

RECYCLED WATER POLICY PEER REVIEW

Dr. Richard M. Gersberg

A. INTRODUCTION

I have been asked to serve as a reviewer and to provide review and make comments on the Draft Final Amendment to the Recycled Water Policy of the State Water Resources Control Board which presents monitoring requirements for constituents of emerging concern (CECs) into permit requirements, and on the scientific framework on which the Draft Final Amendment was based contained in the Final Report of the Science Advisory Panel (SAP) (2010) entitled “Monitoring Strategies for Chemicals of Emerging concern (CECs) in Recycled Water.” I would like to first say that Science Advisory Panel’s Report is of extremely high quality, and considering the state of the science in emerging chemicals of concern, the SAP report makes a deliberate and well thought out attempt in developing a possible framework of monitoring strategies. However, a major concern I have is that when the potential universe of hundreds of chemicals is boiled down (by the SAP and SWRCB in its Draft Final Amendment) to a very few “human health” and “performance indicators” for monitoring compliance, the rationale for choosing these indicators is not clearly presented in either of the aforementioned documents. A main question that arises then, is whether the information gained by monitoring of a small prescribed list of CECs (Table 1 of Draft Final Amendment) is both generalizable to the whole CEC universe of chemicals, and more informative and human health protective than the current approach relying on California Department of Public Health (CDPH) prescribed CEC monitoring as well as technology-based standards and multiple barriers.

To underscore this - on Page iv the Executive Summary of the SAP Report states “The Panel emphasizes that all compounds listed above [that is, the indicators and surrogates ultimately incorporated into the Draft Final Amendment] represent an initial list based on limited data and on a number of qualifying assumptions discussed in the report.” Indeed, my review of the science underlying this “initial list” confirms that it is indeed based on very limited data, and in this way while this list first developed in the SAP Final Report, was very suitable as a framework for method development, occurrence studies and further research, a main question I pose here is whether the underlying science is both sound and sufficiently robust to form the basis for regulatory action. The fact that the Science Advisory Panel itself states itself that the “science of CEC investigation is still

in its infancy” emphasizes the need for further review of the basis and rationale for the CEC monitoring requirements in the Draft Final Amendment as I will present in this document.

B. GENERAL (BIG PICTURE) COMMENTS

However, before I get to my more technical comments on the robustness of the science in the SAP Report and underlying the Draft Final Amendment, I would like to express the more general (big picture) concerns I have, not so much with the SAPs Report itself, but more with the regulatory intent of the Draft Final Amendment based on the SAP’s Report, which presents CEC monitoring requirements for recycled water used for groundwater recharge reuse. These overall concerns are:

1. To the best of my knowledge, most (if not all) reuse projects involving surface spreading or direct injection of treated municipal wastewaters already employ a multiple barrier (technology-based) approach to removing conventional and emerging contaminants alike. Additionally, most (if not all) of these projects are already regulated by the California Department of Public Health (CDPH) which has primary responsibility for establishing water quality standards and treatment technology to protect public health. And even though CDPH’s Draft Groundwater Recharge/Reuse Regulation does not specifically require CEC monitoring, the fact is that most (if not all) of these projects already have some form of CEC monitoring in place, and Independent Advisory Panels formed to work with the water agencies and the CDPH to interpret the CEC data. Therefore, I question whether this Draft Final Amendment to the Recycled Water Policy on CEC monitoring requirements which is primarily focused on public health is duplicative of the regulatory process in California? Indeed, if this new regulatory initiative was instead aimed to promulgate monitoring requirements for discharge of CECs from publically owned wastewater treatment plants (POTWs) into surface waters of the State in order to assess the ecological effects of such discharges, I would find this new regulatory initiative both welcome and needed, as such ecological effects of CECs have been well documented in the scientific literature. However, although initially the SAP Final Report was to consider ecosystem effects of recycling practices, this ecosystem focus rather quickly dropped out of the SAP purview due to the minimal exposure scenarios involved in landscape irrigation. In this way, it is somewhat ironic that the new Draft Final Amendment almost exclusively focuses on the unproven

human health risk of CECs in groundwater recharge projects (which is a sphere already regulated by CDPH), while the known ecological risks of the discharge of sewage and its burden of CECs into surface waters of the State, remain nearly totally unaddressed.

2. Since currently there are no scientific data that suggest that CECs present in recycled water are posing a significant human health risk (indeed none is presented in either the SAP Final report or the Final Draft Amendment), and the science to support precise water quality standard setting (even for indicators) for the myriad of CECs still in its infancy, what I find lacking in the body of the Final Draft amendment is a scientifically-based and logical rationale of why the proposed CEC monitoring program is necessary from a public health/human health risk reduction standpoint, and exactly how (when it is implemented) it will provide a higher level of protection of public health than the technology-based standards (and current CEC monitoring regimes) for recharge/reuse scenarios already in place and overseen by the CDPH. If such justification could be added into the preamble of the Final Amendment, I believe the regulated community would gain confidence that this regulatory action is based more on protecting public health rather than as a reflexive response to the public's (and the media's) perception of the risk of CECs in drinking waters.

3. Additionally, as will be seen in my more technically-based review below, little or no evidence is presented in the Draft Final Amendment (or the SAP Report itself) that the data gained on the very few selected health-based and performance-based indicators (Attachment A of Draft Final Amendment) whose routine monitoring will be required, will provide new information on human health risk and/or insight on treatment performance over and above that which we already know about CEC fate and elimination through the treatment train/multiple barriers of existing reuse and recharge operations. Such a rationale, if provided, would really bolster the need for this new regulatory action and give it human health relevance beyond what it presently conveys.

C. SCIENCE UNDERLYING THE REGULATION

Notwithstanding my general concerns expressed above, I will now detail my specific analysis of certain aspects of the soundness of the supporting science on which the Final Draft Amendment is based.

1. One concern of critical importance, in light of the low (part per trillion) range of the Monitoring trigger Levels (MTLs) for the health indicators, is the accuracy and precision of the analytical methods for detecting these CECs. I do find it disconcerting, that a Drewes et al. (2008) study on which much of the SAP Final Report was based, found a rather high level of uncertainty in the current analytical methods for the CECs. These authors conducted an inter-laboratory comparison of 5 participating analytical laboratories, and concluded that the analytical methods targeted for CEC analysis “exhibited significant variation of recoveries and relative standard deviations indicating the degree of uncertainty that is still involved in the reporting of low ppt-level concentration.” For example, one Lab reported triclosan exceeding 300 ppt in a blank sample, and another reported the presence estradiol at 2 ppt in a blank (the MTL in the Draft Final Amendment of 0.9 ppt).

When the accuracy of the analytical method was tested among the 5 Round Robin Labs, recoveries varied dramatically with a value of 38% -84% for the health and performance indicator (as per the Draft Final Amendment) indicator caffeine, 45% - 93% for the health indicator triclosan, and 34% -166% for the health indicator 17 β -estradiol. And, even more troubling, for two of the other performance indicators (as per the Draft Final Amendment), two out of the 5 labs did not detect any gemfibrozil that was spiked into the samples, and 1 out of 3 Labs that tried could not detect any of the iopromide (and the other only got 22% recovery) that was spiked into the samples.

As for precision in the Round Robin tests, each Lab analyzed 5 replicates of both spiked and unspiked samples of all the targeted CECs. For all the Labs, and all the compounds, precision varied with a large range from 3 -86 %. Moreover, Drewes et al. (2008) concluded that of all the compounds tested “not a single analyte exhibited both high recoveries and low relative standard deviations (RSDs) across all analytical methods employed in this study.” They go on to make the somewhat understated conclusion that “this finding was not expected.” In sum, these authors concluded about CEC analytical testing:

- All methods showed high variations in recovery
- There is a high degree of uncertainty associated with reporting low ppt-level results
- There are clear limitations on the ability of sound laboratory practice to improve recoveries

To summarize, this very well-done inter-laboratory Round Robin revealed the relatively high degree of uncertainty around reporting low ppt (part per trillion) levels. However, since the MTLs for the health indicators (particularly estradiol) are in this same low ppt range, my concern is that the analytical methodology is not yet be sound and reliable enough to support the bright-line thresholds (MTLs) as proposed for the health indicators in the Draft Final Amendment. This will be exacerbated by the fact that as groundwater recharge projects become increasingly common, monitoring might well involve many different laboratories in the future, with some relatively new to CEC analytical chemistry.

2. Another concern I have is that the rationale for each of the health indicators and performance indicators chosen for both surface spreading and direct injection, is never given explicitly and in a straightforward and transparent matter, either in the Draft Final Amendment or in the SAPs Final Report. Why, for example, are caffeine and triclosan chosen as health indicators? Why in particular, are gemfibrozil, iopromide, and DEET chosen for surface spreading applications? And where are the statistical correlations or structure/activity relationships showing how these specific chosen indicators are both generalizable and predictive of the behavior of the whole universe of CECs in a variety of reuse scenarios. Unfortunately, I cannot find the answers to these questions either in the Draft Final Amendment and the SAPs Final Report, and if indeed the selection is based on sound science (even if in its infancy), the scientific rationale is not clearly presented in either the Draft Final Amendment or the SAP Final Report. Rather, the only selection criteria for the indicator compounds that can be found with some detail and support comes from the report by Drewes et al. (2008), where CECs tested were divided into several broad treatment categories with removal ratings for good (> 90%), intermediate (50-90%, and 25-50%, and poor removal (<25%). Apparently, both the SAP Final Report and the Draft Final Amendment relied heavily on the same Drewes et al. (2008) study. In doing so, rather than selecting the indicators based upon their intrinsic structural,

physicochemical, or biodegradation properties, the selection was done empirically from data for a variety of SAT (soil aquifer treatment) or RO operations (from Drewes et al., 2008) thereupon placing the CECs into these very broad treatment categories above. Indeed this classification is so broad, that the overwhelming majority of the CECs tested in both SAT and RO treatment fell into the “good treatment” category, so that that the list turns out to be not very specific (within the very broad % removal ranges given above) or exclusive. Moreover, this selection of a “one size fits all” indicator/surrogate mix found in the SAP Final Report and the Draft Final Amendment (Table 1) seems to be discount one caveat given by Drewes et al. (2008) that “because of geographic and temporal variations in the occurrence pattern of certain wastewater-derived contaminants.....the determination of a given indicator/surrogate monitoring framework for a given treatment train will likely vary from site to site.”

In sum , what the Final Amendment needs in my view is a clear and concise scientific rationale in its preamble of why the specific health and performance indicators shown in Tables 1 and 6 of the Draft Final Amendment were chosen, with an explanation of how (such as with statistical correlations) we can be assured that the behavior of the very few selected health and performance indicators are both generalizable and predictive in the range of treatment trains, as well as geographic and temporal conditions of future recharge scenarios.

3. Now I will turn my attention to the selection of the surrogates to be monitored as per the Draft Final Amendment and the SAP Final Report. As for the choice of performance and health indicators, what science there is to support the surrogate choice was found mostly in Drewes et al., 2008 and 2010a. However, what came as somewhat of a surprise to me was the conclusion stated by Drewes et al. (2008) that the “majority of surrogate parameters are not strongly correlated with the removal of indicator compounds occurring at the nanograms-per-liter (ppt) concentrations.” However, nearly in the same breath, these same authors conclude that “Enhanced removal of select surrogate parameters correlated with improved removal of indicator compounds.” So here, there seems to be contradictory statements being made - that is, does the removal of surrogates correlate

with the indicators or does it not? And if so, which ones correlate, and what is their degree of correlation?

To try to answer these questions we need to turn to the supporting work of Drewes et al, 2008 and 2010a where the practicability of selected surrogates (i.e., the ease with which it can be monitored and its predictive ability) was assessed through pilot- and full-scale monitoring efforts. In Drewes et al. (2008), a variety of surrogate parameters (including BDOC, TOC, TOX, TOI, COD, fluorescence, and UVA as well as nitrate-N and ammonia-N) were evaluated at 3 different full-scale SAT facilities. Significant changes in several of these surrogate parameters were detected and within 2 weeks (of travel time) TOC, UVA, COD, and TOI (TOC and UVA are in the Draft Final Amendment) were reduced to levels such that they became “limited in reflecting additional transformations of the organic matter during subsequent travel in the subsurface.” And in spite of the fact that the authors somehow concluded that “sensitive surrogate parameters exist to describe the biological activity in the subsurface”, no statistical correlation of these surrogate removals with CEC indicator removals could be found in this document by Drewes et al. (2008) (for SAT, at least) that might serve to support the authors conclusion that “Enhanced removal of select surrogate parameters correlated with improved removal of indicator compounds”.

On the other hand, Drewes et al. (2010a) in an analysis of Tucson Water’s Sweetwater Recharge Facility, did find a number of significant ($p < 0.05$) positive correlations between the surrogates TOC and TOX and 7 specific CECs (in Table 3.12 of Drewes et al., 2010a), but from their findings they go on to state that at low TOC concentrations (< 2 mg/L), TOC monitoring “would not be a sufficient surrogate parameter to assess the removal of trace organic chemicals during spreading-basin operation.” This statement itself is somewhat surprising in light of the fact that the Draft Final Amendment (Table 6) promulgates TOC (along with nitrate ammonia and UV absorption) as applicable for surface application recharge scenarios. Additionally, although Drewes et al. (2010a) measured as many as 33 different CECs in Tucson Water’s operation, they say nothing in this document about whether or not any of the correlations were significant for the majority (as many as 26) of the remaining CECs that were monitored. This Drewes et al. (2010a) study then raises more questions than it answers, since for all the health or performance indicators (7 in total)

recommended for surface application in Table 6 of the Draft Final Amendment, only for one of these (gemfibrozil) was a significant positive correlation association noted by Drewes et al. (2010a) as a function of surrogate (in this case, TOC and TOX) removal.

With regard for AOP and RO processes, the degree of evidence presented that surrogate removal is predictive of CEC removal is a bit higher. For example, conductivity has long been a surrogate parameter of choice in assessing membrane performance at RO operations. An earlier study by Drewes et al., (2005) had demonstrated that conductivity in permeate samples was more strongly correlated to the presence of low-molecular weight and neutral trace organic compounds than was TOC. However, a positive correlation from this earlier Drewes et al. (2005) study was shown only for caffeine and conductivity and cited in Figure 5.17 of Drewes et al. (2008), but again this is only for a single CEC indicator. The SAP Final Report and the Draft Final Amendment both propose a variety of indicators and surrogates for surface application and subsurface application using RO, and cite the work by Drewes et al., 2008, 2010a, 2010b, and Dickenson et al., 2009 as showing the predictive ability of surrogates in monitoring treatment, and the SAP Final Report specifically states that “ these studies demonstrate that changes in bulk parameters do correlate with changes of indicator chemicals in the subsurface or during RO treatment leading to direct injection.” However from my review of these same studies, particularly for SAT, but even in the case of RO treatment, the scientific evidence as reflected by significant statistical correlations between surrogates (aside from conductivity) and indicator removal is either lacking or not presented.

4. In both the SAP Final report and the Draft Final Amendment itself (see Attachment A) nowhere could I find an explicit explanation of why and how caffeine and triclosan were chosen among the whole universe of CECs as “health indicators”. I would have hoped to have seen a strong justification for each based on characteristics of their structure/activity or toxicology that make them ideal “health indicators” so that their removal and fate in a specific treatment process makes their behavior as a “health” indicator generalizable to the universe of CECs that may have adverse health effects. Rather, the selection of both caffeine and triclosan as health indicators seemed to be made (Section 8.2 of the SAP Final Report) strictly on the basis of the fact that their measured environmental concentrations

(MEC) were high enough so that a value of 1 for the MEC/MTL (monitoring trigger level) was exceeded.

This, and the fact that in the Draft Final Amendment, the comparison of measured concentrations by a recycled water producer as compared to the MTL (monitoring trigger level), could initiate regulatory response actions suggests that the question of how MTLs were derived for caffeine and triclosan is more than an academic issue. An examination of this selection leads to the finding that both CECs (caffeine and triclosan) have the same MTL of 0.35 µg/L, and these values were in turn based on drinking water guidelines established by Australia. However, even these Australian Guidelines for Water Recycling (2008) did not generate these MTL levels *a priori*, but instead derived them from a pragmatic approach (Threshold of Toxicological Concern or “TTC”) to setting standards for chemicals based on their structure according to the Cramer classification scheme (Kroes and Koziarowski, 2002; Kroes et al., 2004). Further examination shows that Australia derived their caffeine drinking water guideline (which was adopted in the Draft Final Amendment) by assuming that caffeine is a Structural Class III compound which assumes the chemical “for which structural features or likely metabolic pathways either permit no strong presumption of safety, or actually suggest significant toxicity.” But for caffeine, even the SAP Final Report admits that the above assumption of caffeine’s structural class (along with the associated uncertainty factor of 1,500) is an “exceptionally conservative guideline and initial MTL,” and points out that the concentration of caffeine in coffee is about one million times greater than the initial MTL!

And yet, despite the convoluted pathway (from the TTC approach to adoption by the Australian government, to citation in the SAP Final Report, to promulgation in the Draft Final Amendment) this admittedly (by the SAP) overly conservative MTL was adopted in the Draft Final Amendment. So it seems that the Final Amendment is basing both its selection of caffeine as a “health” indicator and its MTL value which could trigger response actions, on a generic toxicity classification scheme (TTC principle) without a *de novo* analysis of the specific toxicity of caffeine. In this way, caffeine is then placed in the same group of Class III substances as 2,4 dinitrotoluene and chlorobenzene, both of these commonly found at

Superfund sites and listed by the ATSDR (Agency for Toxic Substances & Disease Registry) as causing significant threats at these sites to human health.

Moreover, another health indicator as listed in Draft Final Amendment that is placed *in* this same group of Class III substances as caffeine, is triclosan. The fact that both caffeine and triclosan are placed in the same Class III group, then confers on both of these disparate health indicators the exact same MTL of 0.35 micrograms/L, despite the fact that triclosan has been shown to have a variety of sublethal effects of concern. Recently, the Environmental Protection Agency (EPA) noted that a review of triclosan under the Endocrine Disruptor Screening Program (EDSP) provided evidence of its endocrine disrupting potential. For example, in frogs, triclosan can disrupt thyroid hormone-associated gene expression and induce changes in the thyroid hormone-mediated metamorphosis process (Veldhoen et al., 2006). Triclosan can also alter circulating serum concentrations of total thyroxine in rats (Crofton et al., 2007). However, these low-dose endocrine effects were also not included in the TTC assessment as a basis for the derivation of triclosan's MTL in the Final Draft Final Amendment. Moreover, a new study by Clayton et al. (2011) showed that triclosan may negatively affect human immune function as measured by allergy or hay fever diagnosis. However for such immunologically mediated allergic responses, there were also insufficient dose-response data to include in the TTC assessment (Kroes et al., 2004). The fact that both caffeine and triclosan have the same MTL points despite their disparate effects and toxicity endpoints, serves to highlight potential deficiencies of the generic TTC approach on which both the SAP Final Report and the Final Amendment based their derivation of the MTLs.

In sum then the main points I would like to emphasize are:

- No rationale or analyses is presented in either document on the existing degree of public health risk posed by the presence of CECs after surface spreading and direct recharge scenarios, and because of this, the benefit in public health risk reduction by the requirements of this new Final Draft Amendment are not only unknown, but they cannot be compared to the current regulatory approach by the CDPH which consists of both a treatment technique and also certain CEC monitoring requirements.

- No cogent rationale is presented for why the health indicators (particularly caffeine and triclosan) were chosen as “health” indicators, and no good quantitative relationships (as in statistical correlations) could be found in the supporting literature to show that the behavior of the performance indicators chosen is generalizable and well correlated to the whole universe of CECs.
- More evidence showing that the behavior of the chosen surrogates (Table 2 of Draft Final Amendment) correlates well to the behavior of the whole universe of CECs is needed, especially at the low part per trillion levels.
- It would be most helpful if the Final Draft Amendment included a detailed description of how the MTLs were derived for each of the 4 health indicators, with an explicit explanation of underlying the toxicological data for the NOELs, and the safety factors applied.

REFERENCES

Australian Guidelines for Water Recycling: Augmentation of Drinking Water Supplies (2008). National Health and Medical Research Council, Environmental Protection and Heritage Council, and the Natural Resource Management Ministerial Council..

Clayton, E.M.R., Todd, M., Dowd, J.B. and A.E. Aiello (2011). The Impact of Bisphenol A and Triclosan on Immune Parameters in the U.S. Population, NHANES 2003–2006. *Environ Health Perspect.* 119:390-396.

Crofton, K.M., Paul, K.B, DeVito, M.J. and J.M. Hedge (2007). Short-term *in vivo* exposure to the water contaminant triclosan: Evidence for disruption of thyroxine. *Environ. Toxicol. And Pharm.* 24:194-197.

Dickenson, E.R, Drewes, J.E., Sedlal, D.L., Wert, E.C. and S.A. Snyder (2009). Applying surrogates and indicators to assess removal efficiency of trace organic chemicals during chemical oxidation of wastewaters. *Environ. Sci. Technol.* 43(16):6242-6247.

Drewes, J.E., Bellona, C., Oedekoven, M., Xu, P., Kim, T.-U. and G. Amy (2005). Rejection of wastewater-derived micropollutants in high-pressure membrane applications leading to indirect potable reuse. *Environ. Progress* 24(4):400-409.

Drewes, J.E., Sedlak, D., Snyder, S. and E. Dichenson (2008). Development of Indicators and Surrogates for Chemical Contaminant Removal during Wastewater Treatment and Reclamation (WRF-03-014), Alexandria, VA, WateReuse Research Foundation.

Drewes, J.E., Snyder, S. and E. Dickenson (2010a). Developing of Surrogates to Determine the Efficiency of Groundwater Recharge Systems to Remove Trace Organic Chemical. Final Report (WRF-05-004). Alexandria, VA, WateReuse Research Foundation.

Drewes, J.E., Khan, S.J., McDonald, J.A., Trang, T.T.T. and M.V. Storey (2010b). Chemical monitoring strategy for the assessment of advanced water treatment plant performance. *Water Sci. Technol.* 10 (6): 961-968.

Kroes, R. and G. Kozianowki (2002). Threshold of toxicological concern (TTC) in food safety assessment. *Toxicol. Letters* 127:43-46.

Kroes, R., Renwick, A.G, Cheesman, M., Kleiner, J., Manglesdorf, I., Piersma, A., Schilter, B., Schlatter, J., van Schothorst, F., Vos, J.G. and G. Wurtzen (2004). Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food and Chem. Toxicol.* 42: 65-83.

Veldhoen, N., Skirrow, R.C., Osachoff, H., Wigmore, H., Clapson, D.J., Gunderson, M.P., Van Aggelen, G. and C. Helbing (2006). The bactericidal agent triclosan modulates thyroid hormone-associated gene expression and disrupts postembryonic anuran development. *Aquatic Toxicol.* 80: 217-227.