Environmental Science
Processes & Impacts

Summation of Disinfection By-product CHO Cell Relative Toxicity Indices: Sampling Bias, Uncertainty, and a Path Forward

Journal:	*Environmental Science: Processes & Impacts*
Manuscript ID	EM-ART-10-2019-000468.R1
Article Type:	Paper

Summation of Disinfection By-product CHO Cell Relative Toxicity Indices: Sampling Bias, Uncertainty, and a Path Forward

Elizabeth McKenna,[a] Kyle A. Thompson,[b] Lizbeth Taylor-Edmonds,[c] Daniel L. McCurry,[d] David Hanigan[a]*

Environmental Significance Statement:

Recent publications have divided concentration measurements by published cyto-and genotoxicity indices to produce a predicted toxicity metric. This methodology is valuable to determine the relative importance of measured DBPs in a sample. However, using published datasets we show here that statistical uncertainty and sampling bias inherent to predicted toxicity impact the conclusions of studies in which water quality is compared between samples or treatment processes. The conclusions here are important to future regulatory consideration, where predicted toxicity is being considered as a metric to compare treatment technologies which may result in action that is thought to be protective of public health but is detrimental.

Summation of Disinfection By-product CHO Cell Relative Toxicity Indices: Sampling Bias, Uncertainty, and a Path Forward

Elizabeth McKenna,[a] Kyle A. Thompson,[b] Lizbeth Taylor-Edmonds,[c] Daniel L. McCurry,[d] David Hanigan[a]*

[a] Department of Civil and Environmental Engineering, University of Nevada, Reno, NV 89557-0258

[b] Southern Nevada Water Authority, 1299 Burkholder Blvd., Henderson, NV, 89015

[c] Department of Mineral and Civil Engineering, University of Toronto, ON

[d] Astani Department of Civil and Environmental Engineering, University of Southern California, Los Angeles, CA, 90089

*Corresponding Author – David Hanigan, DHanigan@UNR.edu, 775-682-7517

Keywords: summed toxicity, calculated toxicity, predicted toxicity, dibromoacetonitrile, activated carbon, bioassay

Abstract

The cyto- and genotoxic potencies of disinfection by-products (DBPs) have been evaluated in published literature by measuring the response of exposed Chinese hamster ovary cells. In recent publications, DBP concentrations divided by their individual toxicity indices are summed to predict the relative toxicity of a water sample. We hypothesized that the omission or inclusion of certain DBPs over others is equivalent to statistical sampling bias and may result in biased conclusions. To test this hypothesis, we removed or added actual or simulated DBP measurements to that of published studies which evaluated granular activated carbon as a treatment to reduce the relative toxicity of the effluent. In several examples, it was possible to overturn the conclusions (i.e., activated carbon is detrimental or beneficial in reducing toxicity) by preferentially including specific DBPs. In one example, removing measured haloacetaldehydes caused the predicted cytotoxicity of a treated sample to decrease by up to 47%, reversing the initial conclusion that activated carbon increased the toxicity of the water. We also discuss measurements of statistical error, which are rarely included in publications related to predicted toxicity, but strongly influence the outcomes. Finally, we discuss future research needs in the light of these and other concerns.

Introduction

Disinfection by-products (DBPs) form from reactions of inorganic or organic matter with disinfectants during water treatment. The most abundant species by mass in drinking water are trihalomethanes (THMs) and haloacetic acids (HAAs), which are currently regulated by the United States Environmental Protection Agency (EPA).[1] THMs and HAAs are formed to a greater extent by free chlorine than chloramines.[2] Therefore, many treatment plants have switched from free chlorine to chloramination to reduce the formation of THMs and HAAs.[3,4] While lower concentrations of the regulated THMs and HAAs form during chloramination than chlorination, certain other DBPs form to a greater extent.[2,5] Therefore, there are tradeoffs in DBP formation from use of different disinfectants and researchers have focused recent efforts on determining which DBPs are the most important to mitigate formation of to limit the risk to human health.[6,7]

Some DBPs elicit cyto- and genotoxic responses and the "potency" (i.e., the LC_{50}, or concentration required to achieve an effect in 50% of the cells) of roughly 100 individual DBPs has been assessed by multiple *in vitro* and *in vivo* assays.[8-16] The most comprehensive data set uses Chinese hamster ovary (CHO) cells and the published potencies serve as a unique and valuable dataset for comparing the potency of DBPs and of classes of DBPs.[12] The published potencies have also been used to calculate "predicted toxicity" (i.e., the measured concentration of an individual DBP is divided by the published potency to calculate the relative toxicological contribution of each DBP, which are then summed). Predicted toxicity is part of an ever-evolving approach to understanding the human health impact of DBPs and has been used in studies to evaluate treatment process efficacy.[17] This approach is particularly attractive for labs without biological assay capabilities.

It was recently postulated that granular activated carbon (GAC) treatment may increase the toxicity of disinfected water, despite an overall removal of organic matter, based on the observation that GAC does not remove bromide, which may result in higher concentrations of

brominated DBPs.[18] Brominated DBPs are generally more potent than their chlorinated analogues based on results from the CHO comet assay.[19] As hypothesized, in rapid small-scale column tests, predicted toxicity increased due to an increase in brominated DBP formation, in particular, dibromoacetonitrile (DBAN). However, genotoxicity was also directly assayed with the SOS Chromotest and unlike predicted toxicity, the measured genotoxicity was consistently reduced with GAC treatment and tracked well with removal of bulk organic carbon. Of the 30 DBPs measured prior to and following GAC treatment, DBAN accounted for ~53% of the predicted toxicity and it was suggested that further GAC studies focus on HANs, particularly brominated HANs. The conclusion that HANs are the drivers of risk for disinfected water samples has only emerged in the past few years, but has been pervasive among predicted toxicity publications.[12,17,18,20-30]

Previously published studies focusing on predicted toxicity typically measured 30 to 40 DBPs, but a more recent study measured 70.[25,31] The team found that the overall mass of 70 DBPs decreased across GAC, but the number of brominated DBPs, including DBAN, increased. Because brominated DBPs are generally more potent than chlorinated DBPs as measured by the comet assay,[19] it was expected that the predicted toxicity would also increase, following other published studies, despite the overall reduced mass concentration of DBPs. Instead, the investigators found that the predicted toxicity decreased. The authors did not definitively reconcile the opposing conclusions of this research and other published literature, but we attribute the discrepancy to differences in number and speciation of measured DBPs.

Both the published literature and the more recent research discussed above conclude that DBAN precursors are poorly removed by GAC, thus DBAN contributed similar amounts of predicted toxicity before and after GAC.[17,18,23,25] However, by measuring a greater number of DBPs compared with prior studies and including precursors that are well removed by GAC, specifically dibromoacetamide and bromochloroacetamide, the more recent study effectively

diluted the weight of DBAN in the predicted toxicity calculation. This highlights how published literature may have unintentionally biased the toxicity calculations by including a comparatively potent DBP that preferentially forms in conditions that GAC selects for, while neglecting to measure DBPs that are effectively mitigated by GAC. Although inclusion of other DBPs reduced this bias, it is possible that other toxic DBPs which were not measured or remain unidentified could have altered the conclusion. Thus, we find the competing conclusions in the literature to be an excellent example of how predicted toxicity can be difficult to interpret.

We and others have suggested that the overall variability in conclusions across studies and assays is caused by the inherent uncertainty associated with this method of risk attribution. First, DBPs that are not measured or have not yet been discovered or assayed for toxicity might substantially contribute to the predicted toxicity, even at low concentrations, given that DBPs have toxic potencies that span greater than six orders of magnitude (i.e., sampling error or sampling bias).[12] Second, a typical suite of DBPs measured in advanced analytical publications (~30 to 70 DBPs) are representative of only ~30% of the overall DBPs as measured by adsorbable organic halides (AOX),[6,32] which still does not account for DBPs that do not contain halogen atoms. Third, measures of uncertainty are infrequently published, making comparisons difficult to interpret. Finally, published potencies are derived from individual DBP exposures, which ignore agonistic or antagonistic effects of mixtures.[33] Although these limitations are well known among experts in the field and discussed conceptually throughout perspective and review publications,[6,8] they are infrequently discussed in publications in which predicted toxicity is applied, potentially because they are only reviewed broadly, and there is no published demonstration of their potential impacts.

Although the impact of agonistic and antagonistic effects may be extremely important, for brevity, we limited the objective of this manuscript to demonstrating the impact of the number of DBPs measured and the statistical uncertainty on the reported toxicity in surface water

datasets, although our conclusions may be extended to other water sources. First, we removed groups of DBPs from published datasets to determine if the conclusions regarding the efficacy of GAC changed dependent on the number of DBPs measured. The removal of groups of DBPs was not focused on a specific subset of DBPs; we evaluated the theoretical removal of all groups of DBPs individually. Second, we aggregated published haloacetamide (HAcAm) data and inserted the aggregates into datasets from publications that assessed GAC treatment but did not measure HAcAms (i.e., we simulated the measurement of additional DBPs) and compared the conclusions from the publications to hypothetical datasets. We chose to supplement the datasets with HAcAms because they are relatively potent, measured frequently enough for there to be data available, and because HANs can be formed by hydrolysis of HANs.[34] Finally, we discuss measurements of summative error, which are absent in many publications, and comment on the potential impacts of discounting rigorous statistical analysis. Because many DBPs are not genotoxic, published predicted toxicity literature tends to focus on predicted cytotoxicity rather than genotoxicity. We also focus on cytotoxicity because the greater dataset available, but discuss genotoxicity where possible.

Methods

CHO cell DBP potencies were obtained from two publications[9,12] and a personal correspondence.[35] DBP potencies are determined by exposing CHO cells to multiple concentrations of an individual DBP and measuring either cell death (cytotoxicity) or DNA damage (genotoxicity). Predicted toxicity was calculated by dividing measured concentrations of DBPs by their respective geno- or cytotoxic potency (LC_{50} [cytotoxicity], or 50% tail DNA or midpoint of DNA tail moment [genotoxicity]), resulting in a unitless toxicity (see Table S1 for toxic potencies). DBP concentrations from pre- and post-GAC treatment were from multiple publications (see Table S2 for background on the treatments) .[18,25,36] HAcAm concentrations were derived from two publications that measured HAcAms before and after GAC treatment at a

total of 18 drinking water treatment plants (Table 1, pre-GAC concentrations in Table S3).[25,37] The GAC influent water samples were either not oxidized, or pre-oxidized with varying oxidation techniques (chlorine, chloramine, ozone, $NaMnO_4$, $KMNO_4$, see Table S2), representing a broad array of pre-oxidation conditions. HAcAms measured in the GAC effluent samples in both the data that was aggregated from and supplemented to were primarily chlorinated, except two samples, which were chloraminated (Pilot Plant 2015 and 2016 in Figures 1, 2, and 4).[25]

Table 1. Post-GAC HAcAm concentrations derived from two publications. Data from Kosaka *et al.*[37] is the average from 6 treatment plants and Stanford *et al.,*[25] from 12 treatment plants.

	Mean concentration from Kosaka *et al.*[37] (nM)	Mean from Stanford *et al.*[25] (nM)	Mean of both datasets (nM)
DCAM	1.69 ± 0.54	22.47 ± 30	12.08 ± 10.3
DBAM	2.61 ± 1.7	7.84 ± 3.3	5.23 ± 2.6
BCAM	2.13 ± 0.55	8.34 ± 5.2	5.23 ± 3.1
TCAM	0.62	3.08 ± 0.9	1.85 ± 1.2
CAM	1.43 ± 0.50	Not measured	1.43 ± 0.50
BAM	1.57 ± 1.06	Not measured	1.57 ± 1.06

In cases where a HAcAm was not detected, a concentration equal to half the provided MDL was assumed. DCAM, DBAM, BCAM, and TCAM were measured post-GAC by Stanford *et al.,*[25] and therefore the toxic potencies provided in the third column (mean of both datasets) are averages of both data sets. In one instance, Stanford *et al.,*[25] four HAcAms were measured in the additional dataset and therefore the original data from the publication was used, with only two HAcAms supplemented from Table 1. In the study by McKie *et al*.[36] DBAN was not measured, thus, in addition to supplementing the HAcAm data from Kosaka *et al*.,[37] the average of sixteen samples after GAC treatment from Krasner *et al*.[18] and Stanford *et al.*[25] were included (5.12 ± 3.39 nM DBAN, Table S4). DCAM is not genotoxic and thus was not included in genotoxicity. We are unaware of additional sources of HAcAm occurrence data in drinking water facilities with GAC treatment.

Results and Discussion

Simulating the omission of specific DBP subsets in published data

We removed groups of DBPs from published data sets to demonstrate that omission of specific analytes can alter the conclusion of the analysis. We discuss in detail only one example here, but additional data aggregated from publications are provided in the SI, and similar conclusions follow (Figure S1, Panels 2 and 4, and Figure S2, Panels 1 and 2.3). In Figure 1, we show the contribution of individual DBPs to predicted toxicity from the initial data set. In Panel 4, we show that predicted cytotoxicity decreased 20% across GAC (22,000 bed volumes [BV]) when HAcAms were included in the initial measurements. However, removal of HAcAms (pink compounds) from the data (i.e., simulating measurement of fewer compounds) results instead in a 6% increase in predicted cytotoxicity after GAC treatment. Similarly, in Panel 3, GAC treatment reduced the predicted cytotoxicity by 13% to 4%. However, had HALs (purple compounds) been omitted from the analysis, the initial untreated sample would have been predicted to be 5% to 14% more cytotoxic than the GAC treated samples, at the two BV sampled (shown with HALs as the top bars in Figure S3A for clarity). Finally, in Panel 6, GAC treatment increased the predicted cytotoxicity by 15% to 19%, mostly due to increased formation of chloroacetaldehyde. Omission of HALs, including chloroacetaldehyde, caused the predicted toxicity to decrease across GAC by 47% to 28% (also shown in Figure S3B with CAL as the top stacked bar for clarity).

Removal of other DBPs in these three panels or in Panels 1, 2, 5, and 7 resulted in changes to the magnitude of the predicted toxicity change, but generally no change to the initial conclusion, that GAC reduced the toxicity profile of the samples. Thus, omission of specific classes of DBPs does not always change the interpretation of the data and the magnitude of the changes presented here are a relatively small percent contribution to the toxicity profile of the samples. Given that observed reductions or increases in predicted toxicity across the GAC are

relatively small in most cases, we caution that without rigorous statistical analysis, conclusions as to the benefit or cost of a treatment process are not appropriate. Additionally, in some cases, the conclusion that a technology results in better or worse water quality is dependent on which DPBs were measured, which is troublesome considering the investment required to implement such technology in water treatment systems.

The contribution from THMs' predicted toxicity are relatively small compared to those of other DBPs. Total THMs contributed 0.2% to 2.4% of predicted cytotoxicity to each water sample without the addition of the simulated HAcAms (Figure 1). THMs do not elicit a genotoxic response, and therefore did not impact predicted genotoxicity. HAAs contributed 0.2% to 23% of predicted cytotoxicity, which was generally less than other classes of measured DBPs. HAAs dominated genotoxicity in some samples, but not in others (Figure S4). The US EPA currently regulates THMs and HAAs, but these species did not contribute appreciably to predicted toxicity in the cases here or in other publications.[26,38] We believe this is an especially useful application of predicted toxicity; to compare the relative importance of individual compounds or classes of DBPs in a single sample, but not between samples or treatment groups. Finally, to interpret such data as an indication that a certain class of DBPs should be subject to regulation instead of or in addition to THMs and HAAs is likely an overextension of the data (i.e., THMs are probably not important in the given data, but it is not known whether DBAN is important, only that it is more important than THMs [see *Importance of DBAN*]).

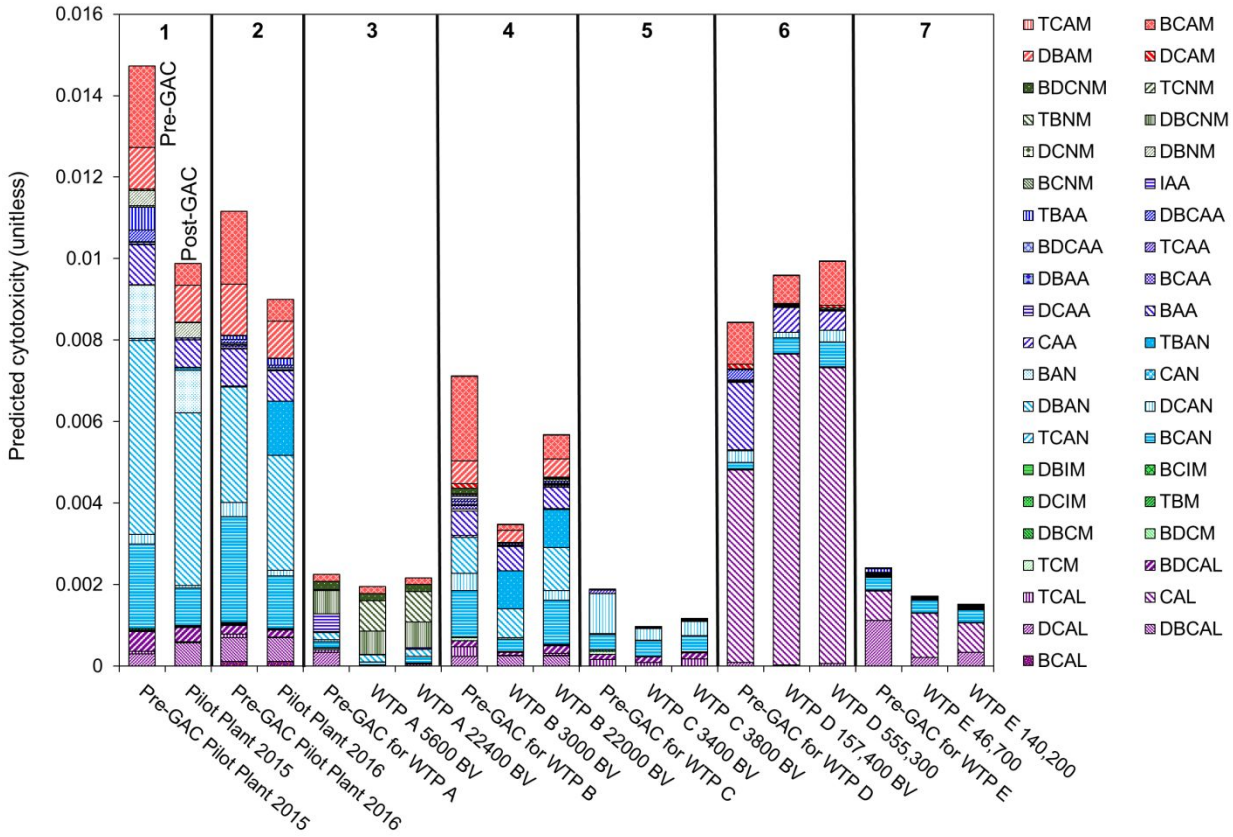

Figure 1 Components of predicted cytotoxicity for data from Stanford *et al.*[25] and Cuthbertson *et al.*[31] (same data in both publications). Pink colored compounds are HAcAms. Left-most bar in each panel is pre-GAC predicted cytotoxicity, other bars are GAC effluent samples. Only 41 DBPs are shown, rather than the 70 that were measured, because 29 DBPs were not detected. Compound abbreviations are provided in Table S1 and raw data provided in Table S5. Panels 3, 4, and 6 are instances where omission or inclusion of specific DBPs or groups of DBPs may cause an inversion of the conclusion that GAC treatment was beneficial or detrimental.

Incorporation of unmeasured DBPs

We initially supplemented aggregated HAcAm data from 18 WTPs (Table 1) into the same pre- and post-GAC example dataset because the number of DBPs measured is relatively comprehensive. We chose to supplement the datasets with HAcAms because they are relatively potent, measured frequently enough for there to be data available, and because HAcAms can be formed by hydrolysis of HANs.[34]

HAcAms were measured in some of the treatment plants and we supplemented the data for other plants or added specific HAcAm compounds to those that did not measure all six HAcAms. The supplemented HAcAm data contributed an average of 51% ± 31% of the predicted toxicity for pre-GAC data and an average of 38% ± 23% for post-GAC data (Figure 2). Predicted cytotoxicity decreased across GAC for five of the seven cases, and the addition of HAcAm data (pink bars) did not change this conclusion. However, in Panel 4, the initial dataset without HAcAms indicates that the predicted toxicity of the GAC effluent initially decreased across GAC (3,000 BV), but then increased to greater than the pre-GAC sample (22,000 BV), suggesting that GAC caused the total predicted toxicity of the treated sample to be greater than the untreated sample. Much of this can be attributed to the increase in tribromoacetonitrile (TBAN) formation. However, with the simulated measurement of HAcAms (i.e., addition of aggregated data), the predicted toxicity of the GAC treated samples tends to increase with increasing GAC use, but does not exceed the predicted toxicity of the pre-GAC sample, suggesting that GAC decreased the predicted toxicity of the water relative to the untreated sample. This is attributable to a decreased weighting of TBAN due to a greater number of compounds measured.

In Panel 2, predicted toxicity decreased relative to the untreated sample despite a large increase in TBAN and independent of the addition of HAcAms. However, had an additional sample been taken at a later point in time, predicted toxicity may have increased because of the large increase in TBAN across GAC and decreasing DBP precursor removal across GAC over time. Amending aggregated HAcAm data would reduce the impact of TBAN and potentially result in decreased predicted toxicity. In Panel 6, predicted cytotoxicity increased independent of the inclusion of HAcAms, but does so to a lesser extent when HAcAms are amended. Again, the relative changes observed here are small and only in select instances, but the impacts on decision making are substantial if the results are assumed to be statistically significant.

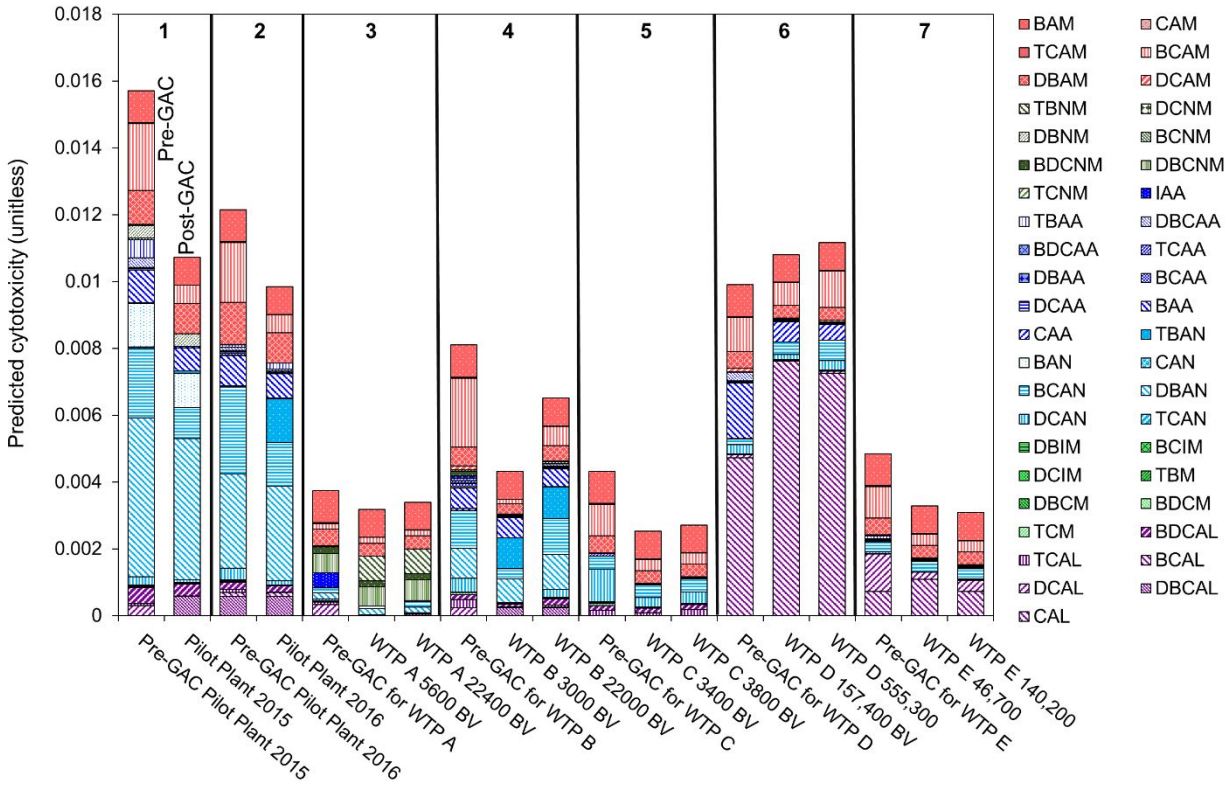

Figure 2 Components of predicted cytotoxicity for data from Stanford *et al.*[25] and Cuthbertson *et al.*[31] (same data in both publications). Pink colored compounds are either measured HAcAms from the study or supplemented HAcAms derived from the mean concentrations at 18 WTPs (Table 1).[25,37] Left-most bar in each panel is pre-GAC predicted cytotoxicity, other bars are GAC effluent samples. Only 41 DBPs are shown, rather than the 70 that were measured, because 29 DBPs were not detected. Compound abbreviations are provided in Table S1 and raw data provided in Table S5. All panels are supplemented with CAM and BAM data from Table 1. Additionally, Panels 1 and 2 are supplemented with TCAM data, Panel 3 is supplemented with DBAM and TCAM data, Panel 5 is supplemented with DBAM and BCAM data, Panel 6 is supplemented with DBAM data, and Panel 7 is supplemented with DBAM, BCAM, and TCAM data. Panels 2 and 4 are instances where inclusion of supplemented HAcAms may have significantly impacted conclusions. Conclusions from other panels are impacted to a lesser extent.

In another published dataset in which a relatively small number of DBPs was measured (N=15), predicted genotoxicity increased slightly across biologically active GAC (e.g., biofiltration) partially due to increased CAA formation (Figure 3 Panels 4 through 8). However, including simulated HAcAm data caused predicted toxicity to decrease by 52% to 75% across

the biofilters. The publication also measured absorbable organic halogens (AOX) and SOS genotoxicity via the SOS Chromotest and found strong correlations between SOS genotoxic response and AOX, THMs, and HAAs, and particularly strong correlation between THMs and SOS genotoxic response after biofiltration (R^2 = 0.97). It is well recognized that THMs and HAAs are not likely to be the primary toxicological drivers based on their potency and occurrence, but they may be well correlated for specific assays. One additional published data set is provided in the SI and simulated addition of HAcAms follows the conclusions here but is not discussed in depth for brevity (Figure S5, Panels 1 and 2.3).

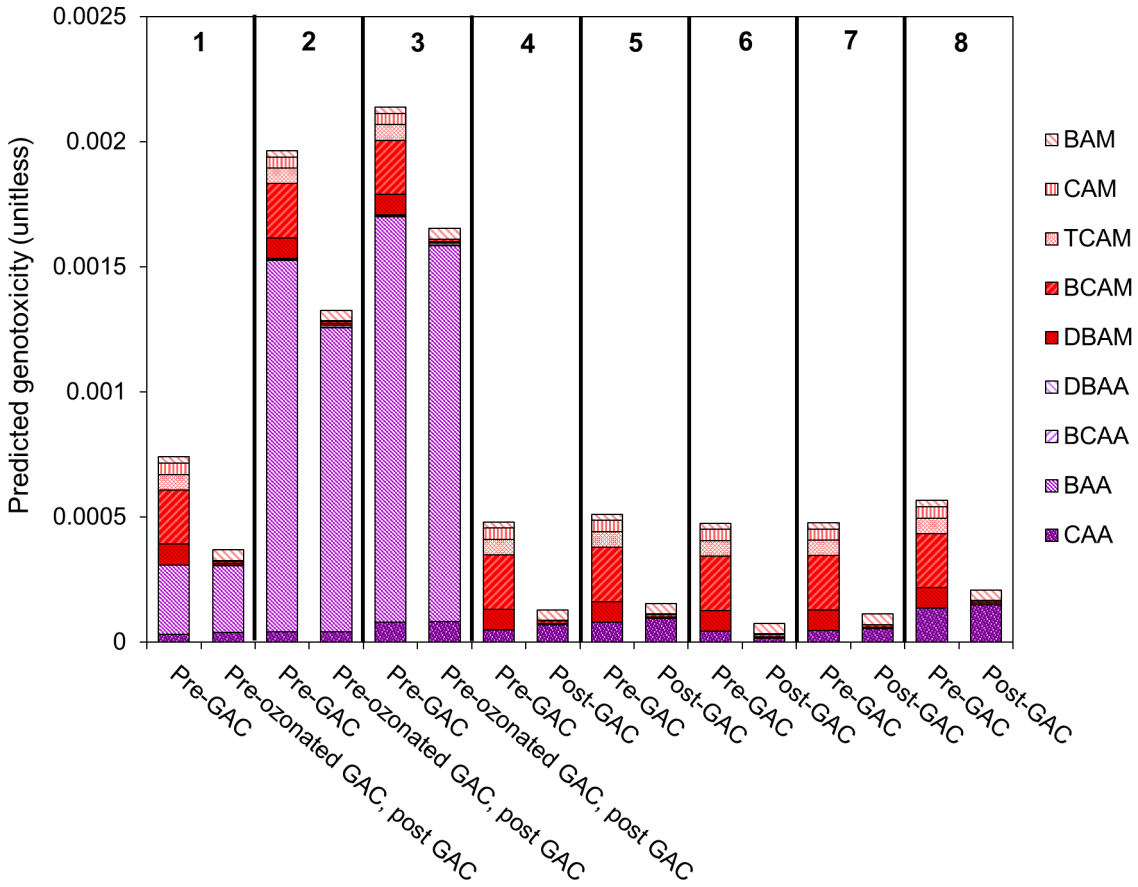

Figure 3 DBP components of predicted genotoxicity from McKie *et al.*[36] Red colored compounds are supplemented HAcAms (see Table S3). Left-most bar in each panel is pre-GAC, other bars are post-GAC. Panels 1-3 represent samples from Lake Ontario taken in three different months. Panels 4-8 are samples from the Otonabee River taken from five separate sampling events. Only 9 DBPs are shown, rather than the 15 that were measured, because 6 DBPs were not

detected or do not have genotoxic values. Compound abbreviations are provided in Table S1 and raw data provided in Table S6. Panels 4 through 8 are strongly influenced by the inclusion of haloacetamides while panes 1 through 3 are driven by BAA.

Importance of DBAN

Because of its extraordinarily high toxicity index, detecting DBAN at the detection limit, typically between 0.2 μg/L[27] and 1 μg/L,[18] results in a contribution of 3.5×10^{-4} to 1.8×10^{-3} to predicted cytotoxicity, the same order of magnitude as the total predicted toxicity for most drinking water samples. Because of this, a large number of publications have implicated DBAN as the primary driver of toxicity,[17,18,20-29] and therefore we also examined the importance of DBAN before and after the addition of aggregated HAcAm data, which also have relatively high toxicity indices, but are not always measured alongside HANs. In Figure 4, we show the contribution of DBAN to the overall predicted cytotoxicity in sampling events from three publications with varying treatment processes and source waters. The addition of HAcAms to the post-GAC samples caused a 10% to 63% percent decrease in the contribution of DBAN to predicted cytotoxicity for two studies.[18,36] DBAN associated toxicity in the third study decreased to a lesser extent because four of the six HAcAms were measured in the initial study, which diluted the effect of adding additional HAcAms. However, inclusion of two additional HAcAms (i.e., CAM and BAM) reduced the percent contribution of DBAN to predicted cytotoxicity by an additional 2% to 4%. In Figure S6, we show the percent contribution from DBAN to predicted genotoxicity, which generally agrees with the conclusions presented for cytotoxicity. Although this exercise might seem intuitive, we note here that increasing the number of total compounds measured will diminish the relative contribution of DBAN to predicted toxicity. Therefore, the conclusion that DBAN drives overall toxicity may be an artifact of 1) the number of DBPs measured, and 2) the relative toxic potency of DBAN.

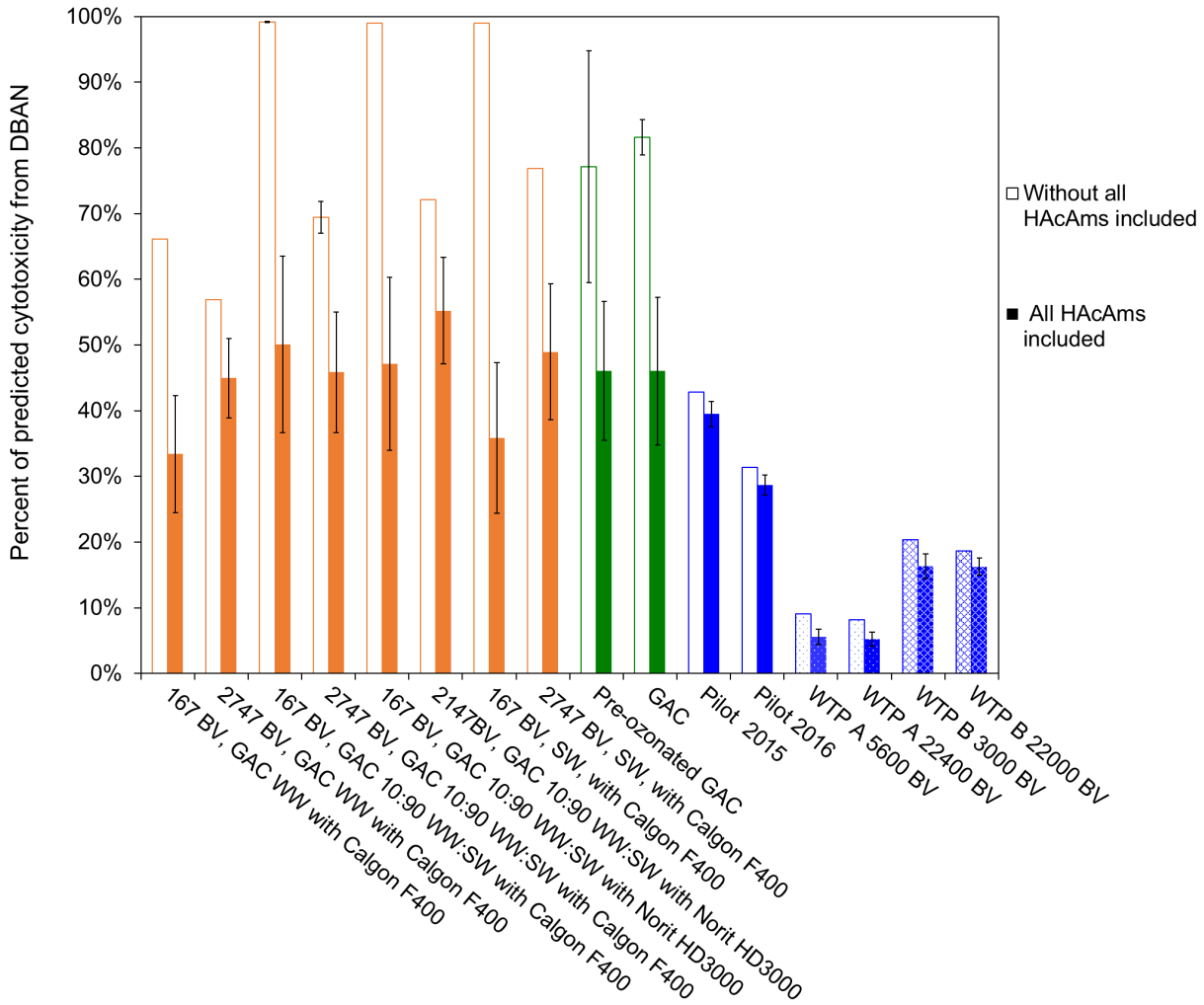

Figure 4 Contribution of DBAN to predicted cytotoxicity of GAC treated samples in which HAcAms were not measured (published data) or with addition of aggregated HAcAm data, indicating that inclusion of additional HAcAms decreases the significance of DBAN. Orange data is from Krasner *et al.*,[18] (Table S6) green data is from McKie *et al.*,[36] (Table S6) and blue data is from Stanford *et al.*[25] and Cuthbertson *et al.*[31] (Table S5) Stanford *et al.*[25] and Cuthbertson *et al.*[31] measured several HAcAms, others were supplemented. The pilot plants included DCAM, DBAM, and BCAM, WTP A included DCAM and BCAM, and WTP B included DCAM, DBAM, BCAM, and TCAM. McKie *et al.*[36] did not measure DBAN; the mean GAC effluent DBAN concentrations from Krasner *et al.*[18], Stanford *et al.*[25] and Cuthbertson *et al.*[31] are presented. Error bars for data including HAcAms (filled bars) are derived from the standard deviation of HAcAm data from Kosaka *et al.*[37] and the HAcAms measured in Stanford *et al.*[25] and Cuthbertson *et al.*[31] (DCAM, DBAM, BCAM, TCAM). Error bars for data without HAcAms (open bars) are derived from the publications. Raw data provided in Table S7.

Statistical Methods in Summed Calculations

Like any measurement, predicted toxicity has some statistical uncertainty. There is uncertainty in both the measurement of a DBP's concentration, and the measurement of its toxic potency. However, the standard deviation of the predicted toxicity is not reported, or is in some cases reported incorrectly, potentially leading to a misunderstanding of the measurement's precision.

Regarding the DBP concentration, during quantification of compounds at low µg/L or low ng/L, relative standard deviation (i.e., standard deviation divided by the mean) of 20% is generally considered acceptable, and some highly genotoxic to CHO cell DBPs regularly occur at or near their limit of quantification (e.g., DBAN). One way to reduce the measurement error is through replicate measurement. However, replicate measurement only accounts for measurement error. If the goal is to compare water treatment processes, it is necessary to measure replicate samples from the experiment to account for both experimental and measurement error. This becomes cost prohibitive, and many data are reported with only measurement replication, rather than experimental.

Regarding measurement of toxic potency, CHO cytotoxicity and genotoxicity assays are considered relatively precise among *in vitro* bioanalytical assays. For example, Wagner and Plewa[12] used a bootstrap method to estimate a relative standard error of 12% for the cytotoxic potency of chloroacetamide. While it is possible to estimate the standard error of the toxic potency of a DBP using the raw data and a bootstrap method, this descriptive statistic has not been published for the majority of DBPs tested with the CHO comet assays. Nevertheless, the toxic potencies measured by these assays also have some uncertainty which should be considered when using them to compare DBPs or water samples.

Multiplying two uncertain values increases the overall standard error. Treating the DBP concentration and its geno- or cytotoxic potency as independent random variables, the standard error of their product is:

$$s_{A\times B} = \sqrt{\left(s_A{}^2 + \bar{x}_A{}^2\right)\left(s_B{}^2 + \bar{x}_B{}^2\right) - \bar{x}_A{}^2\bar{x}_B{}^2}$$

(Eqn. 1)

Where A is the DBP molar concentration, B is the toxic potency ($1/LC_{50}$ or 1/50% DNA tail moment), *s* is standard error, and $\bar{x}$ is mean DBP concentration or mean bootstrap output. For example, for a DBP with concentration measurement relative standard error of 20% and with geno- or cytotoxic potency relative standard error of 12%, the toxicity-weighted concentration standard error is 23.4%. The assumption of independence is valid in this case because there is no relationship between the result of a toxicity assay on a DBP and that DBP's concentration in a sample collected years and miles apart.

When adding random variables, the relative standard error decreases, but to an extent that depends on how much one variable dominates the equation. The standard error for the sum of independent random variables is:

$$s_{(Z_1+Z_2+\cdots Z_n)} = \left(\sum_{i=1}^{i=n} s_{Z_i}{}^2\right)^{\frac{1}{2}}$$

(Eqn. 2)

Where n is the number of variables summed and Z_1, Z_2, etc. are the variables summed. Consider a hypothetical scenario in which a water sample has 30 detected DBPs, each of which has a relative standard error of 20% for the product of DBP concentration and toxic potency. If each DBP contributes to the predicted toxicity equally, the overall relative standard error is just 3.7%. This low relative standard error is because it is unlikely that all 30 DBPs would have been

low estimates in a single sample (assuming independence), and any one extreme value by a single DBP represents a low percentage of the total predicted toxicity. However, if a single DBP contributes 50% of the index (e.g., DBAN) and the other 29 detected DBPs contribute equally to the other 50%, the overall relative standard error is 10.2%. Additionally, the concentration of multiple DBPs measured in a sample may not be completely independent, since the same factors that might dilute, concentrate, or contaminate the measurement of one DBP could also affect the others. Considering covariance, Eqn. 2 becomes:

$$s_{(Z_1 + Z_2 + \cdots Z_n)} = \sqrt{\sum_{i=1}^{i=n} {s_{Z_i}}^2 + \frac{2\sum_{i,j:i<j} cov(Z_i, Z_j)}{N}}$$

(Eqn. 3)

Where N is sample size. Note that a large sample size is needed to provide a valuable estimate of the covariance between each DBP and the number of covariance terms is N*(N-1)/2 (hundreds or even thousands for 30+ DBPs), meaning calculating covariance may not be practical under typical sampling campaigns. But, considering covariance, the true overall standard error of the predicted toxicity could be somewhat higher than calculated by Eqn. 2

As an example of how rigorous analysis of error may alter interpretation, we examine one study in which it was observed that pre-chlorination of surface water before GAC resulted in a lower predicted cyto- and genotoxicity than GAC alone.[25] The predicted cyto- and genotoxicity were reduced 17% and 16%, respectively, if pre-chlorination was applied before the GAC. In Figure 5 we show the predicted cyto- and genotoxicities with error bars assuming a relative standard error of 12% for all DBP toxic potencies and 15% for all DBP concentrations. Based on these assumptions and Eqns. 1 and 2, the relative standard errors of the predicted cyto- and genotoxicities are 10.3% and 10.6%, respectively, before treatment with GAC. After GAC treatment, the relative standard errors of the predicted cyto- and geno-toxicities are 9.3% and

9.9%, respectively. Although in this case a change across GAC is statistically significant, the predicted toxicities with and without pre-chlorination are within two standard errors of each other, and thus not statistically significant (p-value > 0.05). Large experimental replication would have been required to reduce the standard error and verify a change in predicted toxicities of this magnitude. Given that descriptive statistical measures of variance are not generally available and that there is complex interplay between standard errors, small changes in predicted toxicities should be interpreted with caution.

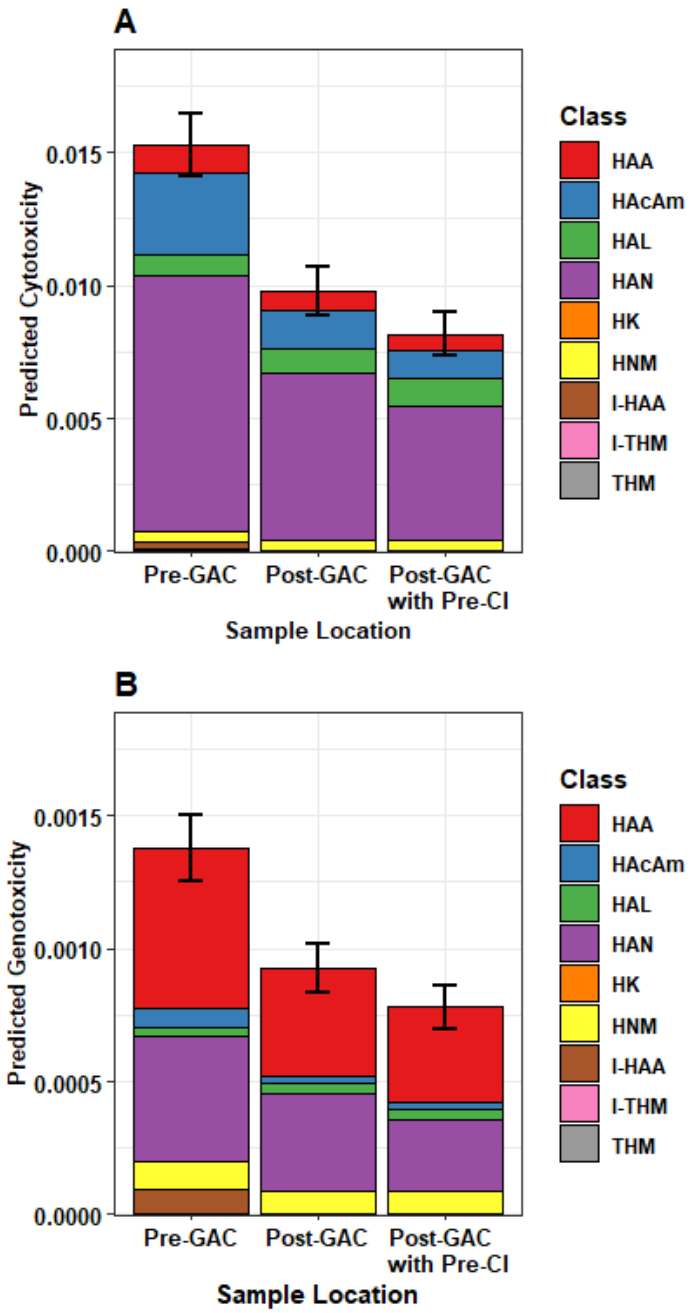

Figure 5 Predicted toxicities of a pilot plant treating surface water with GAC with and without chlorination before GAC, including measurements of error, which are not frequently presented. (A) Cytotoxicity and (B) Genotoxicity. Water quality and treatment details are in Stanford *et al.*[25] HAA = haloacetic acids (non-iodinated), HAcAm = haloacetamides, HAL = haloacetaldehydes, HAN = haloacetonitriles, HK = haloketones, HNM = halonitromethanes, I-HAA = iodinated haloacetic acids, I-THM = iodinated trihalomethanes. THM = trihalomethanes (non-iodinated).

Conclusions and Future Research Needs

Predicted toxicity has been used previously to show that regulated DBPs (THMs and HAAs) contribute much less to the overall toxicological profile of a treated water sample than other DBPs that are present at significantly lower concentrations (i.e., DBAN tends to contribute more to toxicity than THMs and therefore is likely to be more important). This is a function of individual DBPs toxicity index and its concentration. Predicted toxicity is a valuable tool for determining primary contributors to DBPs among the DBPs measured, and is one of many approaches for determining the potential public health effects of DBPs. But we show here that the uncertainties inherent to the method render it challenging and requiring careful interpretation for comparing treatment processes (i.e., GAC treated water is more or less toxic than untreated water). Comparisons between treated and untreated samples using predicted toxicity may be biased towards measured DBP species that have both high toxicity indices in CHO cell assays and precursors that are unaffected by the treatment being studied. Other methods exist to compare toxicity between samples, such as bioassays, but they also have limitations. Primarily that they require extraction of the DBPs to produce a sample that is concentrated enough to produce a response, and the extraction step causes the loss of most volatile DBPs, and likely some unknown DBPs. Further, there are many bioassays that measure various endpoints and it is not yet known which is the most relevant in capturing the human health impacts of DBPs.

In the short term, further research is needed to viably advance predicted toxicity and other toxicity measurements to determine the benefits of a water treatment technology. Additional research to determine how well predicted toxicity and CHO cell toxicity are correlated with other whole mixture bioassays (e.g., SOS Chromotest) would be valuable and would determine if cost effective and quick assays are representative of overall toxicity. Continuation of the discovery of DBPs and their respective toxic potency will continue to improve our understanding of the importance of specific DBPs. If it were possible to measure all DBPs and

their toxicity indices, predictive toxicity would no longer be subject to sampling bias, but this is not possible in the short term, and likely will not be in the long term either, and therefore we must accept that certainty may not be within our grasp. However, better availability and use of metrics of statistical certainty and uncertainty would help to definitively determine if technologies are effective in reducing overall toxicity.

Another short-term goal for DBP researchers should be to assess the role of agonism or antagonism in DBP mixtures, which may be achieved by comparing the predicted toxicity of a clean mixture to that of its actual toxicity to CHO cells.[39] Predicted toxicity assumes that the toxicity of each DBP is additive and ignores the possibility of agonistic or antagonistic effects. Toxicity is generally additive if each compound is toxic through a different mechanism. However, prevailing evidence suggests that DBPs are genotoxic though indirect DNA damage and products of oxidative stress (i.e., similar mechanisms).[11,40,41]

Toxicity threshold values should also be incorporated into predicted toxicity, because some DBPs could be below a threshold concentration at which they would pose no cytotoxic risk. DBPs that are directly genotoxic by chemically reacting with DNA theoretically have no toxicity threshold.[42] However, DBPs that are indirectly carcinogenic through cytotoxicity or oxidative stress are expected to have toxicity thresholds below which they pose zero risk.[42] Ideally, a DBP that is detectable but below this threshold should be excluded from any metric of total DBP risk. Lowest observed effect levels have been published for the CHO genotoxicity and cytotoxicity assays on DBPs,[12] and could be used to exclude DBPs below these concentrations.

A long-term goal may be to assess the differences in toxic response between hamster ovary cells or other bioassays and target human organs. For example, some DBPs are only toxic after hepatic metabolism and liver S9 activation has been developed to act as a surrogate.[43-45] We must accept that both the long- and short-term goals presented here are significant challenges, and that obtaining perfection may not be attainable in the near future.

However, we believe that overcoming the challenges presented will help to guide and understand the implications of future regulatory action.

Acknowledgements

This research was partially supported by the National Science Foundation under Grant No. 1804255 and the Water Research Foundation (Project #5005, managed by Djanette Khiari).

References

1. U.S. EPA, *National Primary Drinking Water Regulations: Stage 2 Disinfectants and Disinfection Byproducts Rule* **2006**: Fed. Regist. 71 p. 388-492.
2. Norman, T. S., Harms, L. L., and Looyenga, R. W., The Use of Chloramines to Prevent Trihalomethane Formation. *Journal American Water Works Association*, **1980**. 72(3): p. 176-180.
3. Seidel, C. J., McGuire, M. J., Summers, R. S., and Via, S., Have Utilities Switched to Chloramines? *Journal (American Water Works Association)*, **2005**. 97(10): p. 87-97.
4. Li, C., Trends and Effects of Chloramine in Drinking Water. *Water Conditioning & Purification*, **2011**. 53(10): p. 52-56.
5. Guay, C., Rodriguez, M., and Serodes, J., Using Ozonation and Chloramination to Reduce the Formation of Trihalomethanes and Haloacetic Acids in Drinking Water. *Desalination*, **2005**. 176(1-3): p. 229-240.
6. Li, X.-F. and Mitch, W. A., Drinking Water Disinfection Byproducts (Dbps) and Human Health Effects: Multidisciplinary Challenges and Opportunities. *Environmental Science & Technology*, **2018**. 52(4): p. 1681-1689.
7. Stalter, D., O'Malley, E., von Gunten, U., and Escher, B. I., Fingerprinting the Reactive Toxicity Pathways of 50 Drinking Water Disinfection by-Products. *Water Research*, **2016**. 91: p. 19-30.
8. Plewa, M. J., Wagner, E. D., and Richardson, S. D., Tic-Tox: A Preliminary Discussion on Identifying the Forcing Agents of Dbp-Mediated Toxicity of Disinfected Water. *Journal of Environmental Sciences*, **2017**. 58: p. 208-216.
9. Plewa, M. J., Kargalioglu, Y., Vankerk, D., Minear, R. A., and Wagner, E. D., Mammalian Cell Cytotoxicity and Genotoxicity Analysis of Drinking Water Disinfection by Products. *Environmental and molecular mutagenesis*, **2002**. 40(2): p. 134-142.
10. Hanigan, D., Truong, L., Simonich, M., Tanguay, R., and Westerhoff, P., Zebrafish Embryo Toxicity of 15 Chlorinated, Brominated, and Iodinated Disinfection by-Products. *Journal of Environmental Sciences*, **2017**. 58: p. 302-310.
11. Lan, J., Rahman, S. M., Gou, N., Jiang, T., Plewa, M. J., Alshawabkeh, A., and Gu, A. Z., Genotoxicity Assessment of Drinking Water Disinfection Byproducts by DNA Damage and Repair Pathway Profiling Analysis. *Environmental Science & Technology*, **2018**. 52(11): p. 6565-6575.
12. Wagner, E. D. and Plewa, M. J., Cho Cell Cytotoxicity and Genotoxicity Analyses of Disinfection by-Products: An Updated Review. *Journal of Environmental Sciences*, **2017**. 58: p. 64-76.
13. Parvez, S., Rice, G. E., Teuschler, L. K., Simmons, J. E., Speth, T. F., Richardson, S. D., Miltner, R. J., Hunter, E. S., Pressman, J. G., Strader, L. F., Klinefelter, G. R., Goldman, J. M., and Narotsky, M. G., Method to Assess Component Contribution to Toxicity of Complex Mixtures: Assessment of Puberty Acquisition in Rats Exposed to Disinfection Byproducts. *Journal of Environmental Sciences*, **2017**. 58: p. 311-321.
14. Pals, J. A., Wagner, E. D., Plewa, M. J., Xia, M., and Attene-Ramos, M. S., Monohalogenated Acetamide-Induced Cellular Stress and Genotoxicity Are Related to Electrophilic Softness and Thiol/Thiolate Reactivity. *Journal of Environmental Sciences*, **2017**. 58: p. 224-230.
15. Li, J., Wang, W., Moe, B., Wang, H., and Li, X.-F., Chemical and Toxicological Characterization of Halobenzoquinones, an Emerging Class of Disinfection Byproducts. *Chemical Research in Toxicology*, **2015**. 28(3): p. 306-318.

16. Laingam, S., Froscio, S. M., Bull, R. J., and Humpage, A. R., In Vitro Toxicity and Genotoxicity Assessment of Disinfection by-Products, Organic N-Chloramines. *Environmental & Molecular Mutagenesis*, **2012**. 53(2): p. 83-93.
17. Chuang, Y. H., Szczuka, A., and Mitch, W. A., Comparison of Toxicity-Weighted Disinfection Byproduct Concentrations in Potable Reuse Waters and Conventional Drinking Waters as a New Approach to Assessing the Quality of Advanced Treatment Train Waters. *ENVIRONMENTAL SCIENCE & TECHNOLOGY*, **2019**. 53(7): p. 3729-3738.
18. Krasner, S. W., Lee, T. C. F., Westerhoff, P., Fischer, N., Hanigan, D., Karanfil, T., Beita-Sandí, W., Taylor-Edmonds, L., and Andrews, R. C., Granular Activated Carbon Treatment May Result in Higher Predicted Genotoxicity in the Presence of Bromide. *Environmental Science & Technology*, **2016**. 50(17): p. 9583-9591.
19. Richardson, S. D., Plewa, M. J., Wagner, E. D., Schoeny, R., and DeMarini, D. M., Occurrence, Genotoxicity, and Carcinogenicity of Regulated and Emerging Disinfection by-Products in Drinking Water: A Review and Roadmap for Research. *Mutation Research/Reviews in Mutation Research*, **2007**. 636(1–3): p. 178-242.
20. Zhang, Y., Chu, W., Yao, D., and Yin, D., Control of Aliphatic Halogenated Dbp Precursors with Multiple Drinking Water Treatment Processes: Formation Potential and Integrated Toxicity. *Journal of Environmental Sciences*, **2017**. 58: p. 322-330.
21. Liu, C., Olivares, C. I., Pinto, A. J., Lauderdale, C. V., Brown, J., Selbes, M., and Karanfil, T., The Control of Disinfection Byproducts and Their Precursors in Biologically Active Filtration Processes. *Water Research*, **2017**. 124: p. 630-653.
22. Le Roux, J., Plewa, M. J., Wagner, E. D., Nihemaiti, M., Dad, A., and Croué, J.-P., Chloramination of Wastewater Effluent: Toxicity and Formation of Disinfection Byproducts. *Journal of Environmental Sciences*, **2017**. 58: p. 135-145.
23. Kristiana, I., Liew, D., Henderson, R. K., Joll, C. A., and Linge, K. L., Formation and Control of Nitrogenous Dbps from Western Australian Source Waters: Investigating the Impacts of High Nitrogen and Bromide Concentrations. *Journal of Environmental Sciences*, **2017**. 58: p. 102-115.
24. Zeng, T., Plewa, M. J., and Mitch, W. A., N-Nitrosamines and Halogenated Disinfection Byproducts in U.S. Full Advanced Treatment Trains for Potable Reuse. *Water Research*, **2016**. 101: p. 176-186.
25. Stanford, B. D., Selbes, M., Reinert, A., Pierce, M., Rosenfeldt, E., Knappe, D. R. U., Maness, C., Zhang, C., Summers, R. S., Mulhern, R. E., Richardson, S. D., Cuthbertson, A., Kimura, S. Y., Liberatore, H., Dickenson, E. R. V., Verdugo, E., Glover, C., Ghosh, A., and Seidel, C. J., Gac Control of Regulated and Emerging Dbps of Health Concern. *Water Research Foundation Final Report*, **2019**. Water Research Foundation: Denver, CO.
26. Chuang, Y.-H. and Mitch, W. A., Effect of Ozonation and Biological Activated Carbon Treatment of Wastewater Effluents on Formation of N-Nitrosamines and Halogenated Disinfection Byproducts. *Environmental Science & Technology*, **2017**. 51(4): p. 2329-2338.
27. Liu, C., Ersan, M. S., Plewa, M. J., Amy, G., and Karanfil, T., Formation of Regulated and Unregulated Disinfection Byproducts During Chlorination of Algal Organic Matter Extracted from Freshwater and Marine Algae. *Water Research*, **2018**. 142: p. 313-324.
28. Li, Z., Liu, X., Huang, Z., Hu, S., Wang, J., Qian, Z., Feng, J., Xian, Q., and Gong, T., Occurrence and Ecological Risk Assessment of Disinfection Byproducts from Chlorination of Wastewater Effluents in East China. *Water Research*, **2019**. 157: p. 247-257.
29. Vatankhah, H., Szczuka, A., Mitch, W. A., Almaraz, N., Brannum, J., and Bellona, C., Evaluation of Enhanced Ozone-Biologically Active Filtration Treatment for the Removal

of 1,4-Dioxane and Disinfection Byproduct Precursors from Wastewater Effluent. *ENVIRONMENTAL SCIENCE & TECHNOLOGY*, **2019**. 53(5): p. 2720-2730.

30. Furst, K. E., Coyte, R. M., Wood, M., Vengosh, A., and Mitch, W. A., Disinfection Byproducts in Rajasthan, India: Are Trihalomethanes a Sufficient Indicator of Disinfection Byproduct Exposure in Low-Income Countries? *Environmental Science & Technology*, **2019**. 53(20): p. 12007-12017.
31. Cuthbertson, A. A., Kimura, S. Y., Liberatore, H. K., Summers, R. S., Knappe, D. R. U., Stanford, B. D., Maness, J. C., Mulhern, R. E., Selbes, M., and Richardson, S. D., Does Granular Activated Carbon with Chlorination Produce Safer Drinking Water? From Disinfection Byproducts and Total Organic Halogen to Calculated Toxicity. *Environmental science & technology*, **2019**. 53(10): p. 5987-5999.
32. Richardson, S. and Postigo, C., *Drinking Water Disinfection by-Products*, in *Emerging Organic Contaminants and Human Health*, D. Barceló, Editor. 2012, Springer Berlin Heidelberg. p. 93-137.
33. Peng, B., Liu, M., Han, Y., Wanjaya, E. R., and Fang, M., Competitive Biotransformation among Phenolic Xenobiotic Mixtures: Underestimated Risks for Toxicity Assessment. *Environmental Science & Technology*, **2019**. 53(20): p. 12081-12090.
34. Yu, Y. and Reckhow, D. A., Kinetic Analysis of Haloacetonitrile Stability in Drinking Waters. *Environmental Science & Technology*, **2015**. 49(18): p. 11028-11036.
35. Plewa, M. J., *Comparative Quantitative Toxicology of the Haloacetonitriles: Forcing Agents of Water Disinfection by-Product Toxicity*. **2019**: Personal Communication.
36. McKie, M. J., Taylor-Edmonds, L., Andrews, S. A., and Andrews, R. C., Engineered Biofiltration for the Removal of Disinfection by-Product Precursors and Genotoxicity. *Water Research*, **2015**. 81: p. 196-207.
37. Kosaka, K., Ohkubo, K., and Akiba, M., Occurrence and Formation of Haloacetamides from Chlorination at Water Purification Plants across Japan. *Water Research*, **2016**. 106: p. 470-476.
38. Chuang, Y.-H., Szczuka, A., Shabani, F., Munoz, J., Aflaki, R., Hammond, S. D., and Mitch, W. A., Pilot-Scale Comparison of Microfiltration/Reverse Osmosis and Ozone/Biological Activated Carbon with Uv/Hydrogen Peroxide or Uv/Free Chlorine Aop Treatment for Controlling Disinfection Byproducts During Wastewater Reuse. *Water Research*, **2019**. 152: p. 215-225.
39. Fairey, J. L., Ghebremedhin, Y., Pham, H., and Zhang, W., Assessing Mutagenicity of Disinfection Byproduct Mixtures Using in Vitro Bioassays. *American Water Work Association Water Quality Technology Conference, Dallas, TX*, **2019**.
40. Pals, J. A., Ang, J. K., Wagner, E. D., and Plewa, M. J., Biological Mechanism for the Toxicity of Haloacetic Acid Drinking Water Disinfection Byproducts. *Environmental Science & Technology*, **2011**. 45(13): p. 5791-5797.
41. Pals, J., Attene-Ramos, M. S., Xia, M., Wagner, E. D., and Plewa, M. J., Human Cell Toxicogenomic Analysis Linking Reactive Oxygen Species to the Toxicity of Monohaloacetic Acid Drinking Water Disinfection Byproducts. *Environmental Science & Technology*, **2013**. 47(21): p. 12514-12523.
42. Hrudey, S. E., Chlorination Disinfection by-Products, Public Health Risk Tradeoffs and Me. *Water Res*, **2009**. 43(8): p. 2057-92.
43. Peto, R., Gray, R., Brantom, P., and Grasso, P., Dose and Time Relationships for Tumor Induction in the Liver and Esophagus of 4080 Inbred Rats by Chronic Ingestion of N-Nitrosodiethylamine or N-Nitrosodimethylamine. *Cancer Research*, **1991**. 51: p. 6452-6469.
44. George, J., Rao, K. R., Stern, R., and Chandrakasan, G., Dimethylnitrosamine-Induced Liver Injury in Rats: The Early Deposition of Collagen. *Toxicology*, **2001**. 156: p. 129-138.

45. Wagner, E. D., Osiol, J., Mitch, W. A., and Plewa, M. J., Comparative in Vitro Toxicity of Nitrosamines and Nitramines Associated with Amine-Based Carbon Capture and Storage. *Environ Sci Technol*, **2014**. 48(14): p. 8203-11.

TOC Art