare potential risk reductions attributable to existing regulations with the potential risk reductions associated with the 20 contaminants that USEPA chose not to regulate in its first two rounds of regulatory determinations. Beyond the CCL, the challenges in developing robust performance measures for the SDWA regulatory program have been summarized in other research (Raucher et al, 2006).

The current CCL process may not be suitable for addressing potential regulatory needs for DBPs. Exposure to DBPs has the potential to be significant because disinfection is required by all surface water systems (serving 191 million people) and by a significant portion of groundwater systems (serving more than 89 million people). The draft CCL3 included only four nitrosamines from among the wide range of known unregulated DBPs (one of the listed nitrosamines has not been identified as a DBP). The draft CCL3 even excluded two of the nitrosamines that are required to be monitored under the second UCMR (USEPA, 2007).

Analytical methods for many unregulated DBPs have been widely available at laboratories for some time. Occurrence data for haloacetonitriles, haloketones, chloropicrin, chloral hydrate, cyanogen chloride, and aldehydes have been previously summarized for the Information Collection Rule (McGuire et al, 2002). More recently, several new DBPs have been identified (Krasner et al, 2006), and the list of known unregulated DBPs is expected to grow with increased research focus and improved analytical methods.

Many classes of unregulated halogenated DBPs that commonly occur are not represented in the CCL3. These include haloacetaldehydes (di- and tri-), halopicrins (mono-, di-, and tri-), halofuranones (MX and analogues), brominated trihaloacetic acids, and haloacetonitriles (di- and tri-).

Even an informed reader struggles to follow the logic in the extensive documentation for inclusion and/or exclusion of these contaminants in the CCL chemical universe or PCCL. The *Federal Register* notice for the draft CCL3 stated that DBPs and treatment additives were automatically added to the universe on the basis of a default occurrence assumption and were moved to the PCCL despite limited availability of health or occurrence data. Although the CCL3 universe included a number of compounds listed previously, many others were skipped, and only one of these DBPs—chloropicrin—moved to the PCCL. Evaluating the CCL process with respect to the appropriateness of the listing outcomes for DBPs proved challenging because data gaps for contaminants that did not pass from the chemical universe to the PCCL were not documented. In addition, an opportunity for identifying research needs was lost.

In the development of the draft CCL3, a different conclusion was reached regarding pharmaceuticals. As summarized in the *Federal Register* notice, 287 pharmaceuticals were included in the universe, and approximately 10% of these moved forward to the PCCL on the basis of the screening process (USEPA, 2008b). Only one—nitroglycerin—ended up on the draft CCL3 because all of these compounds, with the exception of two—diazinon and phenyton—were scored on the basis of production data as opposed to release data or water concentration data, which would put them into the low-certainty bin.

Does the case of the DBPs described previously represent a deficiency in the CCL process? The CCL process does not appear to be able to accommodate the evaluation of unregulated DBPs. Would significant modifications to the existing process make sense, or is a different process, i.e., an alternate regulatory pathway, needed for this class of contaminants?

Resolving these drinking water policy questions is a challenge because the answers are not always evident and different stakeholders have varied perspectives. Currently, minimal investment is being made in the drinking water policy arena compared with other environmental issues such as air, fisheries, and climate change. As a result, opportunities for improved evaluation of the CCL process and the overall SDWA regulatory program have remained static. The project described here facilitated data summaries through visual aggregation, and these data summaries better inform a broader range of stakeholders about both the amount of data and the data quality for a large number of contaminants of potential concern.

Recommendations could help facilitate the process and provide direction for research funds. Development of a CCL research plan and prioritization and completion of the resultant projects are critical to the success of USEPA's regulatory development process. Completion of the priority research projects is essential to both selecting appropriate contaminants to regulate and establishing maximum contaminant levels according to SDWA mandates.

More stakeholder discussions are needed to identify the best means of providing ongoing public access to the databases and the background documentation being developed by USEPA for future CCLs. Broader and ongoing public access could serve several purposes.

- It would help the public be better informed and make better comments on future draft CCLs.
- It would allow for more varied policy analyses by a variety of institutions, such as schools of public policy at universities.
- It would improve sharing of ongoing research and/or gray literature such as master's theses, which may or may not be published and publicly available.

More work is also needed to determine how to best maintain the database developed for this project and how to push this database out to the broader drinking water research community. Although the database contains much useful information that was invaluable in developing comments on the draft CCL3, that information will soon become too outdated to be of use.

CONCLUSION

This project demonstrated a useful method for synthesizing and visualizing a large amount of data and provided a framework to organize the data for effective comparisons among contaminants. The risk-indexes framework facilitates the understanding of the extent of available data, the data quality, and the underlying uncertainties for occurrence and toxicity. The risk-indexes approach is not intended to replace the process that USEPA used for developing the draft CCL3; rather, it should be used as an evaluation tool to supplement that process and to better inform decisions made by USEPA on future CCLs.

The process that USEPA used for developing the draft CCL3 was a significant improvement over the expert-opinion approach that provided the foundation for CCL1 and CCL2. However, more work is needed to refine this process for future CCLs, particularly in the area of ongoing sharing of databases and background documentation.

The main purpose of the CCL needs to be more clearly defined, given that congressional intent for the CCL could be interpreted to be a starting point for the regulatory development process. Although research would therefore become a secondary priority, the amount of research funding might become the driver for determining the number of contaminants on future CCLs that can be adequately evaluated.

ACKNOWLEDGMENT

Funding for this project was provided through a partnership between the Water Research Foundation and the Water Industry Technical Action Fund (WITAF). WITAF is administered by AWWA and is funded through AWWA organizational members' dues. WITAF funds information collection and analysis and other activities in support of sound and effective legislation, regulation, and drinking water policies and programs.

ABOUT THE AUTHORS

J. Alan Roberson *(to whom correspondence should be addressed) is director of Security and Regulatory Affairs, AWWA, 1300 Eye St. NW, #701W, Washington, DC 20005; aroberson@awwa.org. He has a bachelor's degree in civil engineering from Georgia Institute of Technology in Atlanta and a master's degree in civil engineering from the Virginia Polytechnic Institute and State University in Blacksburg. A recipient of the Archie Becher Award, which recognizes exemplary AWWA staff, Roberson has 18 years of experience at AWWA and 12 years of experience as a consulting engineer. Matthew Lueders is a planning engineer with the Indiana American Water Company in Greenwood; at the time that this research was conducted, he was a graduate student at Missouri University of Science and Technology in Rolla. Craig Adams is chair and J.L. Constant Distinguished Professor, Department of Civil, Environmental, and Architectural Engineering at the University of Kansas; at the time that this research was conducted, he was a professor at Missouri University of Science and Technology. Jeffrey S. Rosen is vice-president of Clancy Environmental Consultants Inc., Scituate, Mass.*

Date of submission: 10/10/08
Date of acceptance: 05/19/09

If you have a comment about this article, please contact us at journal@awwa.org.

REFERENCES

AWWA, 2008. Regulatory Issues. Toxicity and Occurrence Charts. www.awwa.org/files/GovtPublicAffairs/PDF/CCLS3FinalDraftCompoundCharts.pdf (accessed Oct. 10, 2008).

Krasner, S.W.; Weinberg, H.S.; Richardson, S.D.; Pastor, S.J.; Chinn, R.; Sclimenti, M.J.; Onstad, G.D.; & Thruston, A.D., 2006. Occurrence of a New Generation of Disinfection By-products. *Envir. Sci. & Technol.*, 40:23:7175.

McGuire, M.J.; McLain, J.L.; & Obolensky, A., 2002. Information Collection Rule Data Analysis. AWWA, Denver.

NRC (National Research Council), 2001. *Classifying Drinking Water Contaminants for Regulatory Consideration.* National Academies Press, Washington.

Raucher, R.S.; Roberson, J.A.; & Harrod, M., 2006. Developing Performance Measures for the SDWA Regulatory Program—Melding What's Countable With What Counts. *Jour. AWWA,* 98:4:82.

Roberson, J.A., 2005. Triggers Behind the Regulatory Development Triad. *Jour. AWWA,* 97:3:66.

Roberson, J.A., 2003. Complexities of the New Drinking Water Regulations—Everything You Wanted to Know But Were Afraid to Ask. *Jour. AWWA,* 95:3:48.

Rosen, J.S. & Roberson, J.A., 2007. A Simplified Approach to Developing Future Contaminant Candidate Lists. *Jour. AWWA,* 99:3:66.

Seidel, C.S.; Roberson, J.A.; Summers, R.S.; Zearley, T.; & Tang, G., in press. Drinking Water Research Plans: Are Results Integrated Into Regulatory Policy? *Jour. AWWA,* in press.

TCRDSAC (Total Coliform Rule/Distribution System Federal Advisory Committee), 2008. Agreement in Principle. www.awwa.org/files/AIP.pdf (accessed Oct. 10, 2008)

USEPA (US Environmental Protection Agency), 2008a. Regulatory Determinations Regarding Contaminants on the Second Drinking Water Contaminant Candidate List; Notice. *Fed. Reg.,* 73:147:44251.

USEPA, 2008b. Drinking Water Contaminant Candidate List 3—Draft; Notice. *Fed. Reg.,* 73:35:9627.

USEPA, 2008c. Contaminant Candidate List 3 Chemicals: Classification of the PCCL to CCL. EPA 815-R-08-004. February 2008-Draft. www.regulations.gov.

USEPA, 2008d. Draft Contaminant Candidate List (CCL) 3 Contaminant Information Sheets. EPA-HQ-OW-2007-1189-0042. February 2008. www.regulations.gov.

USEPA, 2008e. FY 2008 EPA Budget in Brief. www.epa.gov/ocfo/budget/2008/2008bib.pdf (accessed Sept. 23, 2008).

USEPA, 2007. Unregulated Contaminant Monitoring Regulation (UCMR) for Public Water Systems Revisions; Final Rule. *Fed. Reg.,* 72:2:367.

USEPA, 2006. Requests for Nominations of Drinking Water Contaminants for the Contaminant Candidate List; Notice. *Fed. Reg.,* 71:199:60704.

USEPA, 2005. Drinking Water Contaminant Candidate List 2; Final Notice. *Fed. Reg.,* 70:36:9071.

USEPA, 2003. Announcement of Regulatory Determinations for Priority Contaminants on the Drinking Water Contaminant Candidate List; Notice. *Fed. Reg.,* 63:138:42898.

USEPA, 1998. Announcement of the Drinking Water Contaminant Candidate List. *Fed. Reg.,* 63:40:10273.

USEPA, 1996. 1996 Amendments to the Safe Drinking Water Act. Public Law 104-182. www.epa.gov/ogwdw000/sdwa/text.html (accessed Oct. 10, 2008).