National Drinking Water Advisory Council (NDWAC) Contaminant Candidate List (CCL) Classification Process Work Group

March 27-28, 2003
Washington, DC

Meeting Summary

- Final -

Table of Contents

Attachments

A. Work Group Members in Attendance
B. Agenda
C. Methods Group Update
D. Data Group Update
E. Introduction to Attributes
F. Suggested Principles for Attributes
G. "Microbes Are Different"
H. Drinking Water Microbiology: The State of Knowledge
I. "Building the CCL Universe"
J. "Draft Proposed Universe to PCCL Process"

Welcome and Introductions

The fourth meeting of the NDWAC CCL Classification Process Work Group was held on March 27-28, 2003. The meeting objectives were to

- learn more about microbials and the issues to be addressed in developing the CCL
- review the CCL Classification Process Work Group work plan in light of activity group progress and evaluation of tasks
- report and provide feedback on activity group activities to date:
    - identify questions and issues the work group needs to address
    - agree on tasks to be conducted to prepare for subsequent meetings
- identify additional technical expertise needed on the activity groups
- determine activity group tasks between March 27-28, 2003 and May 12-13, 2003

Facilitator Abby Arnold, RESOLVE, welcomed everyone to the meeting and asked the work group members and other meeting participants to introduce themselves (see attachment A). Following introductions, the work group reviewed and approved the meeting agenda (see attachment B).

Reports from Activity Groups on Progress Made between Plenary Meetings

Methods Activity Group

A member of the Methods Activity Group presented an update on the group's work since the February 5-6 plenary meeting (see attachment C). He reported that the group had drafted a rule-based gate approach to screen from the universe of chemicals to the preliminary CCL (PCCL), which would be reviewed by the activity group and presented to the plenary later during the meeting. The next step for the activity group is to continue to develop the proposed approach by identifying additional gates and deciding on rules or sorting criteria for the gates. The group will consider what criteria define "known" occurrence and "known" health effects and what the alternatives are for "potential" occurrence and "potential" health effects. The activity group member explained that four of the gates in the proposed approach represent the four intersections of demonstrated and potential occurrence and demonstrated and potential health effects in the Venn diagram used by the National Research Council (NRC) to present its conceptual approach to identifying contaminants for inclusion on the PCCL (figure 3-1 in *Classifying Drinking Water Contaminants for Regulatory Consideration*, NRC 2001). A fifth gate being considered is a nomination process, and other gates may also be added. The member noted that an agent in the universe can get to the PCCL by "passing through" any one of the gates.

The member reported that the Methods Activity Group also had continued evaluating possible approaches for classifying from the PCCL to the CCL. The group reviewed an example attribute scoring approach that was then used on an example data set to test various classification models. The activity group member commented that selecting and defining the attributes are important

questions for the work group. He presented several examples to illustrate possible approaches for scoring the attributes, noting that the activity group had not yet thoroughly discussed attribute issues. The group's next step is to review the results of the modeling tests run by EPA using both raw data and data scored using the example approach to better understand these models.

In closing the activity group member outlined the tasks and deliverables on which the group will work to prepare for the May plenary meeting.

Data Activity Group

A member of the Data Activity Group presented an update on the group's work since the February 5-6 plenary meeting (see attachment D). He reported that the group had drafted a proposed process for emerging contaminants and an approach to building the CCL universe, both of which would be reviewed by the activity group and presented to the plenary later at the meeting. The member reported that the group also had begun to review more than 200 data sources to determine their quality, reliability, and usefulness. Another activity group member noted that because of members' limited familiarity with many of the data sources, rather than evaluating all of the data sources, the group is considering developing principles for EPA to use to evaluate and select data sources, drawing on additional expertise outside the work group.

The activity group member outlined the group's tasks to prepare for the May plenary meeting, noting that the group would review the tasks and revise as needed based on progress thus far.

Definition and Review of Attributes

Tom Carpenter, EPA, presented an introduction to attributes (see attachment E). The purpose of the presentation was to begin the work group's discussion on attributes by reviewing the NRC definitions and use of attributes, reviewing the approach EPA used to score the example data set for the Methods Activity Group, and identify issues for further discussion. The NRC used the term "attributes" to refer to the characteristics of potential contaminants that contribute to the likelihood that a particular PCCL contaminant or related group of contaminants could occur in drinking water at levels and frequencies that pose a public health risk. Mr. Carpenter explained that NRC recommended using attributes to allow different types of contaminants (chemical and microbial) to be compared in a consistent manner and to allow different types and degrees of information to be associated with one attribute. The NRC recommended three attributes for occurrence: prevalence, magnitude, and persistence/mobility; and two attributes for health effects: potency and severity. Mr. Carpenter presented the NRC definition for each of the attributes and suggested possible steps for the work group to develop attributes:

- determine whether the NRC attributes are the most relevant
- determine if sufficient data are available to develop the value of an attribute
- define meanings/terminology
- determine practical scoring systems for selected attributes
- conduct scoring

He also suggested several principles for consideration in developing the attributes (see attachment F).

Mr. Carpenter then outlined the straw approach EPA used to score the data for 46 chemicals in the example data set being used by the Methods Activity Group to compare different classes of models. The purpose of outlining the approach was to highlight some of the issues encountered and lessons learned from the exercise. Mr. Carpenter focused on the approach for scoring severity and noted two issues that arose: how to account for sensitive subpopulations and how to choose among endpoints and exposure. Two of several lessons learned from the exercise were that detailed descriptions of critical health effects are needed and an approach should be developed for chemicals that lack a critical effect.

In closing, Mr. Carpenter summarized some of the issues involving attributes:

- Develop methods to evaluate and use different levels of data that characterize one attribute
- Identify surrogate data to supplement the lack of data (i.e., chemical production for occurrence)
- How to reconcile the difference in the number of contaminants types (i.e., 100,000 chemicals versus 1,000 microbes)
- Data relation between the Universe, the PCCL, and the CCL
- How should attribute scoring account for data availability, data quality, data gaps, data decisions, and relationships between the attributes

Discussion

Following the presentation and at other points during the meeting the work group began to discuss attribute issues. Members discussed whether there is redundancy among the five NRC attributes. One member commented that prevalence essentially is a measure of the percent of samples in which a contaminant is detected, which is driven by detection capability. He suggested that for prevalence to be meaningful, it would have to be a measure of the percent of samples in which a contaminant is detected at or above concentrations of health concern, which would be the same attribute as magnitude. Other members responded that prevalence as defined by NRC – a measure of how commonly a contaminant does or would occur in drinking water – provides useful information separate from magnitude. A member commented that if prevalence data are available they should be used, but a lack of prevalence data would not exclude a contaminant from consideration for the CCL. An EPA staff member commented that most of the contaminants on the PCCL will not have occurrence data because they will be unregulated contaminants. She added, though, that surrogate measures of occurrence possibly could be developed for some of the contaminants. A member noted that further thought was needed also to define prevalence and the other attributes for microbes.

In regard to severity a member questioned why separate consideration would be given to sensitive subpopulations when reference doses (RfDs) are calculated based on the breadth of sensitivity and therefore already include consideration of sensitive subpopulations. EPA staff explained that many reference doses were calculated in the 1980s, before EPA began applying uncertainty factors for database uncertainty. Another work group member commented that

consideration of sensitive subpopulations may be necessary if a measure other than reference dose is used for severity.

A member suggested that it might be helpful for the group to first define the attributes and decide what characteristics of contaminants are important to capture, and then discuss what data should be used to measure the attributes.

The work group decided to continue its discussion of attributes on a conference call. EPA will work with Nancy Kim to prepare additional materials explaining NRC's deliberations on and use of attributes.

Unique Features of Microbes and Available Data for the Work Group to Consider

Nelson Moyer, The Cadmus Group, and work group members Graciela Ramirez-Toro and Colin Stine presented information on some of the unique features and challenges of microbial contaminants (see attachment G). Dr. Moyer began by explaining that human and animal pathogens have several habitats. Water is not the only method of pathogen transmission, and only some pathogens have the ability to survive in the environment. Dr. Nelson outlined some of the limitations of our knowledge about pathogens: 1) it is historical, based on knowledge and experience, and it is difficult to anticipate or recognize emerging pathogens; 2) it is interactive, as the disease surveillance system depends on interactive cooperation among patient, physician, laboratory staff, and epidemiologist; and 3) it is methods dependent, requiring pathogen specific media and reagents. Information on microbes in the environment has come primarily from ecological research. Ambient pathogen monitoring is rarely performed, due to methods limitations and high costs. Because most monitoring that is done is in response to disease outbreaks, little is known about what happens to pathogens under "normal" circumstances.

Dr. Nelson reviewed the NRC attribute definitions and outlined some of the issues and challenges to scoring the attributes for microbes. He noted that for all of the attributes, data are available for few known pathogens. He commented that probably the most difficult attribute to determine for microbes is magnitude. He stressed also that the range of uncertainty for some microbiological data is orders of magnitude.

Dr. Ramirez-Toro presented some of the limitations and challenges to monitoring for pathogens in drinking water (see attachment H). Because of the difficulty of isolating pathogens in water, monitoring relies on fecal indicators as pathogen surrogates. The relationship between the presence of surrogate organisms (usually bacteria) and the presence of other pathogens, however, is not well established. So, while indicator monitoring provides a measure of water treatment efficacy and system integrity, it does not provide reliable information on what pathogens are in the water; outbreaks can and do occur even when surrogates are within acceptable ranges.

Dr. Ramirez-Toro noted another challenge: the conventional method, which requires first recognizing a disease and then associating that disease with an organism, cannot identify emerging pathogens before they cause disease. Research and surveillance are directed toward known pathogens.

Dr. Stine explained how the use of genomic and proteomic data to characterize virulence factor activity relationships (VFARs) may provide an alternative source of information on pathogen toxicity. Much is being learned as the sequencing of bacteria genomes progresses. Dr. Stine noted that sequenced genomes show a remarkable ability to add or lose genes, and new pathogens may arise by the addition of virulence genes or plasmids. He also explained that genes are related to each other and can be grouped into "families." Noting that 24 million of the 44 million entries in the Chemical Abstracts Services database are strings of genes and proteins, he suggested that the family membership of a gene is information that could be used in the CCL process.

At the end of his presentation Dr. Stine noted the "state of knowledge paradox": though great strides are being made in genomics, the increase in knowledge also serves to identify how much we do not yet know.

In response to questions Dr. Stine explained that no monitoring system (much less a practical monitoring system) has yet been developed to detect gene strings in environmental samples. He also explained that the number has not been determined for the percent homology that allows assumptions to be made about one microbe based on another.

Activity Group Breakout Sessions

Following their meetings in breakout sessions, the Methods and Data Activity Groups reported back to the plenary work group as follows.

Data Activity Group

Members of the Data Activity Group presented the draft document "Building the CCL Universe" (see attachment I) to the plenary work group and highlighted the key points of the proposed approach.

The approach proposes two "inclusionary principles" for building the CCL universe:
- The CCL universe should include those contaminants that have demonstrated or potential occurrence in drinking water
- The CCL universe should include those contaminants that have demonstrated or potential adverse health effects

The draft document includes the following proposed definitions:
- *contaminants:* any physical, chemical, or biological agent that does, or may, occur in water.
- *known contaminants:* physical, chemical, or biological agents that have been identified in the technical literature and adequately characterized to enable a judgment regarding their inclusion in the CCL universe

- *emerging contaminants:* a subset of known physical, chemical, or biological agents previously evaluated as not requiring inclusion in the CCL, for which new information becomes available which heightens concern and triggers re-evaluation.
- *new contaminants:* physical, chemical, or biological agents that are or may be newly-discovered or synthesized, for which little is known about their potential occurrence or adverse health effects

The process for creating and updating the CCL universe includes four components: 1) construction of the CCL universe for known contaminants, 2) a surveillance process for new and emerging contaminants, 3) a nomination and evaluation process for new and emerging contaminants, and 4) an expedited process to move emerging contaminants directly to a higher level of assessment within the CCL or to drinking water standard development.

Construction of the CCL universe for known contaminants would include the following steps:
- identify data sources that provide relevant information according to the inclusionary principles
- identify relevant information and combine into new database organized by contaminant

Suggested aspects of surveillance process for new and emerging contaminants include the following:
- institutionalize a proactive process that seeks relevant information, including surveying key contacts likely to have new data, such as public health organizations, health departments, and research institutes
- review key published data sources approximately every 2.5 years (mid-cycle of the CCL process)
- a means for identifying new information from recent updates of data sources to minimize redundant searching
- a review process that is technically sound and logistically practical
- a means for documenting the process and any decisions reached (for transparency)

Aspects of the proposed nomination and evaluation process for new and emerging contaminants included the following:
- identify and communicate with stakeholders
- design process with appropriate level of documentation
- evaluate submittals in transparent process
- take appropriate action for those found to have merit

The draft document acknowledges that EPA currently has the authority to accelerate review or regulation of contaminants through an expedited process. The document reiterates that it can be important to allow contaminants of immediate concern to be expedited.

Discussion

An EPA staff member expressed concern that budget constraints would not allow for a mid-cycle review for new and emerging contaminants. She asked that in developing a recommendation for

the process the group keep in mind EPA's ability to implement the recommended process. An activity group member responded that the proposed process for new and emerging contaminants was developed in response to a concern among work group members that building the CCL universe from existing databases would miss new and emerging contaminants. Another member commented that the group did not intend to elaborate any further on the proposed process, so if EPA chooses to move forward with the process, EPA could determine the details.

A work group member expressed concern that process proposed for constructing the CCL universe from existing data sources would result in too small a universe, particularly for addressing microbes. Other members pointed out that the data sources would include genebanks and other sources of microbial and genetic data, and that the proposed process for new and emerging contaminants would look beyond existing data sources. A member reminded the group of the need to balance between a universe that includes "everything" and a universe that EPA realistically has the capacity to build.

Through the discussion, the work group reached consensus on the "Building the CCL Universe" draft, contingent upon the agreement of the absent members. Members noted that further discussion will be needed as the details of the approach to building the CCL universe are developed. The next step for the Data Activity Group is to develop guidelines for EPA to use to select data sources to include in the CCL universe.

Methods Activity Group

Methods Group members presented a discussion paper outlining a draft proposed process for screening from the CCL universe to the PCCL (see attachment J, "Draft Proposed Universe to PCCL Process"). The process proposes several parallel paths or "gates" by which a contaminant could pass to the PCCL. The process focuses on two criteria: occurrence and health effects. To get onto the PCCL a contaminant must have demonstrated or potentially significant occurrence at a level to have demonstrated or potentially significant health effects. In discussing what would constitute "demonstrated" and what would constitute "potential" for occurrence and effects, the Methods Group distinguished between data and information. Data are actual measurements of occurrence in water and actual measurements of health effects; information is anything other than these actual measurements. For example, estimates of occurrence based on production data would be considered information. The first four proposed gates are as follows.

- **Gate I.** Have health effects data and concentration data. We know it is there and that it can cause problems. Is it there at high enough concentrations to be of concern? If yes, it is on the PCCL.
- **Gate II.** Have concentration data and health effects information. We know it is there, and suspect it may cause problems. Is it there at high enough concentrations to be of concern? If yes, continue onto the PCCL with medium confidence.
- **Gate III.** Have health effects data and concentration information. We suspect it may be there, and know it can cause problems. Is the potential plausible maximum concentration in

water of the contaminant greater than a demonstrated known or other safe dose? If yes, continue onto the PCCL with medium confidence.

- **Gate IV**. Have information on adverse health effects and possible occurrence, but have no data. It may be there, and it may cause problems. Is the potential maximum concentration in water of the contaminant greater than the potential safe dose? If yes, continue onto PCCL, but with low confidence.

The group is considering a fifth gate that would be a nomination process. Additional gates may also be defined.

Discussion

A work group member noted that given that research will not be done at the PCCL stage, the work group should further consider two crucial questions:

- What are the minimum data a contaminant should have to be on the CCL universe?
- What are the minimum criteria for screening from the universe to the PCCL?

A member raised a specific example for consideration: is it enough to have only a chemical's structure and to make estimates based on the structure for gate IV of the proposed process? If chemical structure is the only information available for a chemical will it even be included in the CCL universe?

A member noted that a distinction should be made between contaminants that are left off the PCCL because they are known not to occur or known not to have health effects and those for which there simply is no information to determine whether they occur or no information to determine whether they have health effects. He suggested that contaminants with known or potential occurrence but no information on effects and contaminant with known or potential effects but no information on occurrence could be candidates for the fifth gate nomination process.

Members discussed the idea of ranking or prioritizing contaminants on the PCCL. A member expressed concern with making prioritization judgments at the PCCL stage when additional factors and data will be considered in classifying contaminants from the PCCL to the CCL. Another member observed that there are two possible components of prioritization: 1) confidence in the data or information and 2) level of concern based on the information. A member of the Methods Group commented that he envisioned the high, medium, or low confidence rating for each gate as a tag to be attached to a contaminant that passed through the gate. The tag would indicate only how good the data were that placed the contaminant on the PCCL, and thereby might provide a useful piece of information in considering the contaminant for the CCL.

A member asked whether the intent was to work toward a common metric for toxicity and a common metric for occurrence. A member responded that it may not be possible to get to one metric for occurrence. Another member commented that though a common metric would be possible for toxicity it would involve high levels of uncertainty, which may not be acceptable.

A member commented that the gates may not be restrictive enough. He suggested that if the resulting PCCL is too large, more stringent screening criteria may be needed. Another member expressed concern with this idea of changing the process based on the size of the resulting list. He suggested that such a decision should be made by EPA rather than the work group. The first member responded that the work group may want to advise EPA on what to do if the PCCL is determined to be too large.

The work group agreed to the concept of the screening approach presented in "Draft Proposed Universe to PCCL Process." Deliberations on the issues and questions raised during this meeting will be continued in the activity groups.

Public Comment

No members of the public expressed an interest in making comments to the work group at this meeting.

Next Steps

Building the CCL Universe and Framework for Screening from Universe to PCCL
The work group agreed to the approach for building the CCL universe (attachment I) as well as a conceptual framework to move from the universe to the PCCL (attachment J). EPA will edit the framework and it will be distributed to the work group for review. Sara Litke or Abby Arnold will be in touch with members not present at the work group meeting to review these decisions.

Table of Contents for Final Report to NDWAC
RESOLVE will take a lead with the facilitation team to develop a draft table of contents, circulate it to Rick Becker, Graciela Ramirez-Toro, and Brian Ramaley for initial review and, after incorporating edits, send it to the full work group for review and discussion at the May meeting.

Glossary Review
EPA will review the glossary and terms used in "Building the CCL Universe" and "Draft Proposed Universe to PCCL Process" and advise the work group on how to make the terms used in these pieces consistent.

Conference Calls

Guidelines for Selecting Data Sources to Include in the CCL Universe
Conference Call: April 14, 11:00 – 1:00 pm Eastern

The next step for building the CCL universe is to develop guidelines EPA and others can use to determine what data and information should be included. EPA will draft a set of guidelines (based on Data Group discussions) for review on a conference call with all interested work group members. The *Data Activity Group* will lead this call.

Review NRC Definition of Attributes and Use of Attributes
Conference Call: April 25, tentative time is 11:00 – 1:00 p.m. Eastern

Nancy Kim will brief work group members on NRC attribute work and then the call participants will develop recommendations on what additional items the work group ought to consider for presentation at the May 12-13 meeting. Additionally, for this call RESOLVE will summarize concerns or issues raised by work group members about the NRC attribute definitions for review by work group members on the call. The *Methods Activity Group* will lead this call, but all interested work group members are encouraged to participate.

Data Activity Group

- Data Elements (Discuss report on results of applying guidelines to data sources for building the CCL universe (quality assurance/quality control))
  Conference Call: May 1, 12:30-2:30 p.m. Eastern.
- Review and Discuss Draft Presentation for May 12-13 Meeting
  Conference Call: May 6, 11:00 - 12:00 noon Eastern.

WG Members: Jamie Bartram, Rick Becker, Jeff Griffiths, Wendy Heiger-Bernays, Buck Henderson, Nancy Kim, Benson Kirkman, Gary Lynch, Ken Merry, Graciela Ramirez-Toro, and other resources as necessary

Methods Activity Group

- Modeling Analysis
  Conference Call: April 8, 1:00 – 3:00 p.m. Eastern
- Review and Comment on Universe to PCCL Gate Approach Working Draft
  Conference Call: April 30, 10;00 – 12:00 noon Eastern
- Open Agenda
  Conference Call: May 5, 1:00 – 3:00 p.m. Eastern

WG Members: Laura Anderko, Doug Crawford-Brown, Mike Dourson, Alan Elzerman, Brian Ramaley, Colin Stine, Craig Stow, Ed Thomas, Lynn Thorp, Dan Wartenberg, and other resources as necessary

VFAR Activity Group

Conference Calls: RESOLVE will assist with scheduling calls.
WG Members: Jeff Griffiths, Graciela Ramirez-Toro, Colin Stine

Future Meetings

The work group chose dates for meetings through 2003 as listed below. It is expected that all meetings will be held at the RESOLVE offices.

- May 12-13, 2003
- July 16-17, 2003
- September 17-18, 2003

- November 13-14, 2003