Review

Development and application of relevance and reliability criteria for water treatment removal efficiencies of chemicals of emerging concern

Astrid Fischer a,b,*, Annemarie P. van Wezel c, Juliane Hollender d,e, Emile Cornelissen f,g,h, Roberta Hofman f, Jan Peter van der Hoek a,i

a TU Delft, Faculty of Civil Engineering and Geosciences, PO Box 5048, 2600, GA, Delft, the Netherlands
b Evides Watercompany, Department of Technology & Sources, the Netherlands
c Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, the Netherlands
d Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600, Dübendorf, Switzerland
e Institute of Biogeochemistry and Pollutant Dynamics (IBP), ETH Zurich, 8092, Zurich, Switzerland
f KWR Watercycle Research Institute, 3433PE, Nieuwegein, the Netherlands
g Singapore Membrane Technology Centre, Nanyang Environment and Water Research Institute, Nanyang Technological University, Singapore, 637141, Singapore
h Particle and Interfacial Technology Group, Ghent University, Coupure Links 653, B-9000, Ghent, Belgium
i Waternet, Strategic Centre, Amsterdam, the Netherlands

Article info

Article history:
Received 22 February 2019
Received in revised form 18 May 2019
Accepted 25 May 2019
Available online 28 May 2019

Keywords:
Data evaluation criteria
Chemicals of emerging concern
Advanced water treatment technologies
Removal efficiency
Drinking water
Wastewater

ABSTRACT

With the growth in production and use of chemicals and the fact that many end up in the aquatic environment, there is an increasing need for advanced water treatment technologies that can remove chemicals of emerging concern (CECs) from water. The current lack of a homogenous approach for testing advanced water treatment technologies hampers the interpretation and evaluation of CEC removal efficiency data, and hinders informed decision making by stakeholders with regard to which treatment technology could satisfy their specific needs.

Here a data evaluation framework is proposed to improve the use of current knowledge in the field of advanced water treatment technologies for drinking water and wastewater, consisting of a set of 9 relevance criteria and 51 reliability criteria. The two criteria sets underpin a thorough, unbiased and standardised method to select studies to evaluate and compare CEC removal efficiency of advanced water treatment technologies in a scientifically sound way.

The relevance criteria set was applied to 244 papers on removal efficiency, of which only 20% fulfilled the criteria. The reliability criteria were applied to the remaining papers. In general these criteria were fulfilled with regards to information on the target compound, the water matrix and the treatment process conditions. However, there was a lack of information on data interpretation and statistics.

In conclusion, a minority of the evaluated papers are suited for comparison across techniques, compounds and water matrices. There is a clear need for more uniform reporting of water treatment studies for CEC removal. In the future this will benefit the selection of appropriate technologies.

© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Introduction

Chemicals are continuously produced for various beneficial purposes, such as protecting crops, conserving food or treatment of diseases. Over 348,000 chemicals are currently registered and regulated via national and international authorities (CHEMLIST). New chemicals enter the market constantly and the global volume of production of chemicals is continuously growing (Bernhardt et al., 2017; CEFIC, 2016; UNEP, 2013). Many of these chemicals and their transformation products enter the aquatic environment during their life cycle (Schwarzenbach et al., 2006). Chemicals of emerging concern (CECs) in the water cycle have been the focus of research for more than 30 years. The main focus has been on assessing their toxicity (Bruce et al., 2010; Schriks et al., 2010), identifying their fate in the environment (Mamy et al., 2015), documenting their occurrence (Loos et al. 2009, 2010, 2013) and minimising their release into the environment together with optimising removal options (Rivera-Utrilla et al., 2013; van Wezel et al., 2017).

To minimise concentrations and thus adverse effects, removal efficiencies of various advanced drinking water and wastewater treatment technologies have been the focus of research (Lee and von Gunten, 2010; Luo et al., 2014; Rivera-Utrilla et al., 2013; Verlicchi et al., 2012; Yang et al., 2017). Advanced water treatment technologies are based on sorption, oxidation and size exclusion principles. The experimental settings in studies on the efficiency of these technologies are not homogeneous. Technologies can be tested at lab-, pilot- or full scale, with different compounds, and under different conditions. Different water matrices can be tested such as de-mineralised water, real or standardized surface water, ground water, drinking water and wastewater. The target compounds can be spiked as single compound, in mixtures with varying concentrations, or environmental samples can be used. There can be variations in the process conditions of the treatment, e.g. dose, contact time or flux. Finally there are variations in how experimental set-ups and results are expressed; in mJ/cm² or W/m² in case of UV oxidation, with Freundlich isotherms or removal percentage in the case of granular activated carbon (GAC), and many others.

These variations, and the resulting ambiguity, obstruct the interpretation and evaluation of data concerning the removal efficiency of CECs of specific treatment technologies. Stakeholders within the urban water cycle have sufficient information on sources, occurrence and risks of CECs and on potential mitigation options, but the relevance and reliability of the information is often unknown (Fischer et al., 2017). As a consequence, a framework for evaluation of scientific and technical information when evaluating removal efficiency studies would be helpful, including criteria for relevance and reliability specified for the technologies. Examples of such data evaluation frameworks from the field of (eco)toxicology are available and well-used, e.g. to identify studies for the derivation of environmental quality standards in a scientifically sound way (Ågerstrand et al., 2011b; Klimisch et al., 1997; Moermund et al., 2016; Roth and Ciffroy, 2016).

The aim of this study is to develop a novel evaluation framework that consists of two criteria sets that can be used to evaluate the i) relevance and ii) reliability of CEC removal efficiency studies for advanced drinking and wastewater treatment technologies. Commonly used advanced drinking and wastewater treatment technologies are activated carbon (granulated or powdered activated carbon, GAC or PAC), the use of ozone (O₃) and UV with or without H₂O₂, and finally nanofiltration (NF) and reverse osmosis (RO) (Luo et al., 2014; van Wezel et al., 2017). Therefore, the criteria developed focus on those technologies. The developed criteria sets are subsequently applied to 244 removal efficiency studies.

2. Methodology

Firstly, existing data evaluation frameworks were explored, to be used as a starting point for the water treatment technology evaluation framework.

Secondly, we identified parameters influencing removal efficiencies of the selected treatment technologies, such as CEC characteristics, characteristics of the water matrix involved and the treatment process settings and characteristics. These impact parameters may differ per treatment technology.

An initial literature review was carried out to gain insight in the selected treatment technologies, using Scopus. Papers describing the water treatment techniques GAC/PAC, Ozone/UV ± H₂O₂ and NF/RO membranes were selected based on their title. The full paper was retrieved based on the abstract, and priority was given to recent reviews. Preliminary impact parameter lists were created and subsequently discussed with five experts from universities, knowledge institutes and water utilities, covering all selected technologies. We used semi-structured face-to-face interviews, and e-mailed the preliminary impact parameter lists ahead of the interview, to give the possibility of addressing additional parameters. The interviews were recorded, and analysed. Based on this, further extensive additional literature research on each individual treatment technique was carried out, to underpin the selected impact parameters, and make sure none were left out. A revised impact parameter list was discussed in a workshop within the EU FP7 SOLUTIONS project. Based on this the final impact parameter list was set.

The selected data evaluation framework was then used to develop a novel framework suited to evaluate the CEC removal...
efficiency studies for advanced water treatment technologies.

Finally we used the developed framework to evaluate the relevance of 244 removal efficiency papers. These papers were found using the Scopus database by searching on removal of CECs and the chosen treatment technologies, and in the reference lists of the retained articles. The 244 papers evaluated for relevance covered 54 journals, with impact factors from 0.7 to 11.6 and included various research fields such as radiation physics, environmental pollution and chemical engineering (Supplementary Information 1). Papers fulfilling the relevance criteria were evaluated with regards to the reliability criteria.

3. Existing data evaluation frameworks

Many frameworks for evaluation of scientific and technical information exist (ECHA, 2011; EPA, 2012; OECD, 1998; USEPA, 2003). Most are developed and applied for the implementation of experimental data in regulatory frameworks or decision making processes. In one of the first generic frameworks for evaluation of scientific and technical information five key factors are highlighted (USEPA, 2003):

1. Soundness - The extent to which the scientific and technical procedures and methods employed to generate the information are reasonable for, and consistent with, the intended application.
2. Applicability and utility - The extent to which the information is relevant for the intended use.
3. Clarity and completeness - The degree of clarity and completeness with which the data, assumptions, methods, quality assurance and analyses employed to generate the information are documented.
4. Uncertainty and variability - The extent to which the quantitative and qualitative variability and uncertainty in the information or in the procedures, measures, methods or models are evaluated and characterized.
5. Evaluation and review - The extent of independent verification, validation and peer review of the information or of the procedures, measures, methods or models.

Besides the generic evaluation frameworks, several scientific fields have their own specific frameworks tailored to their needs. Within the field of (eco)toxicology, thorough and scientifically sound assessment of scientific data is necessary in relation to the hazard and risk evaluation of chemicals (ECHA, 2011; Klimisch et al., 1997; Moermond et al., 2016; Roth and Ciffroy, 2016). In view of the analogy of having to consider the complexity of CECs in various water matrices, these well-developed data evaluation frameworks were considered a useful starting point to develop a data evaluation criteria set for the CEC removal efficiencies.

The first and most used toxicological evaluation framework is the Klimisch approach, focussing on:

1. Reliability - Evaluating the inherent quality of a test report or publication relating to preferably standardized methodologies, and the description of experimental procedures and results to give evidence of the clarity and plausibility of the findings.
2. Relevance - Covering the extent to which data and/or tests are appropriate for a particular intended use of the data, i.e. in the case toxicological evaluation hazard identification or risk characterization.
3. Adequacy - Defining the usefulness of data for the intended (risk assessment) purposes. When there is more than one set of data for each effect, the greatest weight is attached to the most reliable and relevant data set.

The Klimisch approach is used in the REACH legislation for information requirements and chemical safety assessment (ECHA, 2011). The Klimisch framework has been criticised for being too reliant on expert judgement, as it provides few criteria for the reliability evaluation and only mentions relevance with very little guidance on how to evaluate this (Ågerstrand et al., 2011a; Kase et al., 2016; Moermond et al., 2016). Therefore, a more detailed framework for evaluation was developed and tested within the CRED (Criteria for Reporting and Evaluation ecotoxicity Data) project (Kase et al., 2016; Moermond et al., 2016). This framework consists of two criteria sets addressing relevance and reliability. In this framework adequacy is addressed under the heading relevance. The relevance criteria set contains 13 relevance criteria, under the headings general, biological relevance and exposure relevance. The reliability criteria set entails 20 quite specific criteria under the headings general information, test set-up, test compound, test organism, exposure conditions and finally statistical design and biological response. These criteria sets are further elaborated by point by point as to why it is important and how it should be verified (Moermond et al., 2016). This elaboration is followed by a recommendation of 50 points which should be included in a study to be able to evaluate it properly.

The existing (eco)toxicology frameworks have consequently been evaluated by Roth and Ciffroy (2016). As relevance and reliability were seen as important evaluation points, here we considered only frameworks clearly separating and evaluating these criteria, i.e. Ågerstrand, AMORE and CRED (Ågerstrand et al., 2011b; Isigonis et al., 2015; Moermond et al., 2016). The AMORE framework was not selected as it is a computer based decision support system very specifically tailored to the evaluation of ecotoxicity tests and not easily adaptable to other purposes (Ågerstrand et al., 2011b; Isigonis et al., 2015; Moermond et al., 2016; Roth and Ciffroy, 2016). The CRED framework gives a detailed description of what needs to be assessed for the evaluation of relevance and reliability, summarized in two tables. These were used as a starting point to develop a relevance and a reliability criteria set for the evaluation of CEC removal efficiency studies for water treatment technologies.

4. Results

4.1. Impact parameters for water treatment efficiency assessment

The impact parameters which influence CEC removal efficiencies of advanced water treatment technologies concern a) CEC characteristics, b) water matrix characteristics and c) treatment process conditions. These are specified per treatment technology in Table 1. Parameters in bold are unique to each study, other parameters are relevant but can be retrieved elsewhere such as via scientific literature, and databases such as Episuite, Chemsider and OECD toolbox.

4.2. Relevance and reliability criteria for water treatment efficiency assessment

4.2.1. Assessing the relevance of a water treatment study

For relevance we used the definitions of Klimisch et al. (1997) and ECTOC (2009) “Relevance broadly refers to the extent to which data and tests are appropriate (fit-for-purpose) for their intended use, it is a context-dependent quality criterion that is neither intrinsic to a given study per se, nor a function of the information available about that study”. The relevance criteria, appropriateness of data and tests (see Table 2), are used to determine whether a paper is of interest for the specific purpose; if positive the reliability should be assessed.
### Impact parameters per treatment technology

<table>
<thead>
<tr>
<th>Treatment technology</th>
<th>CEC characteristics</th>
<th>Water matrix characteristics</th>
<th>Treatment process conditions</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Granular activated carbon</td>
<td>Molecular charge/pKa Log Kₜw/Log Dₜw</td>
<td>pH NOM concentration (and composition) Temperature</td>
<td>Surface area/grain size Pore volume</td>
<td>De Ridder et al., 2011; Jeiranian et al., 2017; Mailler et al., 2016; Nam et al., 2014; Rossner et al., 2009; Verlicchi et al., 2010</td>
</tr>
<tr>
<td>Powdered activated carbon</td>
<td>Molecular charge/pKa Log Kₜw/Log Dₜw</td>
<td>pH NOM concentration (and composition) Temperature</td>
<td>Surface charge Biological activity Contact time/EBCT Column length Flow through Backflush routine Scale of testing (bench, pilot, full) Prior carbon use (if any)</td>
<td></td>
</tr>
<tr>
<td>Ozone (+H₂O₂) Reactivity Concentration</td>
<td>pH NOM concentration Nitrite/nitrate Bromide/bromate</td>
<td>Dosage of O₃ and H₂O₂ Reactor design (mixing regime) Contact time Scale of testing (bench, pilot, full) Wavelength (lamp type) Irradiation time UV dose and H₂O₂ dosage Reactor design Scale of testing (bench, pilot, full)</td>
<td>Lee et al., 2013; von Sonntag and von Gunten, 2012; Zhang et al., 2012</td>
<td></td>
</tr>
<tr>
<td>UV (+H₂O₂) Reactivity Concentration</td>
<td>pH NOM concentration Nitrite/nitrate Temperature Turbidity</td>
<td></td>
<td>Lester et al., 2008; Liu and Liu, 2004; Pereira et al., 2007; Real et al., 2009; Yang et al., 2014</td>
<td></td>
</tr>
<tr>
<td>NF/RO membranes Molecular weight/size (Molecular weight cut-off) Functional groups Log Kₜw/Log Dₜw</td>
<td>Ionic strength pH</td>
<td>Membrane area Membrane charge/Zeta potential Fouling</td>
<td>Bellona et al., 2004; Hajibabania et al., 2011b; Verliefde et al., 2007b; Yoon et al., 2006</td>
<td></td>
</tr>
</tbody>
</table>


### Relevance of the criteria (criteria numbers from Table 2)

#### 4.2.1.1. Explanation of the relevance criteria (criteria numbers from Table 2)

Relevance of data

1. **Is the scope of the tests appropriate for the evaluation?**

   The requirements for evaluating CEC removal efficiencies by water technology treatments is that the technology is commercially available and full-scale applied in the water sector. Only then can it be used by stakeholders to make informed decisions about relevant treatment technologies. This at current implies that CEC removal studies on activated carbon (GAC and PAC), O₃ (±H₂O₂), UV (±H₂O₂) or NF and RO membranes are considered appropriate. Evidently, in the future this list of commercially available and full-scale applied techniques can be expanded.

2. **Are the data reported appropriate for the evaluation?**

   The purpose is to compare the removal efficiency of one treatment technology with another. This means that if a study reports removal efficiencies in such a way that these cannot be compared across technologies, e.g. break-through curves for granular activated carbon, it is not considered relevant for this purpose. The results should ideally be reported as removal percentage, log units of removal, or influent/effluent concentrations which can be used to calculate the removal efficiency. Studies presenting the removal percentages only in graph form were disregarded as exact removal percentages cannot be retrieved from graphs. Many studies use the connotation > and < before the removal percentage as they cannot be more specific due to the limit of quantification/limit of detection (LOD/LOQ), this severely hampers the possibility of comparing the study to other studies. In case > or < have been used, only studies reporting removal percentages >99% or <1% have been included, as these results are considered equal to 99% or 1% removal. In case of negative removal efficiencies these may reflect < LOQ in influent matrices, fluctuating concentrations, or transformation processes during treatment, and can be relevant to include. However, studies to find the conditions at which a certain percentage of a compound is removed, i.e. where the removal percentage is fixed, are discarded as this is a different way of studying removal.

   **Relevance of tests**
### 3.5 Relevance of target compound(s)

When testing a large number of compounds these should ideally represent a broad range of physicochemical properties. The focus can also be on one compound or a group of compounds that are known to be problematic to remove with other techniques than advanced water treatment techniques, such as persistent mobile organic compounds (PMOCs) (Reemtsma et al., 2016). In case of compounds that degrade easily due to other processes than the treatment studied, it should be verified that the compound is actually removed by the tested treatment technology. This emphasizes the importance of controls. Ideally tested concentrations of the CECs should be representative for environmental levels, but this can be difficult to achieve with compounds with a high LOQ.

### 6-7 Relevance of water matrix

It is important to ensure that the selected water matrix corresponds with the purpose of the study. Especially wastewater has many properties that alter removal efficiencies compared to cleaner matrices (Luo et al., 2014). As an example the often higher wastewater content of DOC/NOM influences the efficiency of O₃ or UV compared to surface water with generally lower DOC/NOM content.

### 8-9 Relevance of treatment technology

Experimental conditions need to be appropriate for the purpose of the tests. It should be considered how the tested conditions relate to full-scale operational ranges (e.g. dosing, exposure time) and whether they are realistic. For example, the maximum dosage of O₃ is in many cases legally restricted because of the formation of bromate that is highly toxic (von Gunten, 2003b).

#### 4.2.2 Assessing the reliability of a water treatment study

When assessing the reliability of a water treatment study for data evaluation, we used the definition by Roth and Ciffroy (2016): “The reliability of a study relates inter alia to the robustness and validity of the method used, the completeness and detail of reporting, the clarity and plausibility of the findings to ensure their reproducibility, but also to the uncertainty of the knowledge base”. With regard to reliability, we divide the criteria into the highlighted components from the above definition: method, reporting information, reproducibility and uncertainty (see Table 3).

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria for assessing the relevance of efficiency studies for advanced water treatment technologies.</td>
</tr>
<tr>
<td>Criterion</td>
</tr>
<tr>
<td>Relevance of data</td>
</tr>
<tr>
<td>Is the scope of the tests appropriate for the evaluation?</td>
</tr>
<tr>
<td>Are the data reported appropriate for the evaluation?</td>
</tr>
<tr>
<td>Relevance of tests</td>
</tr>
<tr>
<td>In case of a formulation or mixture is the compound tested representative and appropriate for the compounds being assessed?</td>
</tr>
<tr>
<td>Are the properties of the compounds chosen appropriate for the purpose of the tests?</td>
</tr>
<tr>
<td>Is the applied compound concentration appropriate for the purpose of the tests?</td>
</tr>
<tr>
<td>Relevance of water matrix</td>
</tr>
<tr>
<td>Is an appropriate type of water chosen for the purpose of the tests?</td>
</tr>
<tr>
<td>Are the properties of the water matrix chosen appropriate for the purpose of the tests?</td>
</tr>
<tr>
<td>Relevance of treatment technology</td>
</tr>
<tr>
<td>Are the properties of the treatment technology chosen appropriate for the purpose of the tests?</td>
</tr>
<tr>
<td>Is the scale of the experiment appropriate for the purpose of the tests?</td>
</tr>
</tbody>
</table>

#### 4.2.2.1 Explanation of the reliability criteria (criteria numbers from Table 3)

**Method**

1-2 Is a guideline or modified guideline used for any part of the experiment? Is the test performed under GLP conditions?

The International Organization for Standardization (ISO) and the Organisation for Economic Co-operation and Development (OECD) are continuously producing guidelines to amongst others standardize various chemical tests (ISO, 2017; OECD, 2017). At the moment no guidelines for testing removal efficiencies for the selected treatment techniques are available. Nevertheless, ISO and OECD have provided guidelines on how to measure CECs and how to take water samples. The American Society for Testing and Materials (ASTM) published several standards related to water treatment technologies on their website, amongst others standards on how to predict adsorption capacity of activated carbon (AC) and some standard procedures for membrane testing e.g. salt rejection (ASTM, 2014a; b; c). These guidelines are relevant for certain studies on CEC removal efficiency. Besides these guidelines, it is also relevant to know if the experiments are performed under good laboratory practices (GLP) conditions.

3 Are validity criteria fulfilled (controls)?

This to verify that only the tested treatment technology is removing the CECs from the water matrix and no other processes such as natural degradation or evaporation takes place.

**Reporting information**

To have all the information necessary to make a thorough evaluation of the limitations and exact conditions of the CEC removal efficiency study the following information (criteria 4-47) should be reported either in the study itself or in the supplementary information.

**Target compound**

4-6 Identification

Many of the below mentioned impact parameters can be found in databases and literature (e.g. Episuite, Chemspider, OECD toolbox etc.) if the compound name and/or CAS number is known.

7 Molecular charge/pKa (AC, Ozone, NF/RO membranes)

The pKa of the compound combined with the pH of the water matrix can be used to determine the charge of the compound if this is not given in the study (de Ridder et al., 2010; Moreno-Castilla, 2004). A charged molecule will be either attracted to or repelled by the charge of the AC surface (De Ridder et al., 2011; de Ridder et al., 2010; Kovalova et al., 2013; Mailler et al., 2014; Margot et al., 2013; Moreno-Castilla, 2004). For the same reason the charge of a compound also has an influence on the NF/RO rejection, usually negatively charged CECs are better rejected than positively charged CECs (Bellona et al., 2004; Verliefele et al., 2008; Yoon et al., 2007). Also for treatment with ozone the charge is often relevant, especially for nitrogen containing compounds as often only deprotonated compounds have enough electron density, see also 10 Functional groups/reactivity, to be attacked by ozone (Borowska et al., 2016).

8 $\log K_{ow}/\log D_{ov}$ (AC, NF/RO membranes)

The hydrophobicity of a compound affects the efficiency of a treatment technique (Hu et al., 1998; Kovalova et al., 2013; Mailler et al., 2014).
The hydrophobicity is often expressed as the octanol-water partitioning coefficient Log $K_{ow}$ or the pH corrected log $D_{ow}$ (De Ridder et al., 2011; Kovalova et al., 2013). Log $D_{ow}$ values can be calculated on the basis of the reported pH of the water matrix, the $pK_a$ and the log $K_{ow}$, so it is not essential to report. For NF/RO membranes more hydrophobic CECs can be adsorbed by the membrane and are initially well removed, but removal may decrease later due to saturation of the membrane.

---

### Table 3

Criteria for assessing the reliability of removal efficiency studies for advanced water treatment technologies, specified per treatment technology. When no technology is specified the criterion applies to all technologies.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>1</td>
</tr>
<tr>
<td>Is a guideline (e.g., OECD/ISO/ASTM) or modified guideline used for any part of the experiment?</td>
<td>1</td>
</tr>
<tr>
<td>Is the test performed under Good Laboratory Practices (GLP) conditions?</td>
<td>2</td>
</tr>
<tr>
<td>Are validity criteria fulfilled (e.g. controls)?</td>
<td>3</td>
</tr>
<tr>
<td>Reporting information</td>
<td>4</td>
</tr>
<tr>
<td><strong>Target compound</strong></td>
<td>5</td>
</tr>
<tr>
<td>Identification</td>
<td>4</td>
</tr>
<tr>
<td>Name</td>
<td>5</td>
</tr>
<tr>
<td>CAS number or other identifier</td>
<td>6</td>
</tr>
<tr>
<td>Form tested (e.g. salt, acid or base)</td>
<td>7</td>
</tr>
<tr>
<td>Impact parameters</td>
<td>8</td>
</tr>
<tr>
<td>Molecular charge/$pK_a$ (AC, ozone, NF/RO membranes)</td>
<td>9</td>
</tr>
<tr>
<td>Log $K_{ow}$/Log $D_{ow}$ (AC, NF/RO membranes)</td>
<td>10</td>
</tr>
<tr>
<td>Molecular weight/size (AC, NF/RO membranes)</td>
<td>11</td>
</tr>
<tr>
<td>Functional groups/reactivity</td>
<td>12</td>
</tr>
<tr>
<td>Supplier, purity of target compound</td>
<td>13</td>
</tr>
<tr>
<td>Water matrix</td>
<td>14</td>
</tr>
<tr>
<td>Matrix being wastewater, wastewater effluent, surface water, ground water, drinking water, demineralized water or synthetic wastewater</td>
<td>15</td>
</tr>
<tr>
<td>Impact parameters</td>
<td>16</td>
</tr>
<tr>
<td>pH</td>
<td>17</td>
</tr>
<tr>
<td>NOM concentration and possibly composition</td>
<td>18</td>
</tr>
<tr>
<td>Temperature</td>
<td>19</td>
</tr>
<tr>
<td>Nitrite/nitrate concentration (ozone, UV)</td>
<td>20</td>
</tr>
<tr>
<td>Bromide/bromate concentration (ozone)</td>
<td>21</td>
</tr>
<tr>
<td>Turbidity (NF/RO membranes)</td>
<td>22</td>
</tr>
<tr>
<td>UV transmittance (UV)</td>
<td>23</td>
</tr>
<tr>
<td>Treatment process conditions</td>
<td>24</td>
</tr>
<tr>
<td>Identification</td>
<td>25</td>
</tr>
<tr>
<td>Type of technology used</td>
<td>26</td>
</tr>
<tr>
<td>Impact parameters</td>
<td>27</td>
</tr>
<tr>
<td>Surface or membrane area (AC, NF/RO membranes)</td>
<td>28</td>
</tr>
<tr>
<td>Pore volume and pore size distribution (AC)</td>
<td>29</td>
</tr>
<tr>
<td>Bed volume (GAC)</td>
<td>30</td>
</tr>
<tr>
<td>Flow through (GAC)</td>
<td>31</td>
</tr>
<tr>
<td>Preloading (GAC)</td>
<td>32</td>
</tr>
<tr>
<td>Surface charge (AC, NF/RO membranes)</td>
<td>33</td>
</tr>
<tr>
<td>Concentration/dosage or intensity (PAC, ozone, UV)</td>
<td>34</td>
</tr>
<tr>
<td>Contact, exposure or irradiation time (AC, ozone, UV)</td>
<td>35</td>
</tr>
<tr>
<td>Biological activity (GAC)</td>
<td>36</td>
</tr>
<tr>
<td>Technology application (diagram)</td>
<td>37</td>
</tr>
<tr>
<td>Wavelength (UV)</td>
<td>38</td>
</tr>
<tr>
<td>Concentration of reagent $H_2O_2$ (ozone, UV)</td>
<td>39</td>
</tr>
<tr>
<td>Fouling (NF/RO membranes)</td>
<td>40</td>
</tr>
<tr>
<td>Trans membrane pressure (NF/RO membranes)</td>
<td>41</td>
</tr>
<tr>
<td>Permeate flux (NF/RO membranes)</td>
<td>42</td>
</tr>
<tr>
<td>Cross-flow velocity (NF membranes)</td>
<td>43</td>
</tr>
<tr>
<td>Backwashing (GAC)</td>
<td>44</td>
</tr>
<tr>
<td>Recovery (NF/RO membranes)</td>
<td>45</td>
</tr>
<tr>
<td>MWCO and/or salt rejection data (NF/RO membranes)</td>
<td>46</td>
</tr>
<tr>
<td>Experimental conditions</td>
<td>47</td>
</tr>
<tr>
<td>Scale of experiment (lab, pilot or full scale)</td>
<td>48</td>
</tr>
<tr>
<td>Stand alone or part of a treatment train</td>
<td>49</td>
</tr>
<tr>
<td>Target compound present in the water matrix tested, or spiked</td>
<td>50</td>
</tr>
<tr>
<td>Tested as single compound or in a mixture</td>
<td>51</td>
</tr>
<tr>
<td>Initial concentration of target compound</td>
<td>52</td>
</tr>
<tr>
<td>Concentrations after treatment of target compound and time points for measuring (duration of experiment).</td>
<td>53</td>
</tr>
<tr>
<td>How are the samples collected and with what method?</td>
<td>54</td>
</tr>
<tr>
<td>Analytical method: description of method, including recoveries and matrix effects and correction for these, Level of Quantification (LOQ)</td>
<td>55</td>
</tr>
<tr>
<td>Reproducibility and uncertainties</td>
<td>56</td>
</tr>
<tr>
<td>Calculation of removal efficiency</td>
<td>57</td>
</tr>
<tr>
<td>Use of statistical methods</td>
<td>58</td>
</tr>
<tr>
<td>Is the result/efficiency statistically significant?</td>
<td>59</td>
</tr>
<tr>
<td>Uncertainty data</td>
<td>60</td>
</tr>
</tbody>
</table>

---

* OECD: The Organisation for Economic Co-operation and Development.
Large compounds tend to be less easily removed with AC (De Ridder et al., 2011; Mailler et al., 2014). The size of a compound, expressed in molecular weight, is believed to be one of the main factors in rejection of CECs by NF/RO membranes (Bellona et al., 2004; Yoon et al., 2007). Besides the molecular weight also the molecular size, width, length, (effective) diameter and the molecular volume are used in studies as a measure of the size of the CEC (Braeken et al., 2005; Comerton et al. 2008, 2009; Jung et al., 2005; Kiso et al., 2001; Sadmani et al., 2014; Yangali-Quintanilla et al., 2009).

Relevant functional groups can be identified on the basis of the compound structure, so in principle they do not have to be reported although it is helpful if the author includes information on this topic. Different functional groups are relevant for different techniques. With regards to AC especially groups allowing H-bonds or π-π binding are relevant (de Ridder et al., 2010; Mailler et al., 2014). For NF/RO membranes functional groups of CECs may play a role in whether the compound is rejected or not, this is, however, mostly related to the resulting size and/or charge of the compound (Bellona et al., 2004; Kosutić and Kunst, 2002; Ozaki and Li, 2002). With regards to O₃ and UV, it is the reactivity of a compound that is important. This is related to the electron density of the compound, which can again be related to the functional groups. Functional groups with a high electron density will increase the reactivity of a compound (von Sonntag and von Gunten, 2012). For treatment with ozone a broad range of functional groups such as nitro groups, amides, primary amines, thioethers and olefins are known to enhance the removal of CECs (Kovalova et al., 2013; Nakada et al., 2007; Snyder et al., 2007a; von Gunten, 2003a). For UV or O₃ treatment in combination with H₂O₂ an even broader scale of chemical structures will be attacked as OH radicals react more unspecific with a very wide variety of functional groups (Snyder et al., 2007a).

Supplier and purity (All)
Supplier and purity of purchased target compounds should be reported, as this cannot be retrieved elsewhere. This is important with regards to both the analysis and the behaviour of the target compounds.

Water matrix
12 Type of water (All)
It is essential to know which type of water was used and whether it has been artificially created or is of environmental origin. This gives an indication of the relevance of the study in relation to full scale treatment and what issues to expect from the water matrix. If the water matrix is artificially created, details on the substituents and preparation procedure are to be reported. In addition, the origin of the composition may be given, as this can be based on literature e.g. greywater (Benami et al., 2016; Pradhan et al., 2019) and wastewater (Wilsenach and Van Loosdrecht, 2004).

13 pH (All)
As previously discussed under 7 Molecular charge/pKa, the pH of the water matrix is relevant. In most full scale treatments the pH lies between 7 and 9 (Kovalova et al., 2013; Mailler et al., 2014; Margot et al., 2013), so it should be explicated if the study has been performed at a significantly different pH. Depending on the pKa a CEC might express higher or lower reaction rates with UV or ozone at a specific pH due to whether it will be present in a deprotonated or protonated state, as mentioned in 7 Molecular charge/pKa (Avissar et al., 2010; Borowska et al., 2016; Real et al., 2009). Furthermore, at a high pH, ozone decomposition to OH radicals increases, and this likely will change the CEC removal (Real et al., 2009). The pH can also have an influence on the surface charge of both NF/RO membranes and AC, as well as impact fouling of the membrane (Bellona et al., 2004; Braghetta et al., 1997; Deshmukh and Childress, 2001; Hagmeyer and Gimbel, 1998).

14 NOM concentration (and composition) (All)
For all treatments the presence of NOM can reduce the efficiency of the treatment in which both the concentration and composition are relevant parameters (Boehler et al., 2012; De Ridder et al., 2011; Loos et al., 2013; Mailler et al., 2016; von Gunten, 2003a). With regards to AC NOM can reduce the removal efficiency of CECs by pore blocking or adsorption competition (Kidduff et al., 1998; Matsui et al., 2002; Newcombe and Drikas, 1997; Pelekani and Snoeyink, 1999). The composition of NOM has proven to be relevant for adsorption competition, where protein-like fluorophores seem to be the most problematic molecular components (Mailler et al., 2016). In advanced oxidation treatment the OH radicals are scavenged by NOM, leading to lower removal efficiencies (Neamtu and Frimmel, 2006; Pereira et al., 2007; Snyder et al., 2007a; von Gunten, 2003a; Yuan et al., 2009; Zwiener and Frimmel, 2000). In the case of UV treatment the presence of NOM can increase the removal efficiency of the treatment as the presence of NOM can initiate the production of OH radicals (Leech et al., 2009; Neamtu and Frimmel, 2006; Pereira et al., 2007; Zhan et al., 2006). On the other hand NOM can also decrease the efficiency of UV treatment as it can adsorb the UV, see 19 UV transmittance. As with UV the effect of NOM on NF/RO membrane removal efficiency is ambiguous, depending on the amount of NOM, the type of CECs, and the type and thickness of the fouling layer, and has been researched in many studies (Comerton et al., 2009; Nghiem and Coleman, 2008; Nghiem et al., 2010; Nghiem and Hawkes, 2007; Verliefde et al., 2009a). Usually NOM in wastewater is negatively charged, assisting in repulsion of negatively charged CECs (Mailler et al., 2014). Even though the effect of NOM present in the water matrix may not be clear, it is important to mention the presence, concentration, and when relevant and possible the composition of the NOM, so the influence can be taken into consideration when evaluating the removal efficiency. In the 49 studies evaluated in this paper NOM is expressed as DOC (18 studies), TOC (17 studies) and/or COD (13 studies). To enable comparison across studies it is recommended to include at least DOC. Only 4 studies mentions NOM composition, 2 PAC studies, one NF and one NF/RO study.

15 Temperature (All)
With regards to AC the temperature can influence adsorption kinetics, with a low temperature giving lower removal efficiency (Nam et al., 2014). In removal with ozone, temperature can influence ozone decay. In a study by Real et al. (2009) the rate constants of removal with UV increased with increasing temperature. However, the difference in minimum and maximum temperature in the study was 30 °C, and this difference is unlikely to occur in full-scale applications. With regards to NF/RO membranes a higher temperature will result in more viscous water, which will easier transport (diffuse) through the membrane and therefore result in a higher flux which also influences the CEC removal, see also 34-35 Transmembrane pressure and permeate flux (Crittenden and Montgomery Watson, 2012).

16-17 Nitrite/nitrate and bromide/bromate (Ozone, UV)
When applying ozone (especially for wastewater) the nitrite concentration is relevant since nitrite reacts quickly with ozone and reduces the available dose (Hollender et al., 2009; Lee et al., 2013). For both UV and ozone the nitrate concentration (especially for drinking water) is also relevant as this can form nitrite and other by-products which are highly toxic. Furthermore, bromide can form the toxic compound bromate with ozone and consequently, above certain bromide concentrations ozonation cannot be recommended (Lee et al., 2013; von Gunten, 2003b; von Sonntag and von...
compensate for the limiting effect of the DOC concentration.

The DOC concentration of the water matrix, as a higher dose can

The technology test.

report any development of these parameters over the course of

Some of these parameters will change over time, so it is important
to understand the ability of the UV radiation to reach the CECs. The
transmittance is an indication of how much of the UV radiation is
adsorbed by dissolved matter in the water matrix, and is related to
the NOM and nitrate content of the water matrix (Kruithof and
Martijn, 2013).

In general, the higher the transmittance the higher the CEC removal (Wols et al., 2015). A low transmittance can be compensated for by using more energy to reach the same UV
dose.

Treatment process conditions

The parameters listed below are all relevant to report as they are
unique to the particular study and cannot be found elsewhere. Some of these parameters will change over time, so it is important
to report any development of these parameters over the course of
the technology test.

20 Type of technology (All)
The exact type of water treatment technology used should be
reported.

21 Surface/membrane area (AC, NF/RO membranes)
The surface area has a significant influence on the removal effi-
ciciency of a specific AC, and is thus important to report. In general,
a smaller surface has a smaller removal efficiency (Mailler et al.,
2016). The surface area can be acquired from the producer, if the
commercial name is provided in the study. For NF/RO membranes
the membrane area tested is important to understand the scale of
the test rather than its influence on the removal capacity (Verliefde
et al., 2009b).

22 Pore volume and pore size distribution (AC)
The pore volume and the pore size distribution is relevant in
relation to the size of the CEC, and the composition of NOM
(Kovalova et al., 2013; Mailler et al., 2016; Moreno-Castilla, 2004).
NOM can access mesopores and block these and the access of the
CECs to underlying micropores (Moreno-Castilla, 2004).

23–25 Bedvolume, flow through and preloading (GAC)
The bedvolume and flow through of the GAC column influences
the removal efficiency of the GAC treatment (Chowdhury, 2013).
Bedvolume and flow through indicate the time the CECs have to
adsorb to the AC, see also 28 Contact time. In case the GAC has been
used before it will be preloaded with NOM, this should be reported
as it can have a negative influence on the efficiency, however
preloaded GAC columns might also exhibit biological degradation
which will have a positive effect on the removal efficiency (Bourgin
et al., 2018).

26 Surface charge (AC, NF/RO membranes)
The impact of the surface charge has been discussed under 7
Molecular charge/pKa. For NF/RO membranes the surface charge
is often quantified as zeta potential (Bellona et al., 2004). NF/RO
membranes are mostly negatively charged at neutral pH, and thus
have a higher affinity for positively, or neutral compounds (Bellona
et al., 2004; Taheran et al., 2016).

27 Concentration/dosage/intensity (PAC, ozone, UV)
For PAC, the dosage is the main controlling factor for the
removal efficiency, the higher the dose the higher the removal
(Boehler et al., 2012; Mailler et al., 2014; Snyder et al., 2007b). Also
for ozone, the dosage has a substantial positive influence on the
removal efficiency (Margot et al., 2013; Snyder et al., 2007a; Zwiener and Frimmel, 2000). Often the ozone dosage is related to the
DOC concentration of the water matrix, as a higher dose can
compensate for the limiting effect of the DOC concentration
(Zwiener and Frimmel, 2000). Especially for wastewater the ozone
dosage is often given in relation to the DOC concentration as g O3/g
DOC (Hollender et al., 2009; Kovalova et al., 2013). Also for UV
in general the higher the dose the higher the removal efficiency,
however, this is only valid for CECs that are already susceptible to
degradation with UV light (Kovalova et al., 2013; Wols and Hofman-
Caris, 2012; Wols et al., 2013). In some of the studies evaluated the
UV intensity is given in einstein s−1, which is not a SI unit and
cannot be compared to the more commonly used mJ/cm².

28 Contact, exposure or irradiation time (AC, ozone, UV)
The contact time can have a great influence on the removal effi-
ciciency of AC (Mailler et al., 2016; Snyder et al., 2007b). However
for compounds with very fast adsorption kinetics this impact factor
will be less relevant (Mailler et al., 2016). In UV studies the irradia-
tion time is included in the calculation of the dose, so it is less
relevant to state. However, when using advanced oxidation, the
exposure time, during which the OH radicals have time to form and
react with the CECs, is relevant.

29 Biological activity (GAC)
Usually it is known if there is a possibility of biological activity in
the GAC experiments. If this is the case it needs to be stated as the
presence of microorganisms in the GAC filter can enhance removal
of biodegradable CECs (Bourgin et al., 2018; Magic-Knez and van
der Kooij, 2004; Rattier et al., 2014). Biological activity depends
on the time the GAC have been in use. When using fresh GAC or GAC
directly after regeneration there will be no biological activity in
the beginning, but it will develop and increase in time during use.

30 Technology application (All)
The way the technology is applied is also relevant to report as for
example the efficiency of ozone depends a lot on the mixing regime
(Dodd et al., 2008; Zucker et al., 2016). It can be difficult to describe
an experimental set-up, so in this case the criterion is considered
fulfilled if a schematic of the set-up has been included.

31 Wavelenght (UV)
The wavelength or the source of the UV light (which type of
lamp) is relevant, especially in the case of treatment without H2O2
(Yang et al., 2014). As mentioned under 10 Functional groups/react-
tivity, this can have influence on whether the CEC will be able to
absorb the radiation, and thus degraded by photolysis (Chen et al.,
2007; Rosenfeldt and Linden, 2004; Yang et al., 2014). Furthermore
a broader UV spectrum can lead to more unwanted by-products
(Hofman-Caris et al., 2015).

32 Concentration of reagent (H2O2)(Ozone, UV)
The addition of H2O2 to ozone or UV treatment increases the
production of OH radicals, and thus the removal efficiency of the
treatment (Ocampo-Pérez et al., 2010; Real et al., 2009; Ternes
et al., 2003; Zwiener and Frimmel, 2000). An increase of the H2O2
dose can be used to compensate for NOM scavenging (Zwiener and
Frimmel, 2000).

33 Fouling (NF/RO membranes)
Fouling and the type of fouling is important for the evaluation of
the removal efficiency of the NF/RO membrane. Organic fouling/biofouling has been referred to under 14 NOM concentration (and
composition). Fouling with NOM can lead to higher removal of
negatively charged CECs and lower removal of positively charged
CECs. Pressure and flux is also influenced by fouling (added resis-
tance) as it takes more pressure to force water through a fouled
membrane while keeping the flux stable, see also 34–35 Trans-
membrane pressure and permeate flux (Taheran et al., 2016). 34–35 Transmembrane pressure and permeate flux (NF/RO membranes)
The transmembrane pressure (TMP) is related to the permeate
flux in a NF/RO membrane, the higher the transmembrane pressure
the higher the flux, when the resistance is constant (Plakas and
Karabelas, 2012). Increased pressure lead to higher removal of

Gunten, 2012).

18 Turbidity (NF/RO membranes)
For NF/RO membranes the turbidity can give an indication of the
pre-treatment needed and whether to expect fouling, see 33 Fouling
(Kim et al., 2002).

19 UV transmittance (UV)
The UV transmittance of the water matrix is relevant to be able
understand the ability of the UV radiation to reach the CECs. The
transmittance is an indication of how much of the UV radiation is
adsorbed by dissolved matter in the water matrix, and is related to
the NOM and nitrate content of the water matrix (Kruithof and
Martijn, 2013). In general, the higher the transmittance the higher
the CEC removal (Wols et al., 2015). A low transmittance can be
compensated for by using more energy to reach the same UV
dose.
CECs (Plakas and Karabelas, 2012). The highest removal efficiency of pesticides was found with NF membranes operation with a high flux by Chen et al. (2004) and Ahmad et al. (2008). In the 49 evaluated papers TMP was not always given. Instead the working pressure, the applied pressure or just pressure was given, obstructing comparison across studies.

36 Cross-flow velocity (NF membranes)

During membrane filtration the phenomenon of concentration polarization occurs, resulting in a higher concentration of solutes in front of the membrane surface, resulting in a lower rejection of solutes compared to the bulk solution (Hajibabania et al., 2011a; Ng and Elimelech, 2004). A higher cross-flow velocity results in a lower concentration polarization, and therefore a higher CEC rejection (Crittenden and Montgomery Watson, 2012). The cross flow velocity cleans the membrane, the higher the velocity the more fouling (resistance) is removed (Crittenden and Montgomery Watson, 2012). The influence of fouling on removal of CECs has been discussed under 33 Fouling.

37 Backwashing (GAC)

Backwashing can influence the CEC removal of the GAC. It removes materials blocking the pores of the GAC enhancing the removal efficiency on the other hand it can also disturb the build-up of the bed and remove organic material degrading the CECs which will decrease the removal efficiency, see also 29 Biological activity and 14 NOM concentration (and composition). It is therefore important to state if backwashing is done and how it has been done.

38 Recovery (NF/RO membranes)

The recovery of a NF/RO membrane has an effect on the removal efficiency (Bellona et al., 2004). Several studies have found that removal efficiency of CECs decreases with increasing recovery (Chellam and Taylor, 2001; Chen et al., 2004).

39 MWCO and/or salt rejection data (NF/RO membranes)

The rejection capacity of NF/RO membranes can be characterized by the salt passage or rejection of standard salts. For RO membranes typically monovalent salts, such as NaCl are used, while for NF membranes bivalent salts such as MgSO₄ are used (Bellona et al., 2004; Kiso et al., 1992, 1996, 2001). It is preferable that these characteristics are noted in the original study as there might be minor changes from batch to batch, producers and production processes change. The molecular weight cut-off (MWCO) is also used in NF/RO, referring to the lowest molecular weight solute (in daltons) in which 90% of the solute is retained by the membrane. The MWCO, however, can be difficult to obtain, and is due to differences in protocols used by manufacturers not always comparable between NF/RO membranes (Bellona et al., 2004; Cleveland et al., 2002).

Experimental conditions

All experimental conditions are essential to report as they are all unique to the specific treatment study, and can influence the removal efficiency of the studied technique.

40 Scale of experiment

As differences can be expected between lab- pilot- and full scale experiments, it is important to know the scale of the study (Hofman-Caris et al., 2017; Verliefde et al., 2009b). Lab scale refers to laboratory scale experiments (proof of principle). This can be either batch experiments (batch volume < 100 L) or continuous flow experiments (flow < 100 L/h). Pilot scale (proof of practice) refers to experiments on a larger scale, often on-site, with a flow rate normally < 10 m³/h. These scales proceed full scale application (proof of market).

41 Stand-alone or part of a treatment train

When the study is done on a treatment train it is important to know the impact parameters of the water matrix, and the removal efficiency of the CECs at each step of the treatment train, at least if the authors want to conclude anything on the individual treatment steps. In all studies it is relevant to know whether the tested water has received any form of pre-treatment in advance of the treatment technology investigated in order to have an understanding of what to expect in the tested water matrix, see also 12 Type of water. The impact parameters of the water matrix should be measured after the pre-treatment. In case any post treatment is needed it can be relevant to mention this, depending on the purpose of the study.

42 CEC already present or spiked

Depending on the purpose of the study spiking can be necessary to ensure that an accurate removal percentage can be calculated. When spiking a factor 100 above the limit of quantification, a removal percentage of 99% can be determined. However, very high spiking at environmentally unrealistic concentrations may have an effect on removal efficiencies and may alter the behaviour of the CEC, which is undesirable.

43 Single compound or mixture

This gives an indication of the boundaries of the study, as in reality the compounds will almost always be present in a mixture.

44-45 Initial and end concentration of CECs

The initial concentration and the end concentration is particularly important in case no explanation is given on how removal efficiency percentage have been calculated.

46 Sample collection and pre-treatment

It is relevant to report how the samples have been collected, e.g. grab sampling or continuously sampling. Also the number of samples and over what time period is important to mention, in order to get an understanding of the representativeness of the samples.

47 Analytical method CECs

Here it is essential to describe the analytical methods used and the treatment of the samples. Furthermore, it is important to give the limit of detection for each CEC specifically, and if possible the recoveries too. This information is very relevant as it gives a detailed understanding of what the removal efficiencies are based on. This can all be added in supplementary information.

Reproducibility and uncertainties

To ensure the clarity and plausibility of the experimental findings, the following information (criteria 48-51) should be available to assess the reproducibility and uncertainties of the study.

48 Calculation of removal efficiency

It should be clear how the removal efficiencies are calculated, and how values <LOQ/LOD are treated, as these choices highly impact the results (Helsel, 2005; Weltje and Sumpter, 2017). When spiking the CECs this problem can be avoided, by spiking a high enough concentration. When not spiking, When not spiking a ‘>' percentage using the LOQ/LOD as end concentration is preferred to stating a 100% removal, as this in cases with a low initial concentration and a high LOQ/LOD can be very misleading. It is also important to note the number of replicates, and how they are used in the determination of the removal efficiencies.

49 Statistical methods

Have appropriate statistical techniques been employed to evaluate variability and uncertainty? If any statistics are used to analyse the experimental findings this should be mentioned including which methods and why.

50 Significance

When significance of the experimental findings is presented it should be noted how this is determined and at what p-value a result is deemed significant.

51 Uncertainties data

A description of the produced amount of data and to what degree it is sufficient to support the conclusions of the study is very helpful to evaluate the reliability of the results. To what extent do the uncertainty and variability of the different measurements impact the conclusions that can be inferred from the data and the
utility of the study? What is the standard deviation of the results?

5. Application of the criteria

The current status of reporting is discussed based on the review of 244 selected papers (Supplementary Information 1), and the benefits of a standardised framework are elaborated based on the findings. Because of the broad selection of techniques and resulting large amount of data, the evaluation of the CEC removal itself will be presented in a follow-up study.

5.1. Relevance criteria

Table 4 details the results of the application of the relevance criteria described in chapter 4.2.1 on 244 peer-reviewed scientific articles on the removal of CECs with one of the selected water treatment techniques.

It is important to realise that the failure to fulfil one of the relevance criteria does not necessarily mean that the paper is of poor quality, but merely that the paper does not comply with the specified relevance criteria and cannot be used for our purpose, i.e. decision-making by stakeholders comparing removal efficiencies between various water treatment techniques. It is however remarkable that so few of the evaluated papers are suitable for a comparison across techniques, compounds and water matrices. In general the first three criteria are the ones most frequently not fulfilled. The papers were either outside the scope we were interested in, did not include removal percentages at all or did not report removal percentages in a way that could be used for our purpose.

Overall, only 20% of the papers evaluated were in compliance with the relevance criteria and could be used for our purpose.

5.2. Reliability criteria

The application of the relevance criteria left 49 studies on which the reliability criteria could be applied. Table 5 details the results of the application of the 51 reliability criteria to the water treatment technology papers.

From Table 5, it is clear that some criteria (in bold) such as compound name, technology/combinations, wavelength, scale of experiment, whether it is a stand-alone, whether the compound is spiked or not, and the pressure for NF/RO membranes are always reported. For compound name, technology/combinations and scale of the experiment this is logical as this is also included in the relevance criteria, and the studies have been selected based on this information.

Several criteria are almost always reported (criteria 9, 12, 13, 14, 21, 27, 30, 32, 36, 39, 43, 44 and 47) however for some of these parameters such as NOM content and dosage of O3 and UV, what is reported is not consistent. NOM was found to be reported in DOC, TOC and/or COD, and O3 as mg/L or mg O3/g DOC which cannot easily be compared if DOC is not given.

The criteria for the target compound properties (criteria 4–10) are poorly reported, varying from not at all to 50% of the studies, except for molecular weight which is reported in 21 of 24 studies. In general these criteria are not essential to report as they can be found elsewhere. However, they facilitate the interpretation of the results in the study itself.

The supplier of the CECs tested (criterion 11) is in general reported, however the purity of the compound (also criterion 11) only in approximately 60% of the studies.

The criteria concerning the water matrix (criteria 12–19) except water type, pH and NOM concentration (and composition) are also not well reported, with the most reported in 50% of the studies and the least in 25%. This is essential information specific for the study, and cannot be found elsewhere.

For the treatment process conditions (criteria 20–39) not yet mentioned, reporting is also not complete. In general, they are reported in less than 40% of the studies, this information is also specific for the study and cannot be found elsewhere.

With regards to the experimental conditions (criteria 42–48), sample collection and concentration after treatment are rarely reported. Especially the lack of end concentration makes it very difficult to understand and reproduce the calculation of removal efficiencies.

Understanding removal efficiencies is even more difficult as LOD/LOQ and recoveries for the analytical methods are not always reported (criterion 47), especially when the method used to calculate the efficiencies, the number of replicas, the standard deviation and how values below the LOD/LOQ are dealt with are reported in as little as 15%–47% of the studies. This makes it almost impossible to validate the findings of the studies. There is also virtually no reporting on whether statistics have been used in the studies (criterion 49) and how. This together with details on sample collection (criterion 46) are the most underreported areas found in the studies.

No papers fulfilled all the criteria. The GAC papers fulfill 36–50% of the criteria, PAC 41–73%, ozone (±H2O2) 38–74%, UV (±H2O2) 33–64%, NF/RO 39–71%. This does not mean that the papers cannot be used, but that not all the information needed for our evaluation is available. This will make the assessment of the efficiency of the technique less straightforward as impact parameters that can

<table>
<thead>
<tr>
<th>Criteria applied</th>
<th>Description</th>
<th>GAC</th>
<th>PAC</th>
<th>O3 ± H2O2</th>
<th>UV ± H2O2</th>
<th>NF</th>
<th>RO</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of papers evaluated</td>
<td></td>
<td>32</td>
<td>40</td>
<td>64</td>
<td>82</td>
<td>62</td>
<td>29</td>
<td>244</td>
</tr>
<tr>
<td>1</td>
<td>Outside scope</td>
<td>1</td>
<td>10</td>
<td>8</td>
<td>14</td>
<td>12</td>
<td>4</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>Without removal percentage</td>
<td>25</td>
<td>17</td>
<td>22</td>
<td>20</td>
<td>5</td>
<td>2</td>
<td>75</td>
</tr>
<tr>
<td>3</td>
<td>Data unusable</td>
<td>3</td>
<td>6</td>
<td>18</td>
<td>29</td>
<td>30</td>
<td>16</td>
<td>79</td>
</tr>
<tr>
<td>4</td>
<td>Not appropriate compound tested</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>No properties of compounds</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>No concentration of compound</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>No type of water matrix</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>No properties of water matrix</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>No properties of treatment technology</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>No scale of experiment</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Papers fulfilling all the criteria | 3 | 7 | 15 | 18 | 13 | 6 | 49 |

Data presented as > or < values out of the chosen range, only graphs no tables.
influence the removal efficiency are not reported and/or removal efficiencies have not been described.

6. Conclusions and recommendations

Our literature survey revealed that there is currently no uniform approach to study the CEC removal efficiency of advanced water treatment technologies. Removal is tested at various scales, with different water matrices and numerous CECs. There are large variations in the process conditions of the treatment, in the experimental set-ups and in the way results are expressed and reported. These variations hinder the interpretation and evaluation of the
removal efficiency data. This in turn makes it difficult for stakeholders to make an informed decision with regards to which treatment technology will be relevant for their specific needs. Therefore, in this study a framework for evaluation of scientific and technical information that describes what is important when evaluating removal efficiency studies was developed.

For this framework, two criteria sets were developed: 9 relevance criteria and 51 reliability criteria. These two criteria sets offer a thorough, unbiased and standardised method to select studies to evaluate and compare the CEC removal efficiency of advanced water treatment technologies in a scientifically sound way.

The relevance criteria have been applied to 244 treatment technology studies, and 49 of these papers fulfilled the criteria, with non-compliance with the criteria outside scope and appropriate data being the main reason for the papers being discarded. Especially the lack of removal percentage, and data reported in a way that it cannot be used, affect the potential of these studies to provide additional information and be suitable for a comparison across techniques, compounds and water matrices.

The reliability criteria were applied to the 49 remaining papers. The main finding here was a severe lack of information on how removal percentages had been calculated in terms of how LOD/C19
could have been applied, and what the standard deviation was. This hinders the interpretation of the reported result. Furthermore, many papers did not report on all the identified impact parameters which makes comparison across studies difficult.

These findings clearly demonstrate the need for a more uniform approach. When the developed framework is used as a guideline by scientists to document their studies it will facilitate the comparison of different treatment technologies regarding the removal of CECs. This in turn will enhance the interpretation of the findings of the studies and consequently benefit the selection of appropriate technologies by water managers and other stakeholders.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the experts interviewed for the identification of impact parameters, and the participants of the criteria workshop at the SOLUTIONS general assembly. We would also like to thank John Munteh and Christa McArdell for their input to this paper. We acknowledge funding for the project TAPES from the European Regional Development Fund through INTERREG IV B, and the European Commission for funding the FP7 project SOLUTIONS, contract number 603437.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.watres.2019.05.088.

References
