Identifying organic chemicals not subject to bioaccumulation in air-breathing organisms using predicted partitioning and biotransformation properties

Frank Wania, Ying Duan Lei, Sivani Baskaran, and Alessandro Sangion
Department of Physical and Environmental Sciences, University of Toronto Scarborough, Toronto, Ontario, Canada

Abstract
Because the respiration processes contributing to the elimination of organic chemicals deviate between air- and water-breathing organisms, existing and widely used procedures for identifying chemicals not subject to bioaccumulation in aquatic organisms based on the octanol–water partition ratio $K_{OW}$ need to be complemented with similar procedures for organisms respiring air. Here, we propose such a procedure that relies on the comparison of a compound’s predicted $K_{OW}$, octanol–air partition ratio $K_{OA}$, and biotransformation half-life $HL_B$ with three threshold values, below which elimination is judged to be sufficiently rapid to prevent bioaccumulation. The method allows for the consideration of the effect of dissociation on the efficiency of urinary and respiratory elimination. Explicit application of different types of the prediction error, such as the 95% prediction interval or the standard error, allows for variable tolerance for false-negative decisions, that is, the potential to judge a chemical as not bioaccumulative even though it is. A test with a set of more than 1000 diverse organic chemicals confirms the applicability of the prediction methods for a wide range of compounds and the procedure’s ability to categorize approximately four-fifth of compounds as being of no bioaccumulation concern, suggesting its usefulness to screen large numbers of commercial chemicals to identify those worthy of further scrutiny. The test also demonstrates that a screening based solely on $K_{OW}$ and $K_{OA}$ would be far less effective because the fraction of chemicals that can be judged as sufficiently volatile and/or sufficiently water soluble for rapid respiratory and urinary elimination based on the partitioning properties predicted for their neutral form is relatively small. Future improvements of the proposed procedure depend largely on the development of prediction methods for the biotransformation kinetics in air-breathing organisms and for the potential for renal reabsorption. *Integrated Environmental Assessment and Management* 2022;18:1297–1312. © 2021 The Authors. *Integrated Environmental Assessment and Management* published by Wiley Periodicals LLC on behalf of Society of Environmental Toxicology & Chemistry (SETAC).

KEYWORDS: Bioaccumulation, Biotransformation, Chemical assessment, Partitioning

INTRODUCTION
The process that causes a chemical’s concentration in an organism to exceed the concentration in the medium that this organism respires or in its diet, or both, has been termed bioaccumulation (B) (Gobas et al., 2009). While bioaccumulation is not a prerequisite for toxicity, a bioaccumulating substance is more likely to achieve levels within an organism that result in adverse effects than a substance that does not bioaccumulate. Because adverse effects of chemicals on organisms and ecosystems are often difficult to establish, bioaccumulation itself is regarded as an important element of chemical risk assessment (Beek et al., 2000; Franke et al., 1994). In fact, it is generally regarded as one of the most troublesome characteristics of a compound and should be avoided in commercially produced and used substances. Consequently, it is important to screen such substances to identify those that are potentially bioaccumulative.

Traditionally, chemical regulation sought to identify compounds that are bioaccumulative in aquatic organisms. The first step in that identification process is to confirm that the partitioning properties of a chemical allow for B. The relationship between chemical partitioning properties, such as the octanol–water partitioning ratio $K_{OW}$ and bioaccumulation in aquatic organisms, has been known for a long time (Chiou et al., 1977; Mackay, 1982) and is a result of lipophilic compounds not being readily eliminated.
through the gills of a fish. Bioaccumulation is therefore often judged by comparing a chemical’s \( \log K_{\text{OW}} \) against a threshold value.

A limitation of this approach is that there are compounds that are bioaccumulative in air-breathing animals, but not in aquatic organisms (Czub & McLachlan, 2004; Kelly et al., 2003, 2007). Such chemicals can be eliminated efficiently by respiration of water, which is an elimination process not available to organisms respiring air. Therefore, a B assessment based on \( K_{\text{OW}} \) alone is not protective of air-breathing animals. Conversely, there may also be chemicals readily eliminated by respiration of air that accumulate in water-breathing organisms. This means that separate B assessments for air- and water-breathing organisms are necessary.

One of the difficulties associated with decisions made based on a comparison of a chemical property with a threshold value is that every chemical property value has an uncertainty, irrespective of whether it is derived experimentally or through a prediction method. Zhang et al. (2010) have, for example, shown that such decisions can depend on which property prediction method is applied. Ideally, therefore, a method that relies on the comparison of chemical property values with thresholds should (i) quantify the uncertainty of the property value and (ii) estimate the likelihood of making a wrong decision because of that uncertainty.

In summary, it is important to have a procedure allowing for the identification of chemicals that may bioaccumulate in air-breathing organisms that complements existing methods for B assessment in fish. Here, we propose and describe such a method, whereby the method’s development was guided by the following considerations:

- Taking its cue from the B assessments relying on a log \( K_{\text{OW}} \) threshold, the method should ideally be simple and be based, to the extent possible, on chemical equilibrium partitioning properties. Because a method based on partitioning alone proved quite inefficient, we also show how elimination by biotransformation can be integrated into the assessment process.
- The assessment procedure should be efficient in that it requires little time and financial resources. For a procedure to be efficient, experiments to measure property values should only be necessary in exceptional cases. The method should therefore rely on predicted partitioning and biotransformation properties, so that it could be used both for (i) the relatively rapid screening of fairly large numbers of chemicals, and therefore, the focusing of more refined and detailed assessments on chemicals that are initially judged to be potentially bioaccumulative, and (ii) the quick assessment of individual chemicals, with the objective being to decide whether experimental determinations are required to appraise their potential for B.
- The method should explicitly take into account the uncertainty of predicted properties. Consideration should be given to the confidence with which these chemical properties are known and therefore the statistical likelihood that a value is indeed on one side of a threshold.

We also tested the efficiency of the proposed procedure by applying it to a set of more than 1000 organic compounds.

**METHODS**

**Outline of the assessment procedure**

The purpose of the procedure is to identify chemicals that are clearly not bioaccumulative in air-breathing organisms, and therefore, no further assessment of their potential for bioaccumulation should be necessary. Important, this is not synonymous with classifying chemicals as bioaccumulative. It is thus a procedure for de-prioritization rather than prioritization. The basic premise of the proposed method is that a chemical is not bioaccumulative in air-breathing organisms if the organism is capable of rapidly eliminating the compound. The log \( K_{\text{OW}} \) threshold used in the assessment of B in aquatic organisms could be seen as being based on the same premise, as chemicals with low log \( K_{\text{OW}} \) are more readily eliminated by respiration across the gills of a fish. There are primarily three pathways for air-breathing animals to eliminate a chemical: urinary excretion, respiratory exhalation, and enzymatic biotransformation. While it is also possible for compounds to be excreted with feces (Rozman, 1985), this process is too slow to prevent the bioaccumulation of hydrophobic substances. This is because the efficient assimilation of lipids during digestion assures that dietary intake rates always markedly exceed fecal elimination rates for bioaccumulative substances. While there are also substances that do not bioaccumulate by virtue of inefficient uptake rather than efficient elimination, we do not account for this possibility here because of the very large uncertainty of both the relevant threshold and the high log \( K_{\text{OW}} \) values of such substances. Furthermore, screening approaches for bioaccumulation in aquatic organisms generally neither include an upper log \( K_{\text{OW}} \) threshold indicative of inefficient dietary uptake.

A chemical’s propensity for urinary excretion and respiratory exhalation can be judged based on its equilibrium partitioning properties. In particular, ECHA (2017a, 2017b) recommends that chemicals with a log \( K_{\text{OW}} > 2 \) and a logarithmic octanol–air equilibrium partitioning ratio log \( K_{\text{OA}} > 5 \) are considered potentially bioaccumulative in air-breathing organisms. These thresholds are based on model calculations for terrestrial food webs that have suggested that any chemical with a log \( K_{\text{OW}} \) below 2 can be readily eliminated by urination, and any chemical with a log \( K_{\text{OA}} \) below 5 can be readily eliminated by exhalation (Kelly & Gobas, 2003; Kelly et al., 2003).

Bioaccumulation in humans has been shown to be strongly determined by a compound’s susceptibility to enzymatic biotransformation (McLachlan et al., 2011) and this is likely also true for other air-breathing organisms. The potential for rapid elimination by biotransformation cannot
be assessed with partitioning properties, but requires a parameter describing the kinetics of biotransformation in the organism. While this is generally not done in a tier 1 assessment of B in aquatic organisms, there has been considerable progress in predicting biotransformation kinetics of organic chemicals (Arnot et al., 2014; Brown et al., 2012; Papa et al., 2018). Furthermore, models can be used to estimate biotransformation half-life $H_L$ thresholds that assure rapid elimination, whereby the numerical thresholds depend on the size of the organism (Goss et al., 2013).

The proposed procedure therefore relies on the comparison of the compounds’ log $K_{OW}$, log $K_{OA}$, and $H_L$ with three thresholds. If the properties for a chemical are above all three of them, it is judged to be potentially bioaccumulative and therefore in need of further, more detailed assessment. If one of the chemical properties can, with confidence, be said to fall below any one threshold, it is deemed not to be susceptible to bioaccumulation in air-breathing organisms. While it is conceivable that any one elimination process is insufficient to prevent bioaccumulation of a chemical, whereas the combination of two or more of these processes would be sufficient, we ignore this possibility. This is defensible as it only increases the chance of false-positive decisions (i.e., chemicals being recommended for higher-tiered assessments even though they are not bioaccumulative), but not false negatives (i.e., chemicals are judged not B, even though they are).

One complication arises for compounds that dissociate, that is, are not completely neutral at the physiological pH prevailing in the body of the air-breathing organism. The partitioning of a dissociating compound between octanol and water is described with the distribution ratio, $D$, which is the ratio of the sum of the concentrations of the charged and neutral forms of a compound in octanol and water at equilibrium. $D$ depends on the pH of the aqueous phase. A chemical that is predominantly present in the charged form is more likely to be eliminated by urinary excretion, but less likely eliminated by respiration than its neutral conjugate. Our approach therefore includes two steps that consider the potential for dissociation. If dissociation is ignored, it might result in some substances being assessed as requiring further assessment of their bioaccumulative potential, even though they could be readily eliminated via urination in their charged forms. For chemicals whose neutral form is judged to have a log $K_{OW}$ above the threshold for efficient urinary excretion, we therefore also compare the log $D$ at the physiological pH of 7.4 with this threshold. Even more troublesome is that ignoring dissociation might lead to some substances being deemed not B by having a predicted log $K_{OA}$ below the threshold, even though in their charged form they are not readily exhaled. Therefore, a chemical with a log $K_{OA}$ for the neutral species judged to be below the threshold is not assessed to be readily lost by exhalation if it is predicted to be entirely present in charged form at a pH of 7.4, that is, if the neutral fraction in blood is estimated to be less than 1%.

**Detailed description of procedure**

Comparing predicted values against threshold values. The core elements of the approach are comparisons of a predicted property of chemical $i$ (X) with a threshold ($T$) to establish whether the property is above or below $T$. Because it is preferred to have false positives than to have false negatives, all the comparisons take the form:

\[
\text{If } (X_i + PE_i) < T, \text{ then chemical is not B}, \quad (1)
\]

where $PE_i$ is the prediction error of $X_i$. The choice of the parameter to use as $PE_i$ in these comparisons depends on the tolerance for different types of errors in an assessment. A very stringent approach would use the 95% prediction interval ($PI_{95}$) for $PE_i$:

\[
\text{If } (X_i + \text{ upper } PI_{95}) < T, \text{ then chemical is not B}. \quad (1')
\]

Using such an approach, one minimizes the potential of a false-negative assessment, that is, reduces the chance of “missing” a potentially bioaccumulative substance. There is a 95% likelihood that the true value of a property is between $X_i$ minus the lower $PI_{95}$, and $X_i$ plus the upper $PI_{95}$. There is thus a 2.5% likelihood that the true value is higher than $X_i$, plus the upper $PI_{95}$. For every comparison of type (1), there is thus at most a 1 in 40 chance of categorizing a chemical wrongly as not B, and only if $X_i$ plus upper $PI_{95}$ equals $T$ (Figure 1). In most cases, when this categorization of not B is made, $X_i$ plus upper $PI_{95}$ is smaller (often very much smaller) than $T$ and the likelihood of misclassification is in fact much smaller than 1 in 40. This approach therefore is very cautious.

If one were willing to tolerate a higher number of false negatives, for example, if the threshold value $T$ itself is

**FIGURE 1** Illustration of the use of different quantities to estimate the error of a predicted property $X_i$ and of the maximum likelihood of making a false-negative decision when comparing $X_i$ with a threshold $T$.
already cautiously selected, one could use the standard error of the prediction (SE) as the PE:

\[
\text{If } (X_i + \text{upper SE}) < T, \text{ then chemical is not B. (1')}
\]

This choice leads to fewer false-positive assessments, that is, reduces the fraction of chemicals requiring a higher-tier assessment. However, for every comparison of type (1'), there is at most a 1 in 6 chance of categorizing a chemical wrongly as not B, and only if \(X_i + \text{SE} \) equals \(T\). It is of course also possible, and presumably quite common, to ignore the uncertainty of a prediction, that is, use the comparison:

\[
\text{If } X_i < T, \text{ then chemical is not B. (1'')}\]

For every comparison of type (1' ''), there is at most a 1 in 2 chance of categorizing a chemical wrongly as not B, and only if \(X_i \) equals \(T\).

Specifically, the proposed approach relies on the following four comparisons:

\[
\text{If } (\log K_{OW} + \text{PE}_{\log K_{OW}}) < T_{\log K_{OW}}, \text{ then chemical is not B. (2)}
\]

\[
\text{If } (\log K_{OA} + \text{PE}_{\log K_{OA}}) < T_{\log K_{OA}} \\
\text{and neutral fraction at pH 7.4 > 0.01, then chemical is not B. (3)}
\]

\[
\text{If } (\log D_i \text{ at pH 7.4} + \text{PE}_{\log D_i}) < T_{\log K_{OW}}, \\
\text{then chemical is not B. (4)}
\]

\[
\text{If } (\text{HLB} \cdot \text{PE}_{\text{HLB}}) < T_{\text{HLB}}, \text{ then chemical is not B. (5)}
\]

where \(T_{\log K_{OW}}, T_{\log K_{OA}}, \text{ and } T_{\text{HLB}}\) are suitably selected thresholds for \(\log K_{OW}\), \(\log K_{OA}\), and biotransformation half-life, respectively. Whereas the upper prediction error is multiplied by the property value in Equation (5), in Equations (2)–(4) it is added to the property value. This is because the error is always assumed to be log-normally distributed.

In principle, the sequence of these comparisons should not matter. However, if any one of the comparisons leads to a decision of not B for a chemical, the subsequent comparisons become obsolete for that chemical. It is therefore desirable to start with comparisons involving properties that can be predicted with techniques that (i) yield predictions for a wide range of diverse compounds and (ii) are accessible easily and for free. Therefore, we start with the comparisons involving \(K_{OW}\) and \(K_{OA}\). Note that the approach can still be applied if a prediction technique for \(\log D_i\) or \(\text{HLB}\) is not accessible or if it does not yield a result for a compound. In that case, any comparison involving \(\log D\) and \(\text{HLB}\) can be omitted from the procedure, which is equivalent to assuming that the chemical exceeds the threshold for that property. It simply means that the likelihood of a judgment of “potential B,” that is, the number of false positives, increases.

To perform comparisons (2)–(5), we therefore need numerical values for the three thresholds \(T_{\log K_{OW}}, T_{\log K_{OA}}, \text{ and } T_{\text{HLB}}\) as well as, for each assessed chemical, values of \(\log K_{OW}\), \(\log K_{OA}\), \(\log D\), and \(\text{HLB}\), and, importantly, an estimate of the error for those predicted values. The following sections describe how we propose these values to be obtained.

**The thresholds for \(\log K_{OA}\), \(\log K_{OW}\), and the biotransformation half-lives \(\text{HLB}\)**

Selecting numerical threshold values. For the thresholds \(T_{\log K_{OW}}\) and \(T_{\log K_{OA}}\), we used the values of 2 and 5, which we believe have their origin in the food chain bioaccumulation calculations by Kelly and Gobas (2003) and Kelly et al. (2003). This choice is based on the fact that these values have been adopted by ECHA (2017) and thus have regulatory relevance. These should, however, not be considered necessarily as set in stone. Armitage and Gobas (2007) proposed a \(T_{\log K_{OW}}\) of 1.75 and a \(T_{\log K_{OA}}\) of 5.25 based on simulations of the soil–earthworm–shrew food chain. Simulations for food webs leading up to large, warm-blooded carnivores by Czub and McLachlan (2004) and Kelly et al. (2007) yielded a higher log \(K_{OA}\) threshold of 6. We further note that Czub and McLachlan (2004) refer to their choice of thresholds as “reasonable, but not especially conservative.”

Selecting the threshold \(T_{\text{HLB}}\) is trickier, as no regulatory recommendation currently exists and because elimination half-lives for different organisms (especially of different sizes and physiologies) are not directly comparable. In other words, the selected \(\text{HLB}\) threshold has to be appropriate for a particular organism. Because the prediction methods for \(\text{HLB}\) that we propose to use are for humans, the threshold should be appropriate for humans,

Goss et al. (2013) have suggested that a total elimination half-life of 70 days, equivalent to an elimination rate constant of 0.01 day\(^{-1}\), corresponds to a lipid-normalized biomagnification factor (BMF) of 1 under a certain set of assumptions, such as the daily food intake rate equaling 1% of body mass, a dietary uptake efficiency of 95%, and a dietary lipid content that is equal to the lipid content of the organisms. However, while these assumptions seem reasonably appropriate for humans, equally plausible values for the food intake rate or the lipid content of the food would yield different numerical values for the \(T_{\text{HLB}}\). Of course, different air-breathing organisms can also differ in their dietary characteristics. It is furthermore possible that the biotransformation rates in animals deviate from those in humans (Watkins & Klaasen, 1986; Williams, 1974). Nevertheless, we assume here that if an air-breathing animal with the average body size of a human can biotransform a chemical with an \(\text{HLB}\) of 70 days, it is unlikely to be bioaccumulative. There is some conservatism built in this assumption because biotransformation half-lives are compared with a threshold for the total elimination half-life, that is, the contribution of other elimination processes is neglected.
There is some hesitation to perform an assessment of B on elimination half-lives. The primary concerns are that the half-life that produces no biomagnification is highly dependent on the species and a number of assumptions, and that chemicals that are not efficiently eliminated do not necessarily bioaccumulate (e.g., if they are not taken up efficiently). However, the numerical value of BMF or the thresholds for log $K_{OA}$ and log $K_{OW}$ are similarly dependent on a number of assumptions, such as the lipid content of an organism or the temperature of exhaled air. Furthermore, here, we propose the use a threshold that is specific for a particular type of organism and we apply this threshold in such a way that if $H_L < T_{HL}$, then the chemical is not B, which does not imply that a chemical with an $H_L > T_{HL}$ is necessarily B. After all, a chemical with a log $K_{OA}$ or a log $K_{OW}$ or log $D > T_{logK_{OW}}$ is not necessarily bio-accumulative either.

Consistency of threshold values. It is interesting to explore whether the three thresholds for respiratory, urinary, and biotransformation loss that we adopt here are internally consistent. It is possible to estimate the respiratory elimination half-life of chemicals from the human body by assuming that exhaled air is in chemical equilibrium with the human body.

$$H_{respiration} = \frac{\ln 2 \cdot K_{OA} \cdot f_{lipid}}{G_{respiration}}, \quad (6)$$

where $G_{respiration}$ is the body mass-normalized respiration rate (L air kg$^{-1}$ body mass day$^{-1}$), $f_{lipid}$ is the lipid-equivalent fraction of the body (L lipid kg$^{-1}$ of body mass), and $K_{OA}$ applies to the temperature of exhaled air. Similarly, we can estimate the urinary elimination half-life of chemicals assuming that urine is in chemical equilibrium with the human body.

$$H_{urination} = \frac{\ln 2 \cdot K_{OW} \cdot f_{lipid}}{G_{urination}}, \quad (7)$$

where $G_{urination}$ is the body mass-normalized urination rate (L urine kg$^{-1}$ body mass day$^{-1}$).

Using Equations (6) and (7) with values for $f_{lipid}$ of 0.22 L kg$^{-1}$, $G_{urination}$ of 0.017 L kg$^{-1}$ day$^{-1}$, and $G_{respiration}$ of 250 L kg$^{-1}$ day$^{-1}$ (Goss et al., 2013), we estimate that an elimination half-life in humans of 70 days corresponds to a log $K_{OA}$ at body temperature of 5.06 and a log $K_{OW}$ of 0.90. Using a slightly different approach, Goss et al. (2013) obtained similar values of log $K_{OW}$ 1.3 and log $K_{OA}$ 5.5. If we assume a body temperature of 37°C and an internal energy of phase transfer from octanol to the gas-phase $\Delta U_{OA}$ of 50 kJ mol$^{-1}$ (Goss & Schwarzenbach, 1999), this corresponds to a log $K_{OA}$ at 25°C of 5.40. In other words, the log $K_{OA}$ threshold is broadly consistent with the chosen $H_L$ threshold, while the log $K_{OW}$ threshold seems to be too high. A chemical with a log $K_{OW}$ of 2 is estimated to have a whole-body elimination half-life by urination in humans of more than two years. A closer look at the boundaries in the chemical partitioning space plots displaying modeling results on biomagnification (Czub & McLachlan, 2004; Kelly et al., 2007) actually suggests that thresholds of a log $K_{OA}$ of 6 and a log $K_{OW}$ of 2 should be indicative of similar food chain biomagnification behavior. While these thresholds were derived from food web simulations and Equations (6) and (7) apply to a single human, both approaches agree that a $T_{log K_{OW}}$ of 2 and a $T_{log K_{OA}}$ of 5 are inconsistent. It may thus be worthwhile to revisit the regulatory thresholds for B in air-breathing organisms (ECHA, 2017).

Incidentally, Equations (6) and (7) illustrate that fixed log $K_{OW}$ and log $K_{OA}$ thresholds result in different elimination half-lives in different air-breathing organisms because organisms differ in terms of $f_{lipid}$, $G_{respiration}$, and $G_{urination}$, as well as in body temperature (poikilotherms vs. homeotherms). In other words, when applied to different organisms, not only will $T_{HL}$ require adjustment to achieve threshold consistency but quite possibly also $T_{log K_{OW}}$ and $T_{log K_{OA}}$. The reasons are differences in allometric scaling with respect to the lipid content and the rates of feeding, respiration and urination.

In summary, while we apply here numerical values of 2, 5, and 70 days for $T_{log K_{OW}}$, $T_{log K_{OA}}$, and $T_{HL}$, we stress that these values should be considered tentative because a $T_{log K_{OW}}$ of 2 appears not to be consistent with a $T_{log K_{OA}}$ of 5, and because all thresholds, and especially $T_{HL}$, are dependent on a number of assumptions. The proposed procedure can obviously be applied to different thresholds, should a set of consistent ones be agreed upon in the future.

Predicting log $K_{OW}$ and log $K_{OA}$. There are several techniques for predicting the log $K_{OA}$ of a chemical (Baskaran et al., 2021a) and many more for the prediction of log $K_{OW}$ (e.g., Cappelli et al., 2015). The criteria that we applied for selecting the prediction technique for these two parameters included the following: (i) It should have a good predictive performance, particularly within the range close to the thresholds of interest. (ii) It should be possible to make these predictions easily, without having to access commercial software. (iii) It should be possible to derive a confidence interval on the predictions.

The work by Endo and Goss (2014) and co-workers (Stenzel et al., 2013) has demonstrated the superior performance of poly-parameter linear free-energy relationships (ppLFER) for the prediction of the solvation properties of a large diversity of neutral organic compounds. For example, Stenzel et al. (2013) reported a root mean square error (RMSE) of 0.39 log units from the comparison of measured and predicted $K_{OW}$ for 111 substances (86 pesticides and pesticide transformation products, 10 PCBs, six heterocyclic aromatic compounds containing either nitrogen (N) or oxygen (O), and seven nitroaromatic compounds, one hormone and one antibacterial). In the specific case of the log $K_{OA}$, Baskaran et al. (2021a) have shown that ppLFERs yield predictions with smaller confidence intervals than other easily accessible and commonly used prediction techniques. The noncommercial UFZ LSER website further assures that anyone can make such predictions and derive their

Integr Environ Assess Manag 2022:1297–1312 DOI: 10.1002/ieam.4555 © 2021 The Authors
confidence limits (Ulrich et al., 2017). Therefore, our proposed procedure relies on the following ppLFERs to predict the $K_{OA}$ and $K_{OW}$ values of compound $i$:

\[
\log K_{OWi} = -\left(1.41 \pm 0.05\right) \cdot S_i - (0.18 \pm 0.04) \cdot A_i \\
- (3.45 \pm 0.05) \cdot B_i + (2.41 \pm 0.10) \cdot V_i \\
+ (0.43 \pm 0.02) \cdot L_i + (0.34 \pm 0.04),
\]

\[
\log K_{OAl} = (0.42 \pm 0.03) \cdot S_i + (3.52 \pm 0.07) \cdot A_i \\
+ (0.84 \pm 0.05) \cdot B_i + (0.92 \pm 0.01) \cdot L_i
\]

where $S_i$, $A_i$, $B_i$, $V_i$, and $L_i$ are solute descriptors (SDs) quantifying a compound's ability to undergo various types of intermolecular interactions. Specifically, $S_i$ is the polarizability/dipolarity parameter, $A_i$ and $B_i$ are the H-bond acidity and basicity, $V_i$ is the McGowan molar volume, and $L_i$ is the decadic logarithm of the hexadecane–air partitioning ratio. Equation (8) is taken from Endo and Goss (2014), whereas Equation (9) is a version of a ppLFER for log $K_{OA}$ presented by Endo and Goss (2014), which eliminates the term containing $V_i$ because it has no statistical significance (Baskaran et al., 2021a). Please note that Equations (8) and (9) are currently not implemented on the UFZ LSER website (Ulrich et al., 2017).

The performance of ppLFER predictions depends on the origin of the SDs, with smaller prediction errors obtained with experimentally derived SDs. In the case of predicting log $K_{OA}$ with Equation (9), this has been shown by Baskaran et al. (2021a). Experimentally determined SDs, including those in the ABSOLV database (ACD/Labs), are available for close to 10,000 organic chemicals and accessible through the UFZ LSER database (Ulrich et al., 2017). If no experimental SDs are available, they can be predicted with quantitative structure–property relationships (QSPRs) (e.g., Platts et al., 1999, 2000), but one should then expect considerably higher PEs, in particular, if a compound's structure requires an extrapolation beyond the applicability domain (AD) of the QSPRs used for SD prediction. The QSPRs, developed using the Iterative Fragment Selection (IFS) algorithm (Brown et al., 2012) for the SDs in Equations (8) and (9), are also available (e.g., Brown, 2014) and accessible through the UFZ LSER website (Ulrich et al., 2017) and only require a chemical's molecular structure as a simplified molecular input line entry system (SMILES) notation as input.

The proposed procedure uses experimental SDs to predict log $K_{OW}$ and log $K_{OA}$, if they are available, and otherwise uses the ifs-QSPR predicted ones. The upper PEs are obtained differently depending on whether experimental or predicted SDs are used.

### Estimating the error of log $K_{OW}$ and log $K_{OA}$

Estimating the error of log $K_{OW}$ and log $K_{OA}$ when using experimental SDs. If experimental SDs are used, we propose using a PE that is empirically derived from the comparison of measured and ppLFER-predicted partition ratios in the threshold region (i.e., within two and a half orders of magnitude of $T$) and is a constant for all chemicals. In the case of log $K_{OA}$, we extracted all measured values within the range 2.5 to 7.5 from the database compiled by Baskaran et al. (2021b). In the case of log $K_{OW}$, we extracted all measured values, excluding salts, within the range $–0.5$ to $4.5$ from the OPERA database (Mansouri et al., 2018). To remove any inconsistency and correctly assign a structure to a chemical, SMILES and CAS number data were standardized and curated using an in-house routine based on the software OpenBabel v3.1.1 (O’Boyle et al., 2011) and the R package Webchem v0.5.0 (Szöcs et al., 2020). We then calculated the difference between the experimental values and the values calculated using Equations (8) and (9) together with experimental SDs from the UFZ LSER database (giving preference to UFZ-preselected values over values from the ABSOLV database). The statistics on these differences are presented in Table 1. The PE is derived as either $PI95 = \text{mean} \pm 1.96 \cdot \text{StDev}$ or $SE = \text{mean} \pm \text{StDev}$, whereby mean and StDev refer to the mean and the standard deviation of the differences between the measured and predicted values.

Based on this analysis, we use the following versions of Equations (2) and (3) when experimental SDs are available. If we apply the 95% prediction interval $PI95$:

If $(\log K_{OWi} + 0.58) < 2$,

i.e., $(\log K_{OWi} < 1.42)$, then chemical is not B, \( (2') \)

If $(\log K_{OAl} + 0.35) < 5$,

i.e., $(\log K_{OAl} < 4.65)$, then chemical is not B. \( (3') \)

Also, when using the standard error of the prediction $SE$:

If $(\log K_{OWi} + 0.26) < 2$,

i.e., $(\log K_{OWi} < 1.74)$, then chemical is not B, \( (2'') \)

If $(\log K_{OAl} + 0.18) < 5$,

i.e., $(\log K_{OAl} < 4.82)$, then chemical is not B. \( (3'') \)
Estimating the error of log $K_{OW}$ and log $K_{OA}$ when using predicted SDs. If predicted SDs are used, the error of the prediction is expected to vary based on the fit of a chemical with the AD of the QSRRs used in the prediction of the SDs. The IFS-based QSRRs quantify this fit and assign a standard error $SE_x$ to an SD $x$, predicted for compound $i$ ($X_i = S_i, A_i, B_i, V_i, L$). As can be seen in Equations (8) and (9), the system constants $x$ in a ppLFER have an uncertainty $SE_x$ as well ($x = s, a, b, v, l$, and intercept $c$). These constants are derived during the original calibration of a ppLFER, which involves multiple linear regression of measured partition ratios against experimental SDs. In other words, in a ppLFER equation, uncertain system constants are combined with uncertain SDs to predict a chemical property. We therefore propose to estimate the compound-specific standard error $SE_{\log K_{OA}}$ of the predicted log $K_{OA}$ by propagating the standard errors $SE_x$ of the SDs and the standard error $SE_c$ of the fitted system constants in Equation (9) as:

$$SE_{\log K_{OA}} = \sqrt{(S_i \cdot SE_x)^2 + (s \cdot SE_v)^2 + (a \cdot SE_a)^2 + (A_i \cdot SE_a)^2 + (b \cdot SE_b)^2 + (B_i \cdot SE_b)^2 + (L \cdot SE_l)^2 + (L \cdot SE_l)^2 + (l \cdot SE_l)^2 + SE_c^2}$$

(10)

where $SE_c$ is the error of the intercept in Equation (9). An equivalent equation with additional square terms for the error of $\nu_i$ in Equation (8) is used to estimate the $SE_{\log K_{OW}}$ of the predicted log $K_{OW}$. In doing so, we neglect the contribution of any covariance terms. Finally, we multiply the $SE_{\log K_{OA}}$ or $SE_{\log K_{OW}}$ by 1.96 to estimate the $PI_{95}$ log $K_{OA}$ and $PI_{95}$ log $K_{OW}$.

Note that the same compound-specific approach is not feasible for predictions relying on experimental SDs because the uncertainty of those SDs is rarely known quantitatively. The uncertainty of experimentally determined SDs has occasionally been reported for individual chemicals (Tulp et al., 2008). In other cases, an estimate of the uncertainty of a set of SDs $S_i, A_i, B_i, V_i,$ and $L_i$ for a compound is estimated from the RMSE between the ppLFER predicted and experimental log $K$ for the systems used in the calibration of that set of SDs (called “internal consistency” by Stenzel et al., 2013). Often, however, no uncertainty information is available, which is the case for the bulk of the SDs available through the UFZ LSER database (Ulrich et al., 2017). Therefore, it is generally not possible to estimate the uncertainty of a partition ratio predicted with a ppLFER and experimental SDs on a compound-specific basis.

To confirm the appropriateness of the error propagation calculation using Equation (10), and specifically of the assumption that the SDs are independent and not correlated with each other and that the covariance terms can therefore be neglected, we used two other methods to estimate the error of a ppLFER prediction using predicted SDs. One of these relies on a comparison of measured and predicted $K_{OA}$ and $K_{OW}$ values, that is, it uses the approach applied to predictions with experimental SDs. The other is a Monte Carlo simulation. Details for both are provided in the Supporting Information. Monte Carlo simulations yielded $PI_{95}$ that are essentially identical to those obtained by error propagation. The SE and $PI_{95}$ obtained by error propagation are of the same order of magnitude as those obtained from the model–measurement comparison, even though they relate to completely different sets of chemicals. Again, details are provided in the Supporting Information.

Accordingly, the following versions of Equations (2) and (3) are used when predicted SDs are used. If we apply the 95% prediction interval:

If $log K_{OWi} + PI_{95 log K_{OWi}} < 2$, then chemical is not B,

(2”)

If $log K_{OAi} + PI_{95 log K_{OAi}} < 5$, then chemical is not B.

(3”)

Also, when using the standard error of the prediction:

If $log K_{OWi} + SE_{log K_{OWi}} < 2$, then chemical is not B,

(2”*)

If $log K_{OAi} + SE_{log K_{OAi}} < 5$, then chemical is not B.

(3”*)

Finally, irrespective of the use of experimental or predicted SDs, we use:

If $log K_{OWi} < 2$, then chemical is not B,

(2”**)

If $log K_{OAi} < 5$, then chemical is not B.

(3”**)

when ignoring the error of the prediction.

How does this approach account for the AD of the individual models? Because different models quantify the AD differently, they are generally not directly comparable. The AD information serves as a warning that the prediction for a particular chemical is less reliable; it is not necessarily wrong, but simply has a higher uncertainty. Thus, by including the uncertainty in the assessment, we indirectly include information about the AD of the models. Chemicals that have a poor fit with the AD are assigned larger errors. This also serves as an incentive to seek out models appropriate for each chemical, since this will reduce the error associated with a prediction and thereby increase confidence in the assessment.

**Accounting for dissociation and the formation of charged species.** Dissociation and the formation of charged species can affect the elimination of organic chemicals from air-breathing organisms in two ways. It will increase the water solubility of a compound and therefore increase the possibility for urinary excretion and it will decrease the volatility of a compound and therefore might prevent exhalation loss, especially if the chemical is entirely present in its charged form. Because Equations (8) and (9) only predict the partitioning behavior of the neutral form of a compound, this could result in misclassifications. If it is possible to gain access to a...
technique predicting the dissociation of neutral chemicals, the proposed procedure should also contain additional steps.

If a chemical has been judged not B based on Equations (3)”–(3″″′), that is, if its neutral species is assessed to have a log $K_{OA}$ below the threshold of 5, it should be confirmed that the compound can exist in neutral form at the physiological pH of 7.4. If the fraction of the neutral species is predicted to be less than 1%, the chemical cannot be assumed to be not B. Underlying this approach is the assumption that 1% of a chemical with a log $K_{OA}$ less than 5 being present in the neutral form is sufficient for rapid respiratory elimination. We used the Global, Adjusted Locally According to Similarity (GALAS) algorithm (Sazonovas et al., 2010) by ACD/Percepta (2021) to estimate the percent of a chemical in the neutral form.

Even if a chemical has been judged to have a log $K_{OW}$ above 2 based on Equations (2)”–(2″″′), it may still be possible to exclude it as a bioaccumulation concern if it is dissociating and has a log D below that threshold. We therefore include a comparison of a chemical i’s log D, at pH 7.4, again predicted using the GALAS algorithm by ACD/Percepta (2021), with the log $K_{OW}$ threshold as in Equation (4). This software does not provide an estimate of the reliability of a log D prediction that would be suitable to estimate a prediction error. However, it does provide a reliability index (RI) for its log $K_{OW}$ predictions. We used the same data set of measured log $K_{OW}$ values that we used in the derivation of a prediction interval for the ppLFER-predicted log $K_{OW}$ to evaluate the performance of GALAS for predicting log $K_{OW}$ values in the threshold region between −0.5 and 4.5. We calculated the statistics on the differences between the measured and predicted log $K_{OW}$ for four groups of compounds, depending on the value of the RI (Table S2). The PE of the log D will have to be larger than the PE of the log $K_{OW}$ from which it is calculated because of the additional uncertainty introduced by the prediction of the acid dissociation constant(s) $pK_a$ and the calculation of dissociation. On this basis, we apply an $SE_{logD}$ of 0.38, 0.75, and 1.5 and a $PI_{95_{logD}}$ of 0.75, 1.5, and 3, if the RI for log $K_{OW}$ is greater than 0.75 (high reliability), between 0.5 and 0.75 (medium) and below 0.5 (low/poor), respectively.

Depending on what error is applied, Equation (4) thus takes either the form:

If (log D, at pH7.4 + $PI_{95_{logD}}$) < 2, then chemical is not B,

(4′)

If (log D at pH7.4 + $SE_{logD}$) < 2, then chemical is not B.

(4″)

Estimating biotransformation loss. Predicting biotransformation half-lives $H_{LB}$ and their error: We are not aware of techniques predicting rates of biotransformation in air-breathing animals (Papa et al., 2017). However, Arnot et al. (2014) and Papa et al. (2018) have presented prediction techniques for the biotransformation half-life $H_{LB}$ in humans, which are based on the IFS algorithm and global molecular descriptors selected by the Genetic Algorithm implemented in the QSARINS software (Gramatica et al., 2013), respectively. Both were trained with the same empirical data set (Papa et al., 2018). In the absence of prediction techniques for air-breathing animals, we suggest that it is defensible to rely on predicted biotransformation rates in humans because humans have enzymatic systems for the biotransformation of xenobiotic compounds that are not fundamentally different from those of other air-breathing organisms (Williams, 1974). Because it is not known whether the IFS-QSOPR or QSARINS provides more reliable predictions, we propose relying on the geometric mean of the $H_{LB}$ values provided by the two techniques:

$$H_{LB} = 10^{(log_{HLB_{IFS}}+log_{HLB_{QSN}})/2}.$$  (11)

The uncertainty of the $H_{LB}$ prediction is related to how well a molecule fits within the AD of the prediction methods, and both IFS and QSARINS estimate a prediction error from that fit. Specifically, a prediction interval of $PI_{95_{H_{LB}}}$ indicates that there is a 95% likelihood that the true value of $H_{LB}$ is between $H_{LB}/PI_{95_{H_{LB}}}$ and $H_{LB} \cdot PI_{95_{H_{LB}}}$. Similarly, we can estimate standard errors $SE_{H_{LB}}$ using $log_{10}(PI_{95_{H_{LB}}})/1.96$, such that there is a 68% likelihood that the true value of $H_{LB}$ is between $H_{LB}/SE_{H_{LB}}$ and $H_{LB} \cdot SE_{H_{LB}}$.

The value that should be compared with the threshold $T_{H_{LB}}$ therefore is the geometric mean of the upper limits of the two techniques, that is, $H_{LB_{IFS}} \cdot PE_{H_{LB_{IFS}}}$ and $H_{LB_{QSN}} \cdot PE_{H_{LB_{QSN}}}$. When using the 95% prediction error, statement (5) therefore reads:

If $10^{(log_{HLB_{IFS}}+log_{PI95_{H_{LB_{IFS}}}+log_{HLB_{QSN}}+log_{PI95_{H_{LB_{QSN}}}}})/2} < T_{H_{LB}}$, then chemical is not B.  

(5′)

When using the standard error, we use:

If $10^{(log_{HLB_{IFS}}+log_{SE_{H_{LB_{IFS}}}+log_{HLB_{QSN}}+log_{SE_{H_{LB_{QSN}}}}})/2} < T_{H_{LB}}$, then chemical is not B.  

(5″)

When we ignore the error of the prediction, we simply compare the geometric mean of the predictions with the threshold:

If $H_{LB} = 10^{(log_{HLB_{IFS}}+log_{HLB_{QSN}})/2} < T_{H_{LB}}$, then chemical is not B.  

(5‴)

Schematic of the overall approach. Figure 2 shows the entire approach in a flow chart. In particular, it clarifies the sequence of the various comparisons with threshold values. The approach is hierarchical in the sense that if a chemical is judged by any one comparison to be not bioaccumulative, subsequent comparisons are no longer required. This is why the comparisons involving the freely and easily accessible ppLFER predictions precede those requiring the commercial ACD/Labs predictions.
We suggest that the comparison involving the log $K_{OH}$ precede the one involving the log $K_{OA}$ because the former turns out to be more effective than the latter, that is, it can judge a large fraction of chemicals as being not B (see the Results section below). The comparison involving the biotransformation half-life is last because not everyone may agree that a first tier should include information beyond partitioning ratios (in analogy to a tier 1 assessment for aquatic bioaccumulation solely based on a log $K_{OW}$ value).

**Testing the applicability and effectiveness of the approach**

We applied the approach illustratively to 1158 organic chemicals to test (i) how readily applicable the various property estimation techniques are to a large number of diverse organic chemicals and (ii) how effective the approach is in eliminating chemicals as potential bioaccumulation threats in air-breathing organisms. The data set was extracted from a list of compounds collected in the context of the CEFIC LRI ECO41 and ECO44 projects (https://cefic-lri.org/projects/) aimed at gathering in vitro and in vivo toxicokinetic data for mammals and included in the EAS-E Suite database (www.eas-e-suite.com). The complete data set is very heterogeneous, comprising pharmaceuticals and environmentally relevant compounds (e.g., PCBs, PBDEs, and PFAs), covering a broad range of chemistries (e.g., neutral and ionizable compounds) and featuring partitioning properties (i.e., $K_{OW}$ and $K_{OA}$) spanning over 10 orders of magnitude. To test our approach on a relevant area of the chemical space (i.e., substances that could possibly be subject to respiratory elimination), we selected compounds with a log $K_{OA}$ estimated with EPISuite’s KOAWIN (US EPA, 2021) below 10, thus limiting the test data set to the 1158 chemicals reported in the Supporting Information. By excluding very involatile substances, the data set is biased toward relatively volatile organic compounds.

We performed the screening procedure for the set of chemicals three times, either using the PI95 or the SE as the prediction error or ignoring the uncertainty of the predicted property values altogether. Using SE instead of PI95 corresponds to a larger tolerance of false-negative categorizations. A justification for proceeding this way, could be, for example, if the threshold values have already been selected conservatively, that is, are in fact lower than they really need to be protective. Using the SE to express the prediction error can be used to assess how much more effective the procedure would be in identifying rapidly eliminated substances if more room for a wrong decision is provided. The screening without consideration of prediction errors can be used to assess the number of false negatives.
that one would incur, if one does not quantify and consider the prediction uncertainty.

**RESULTS**

**Applicability of property prediction techniques**

The ppLFER predictions of log $K_{OW}$ and log $K_{OA}$ could be made for all except one of the 1158 chemicals, that is, the prediction only failed for paraquat (CAS 004685-14-7). Experimental SDs could be found for 532 chemicals, implying that they had to be predicted for the remaining 625 substances. The $HL_{B}$ predictions by QSARINS were successful for all chemicals, whereas the IFS-$HL_{B}$ QSPR failed again only for paraquat. Predictions of the neutral fraction and log $D$ could be made by ACD/Labs for all chemicals.

Overall, the fraction of chemicals with successful application of the property prediction techniques was therefore very high, and the likelihood that no judgement can be passed as a result of the failure to make a prediction is very low.

**Effectiveness of eliminating nonbioaccumulative substances**

The flow of the 1157 chemicals through the total procedure is depicted in Figure 3. The three panels within that diagram indicate the flow through the procedure by applying either the 95% PIs or the SEs or by assuming that the predicted properties have no error. The flow path of the 532 chemicals with experimental SDs is shown in red, whereas the 625 chemicals for which SDs needed to be predicted are shown in blue. Arrows leaving a diamond towards the left indicate a “Yes” decision, generally indicating that a chemical is judged to be below a threshold, whereas an arrow leaving a diamond on the right indicates a “No,” in most cases indicating that a chemical could not be confidently judged to have a property below a threshold.

If the most stringent uncertainty estimate ($PI_{95}$) are applied, 960 out of 1157 chemicals (83%) are eliminated as not being a bioaccumulation concern in air-breathing organisms. This fraction increases to 89% and 95% if the standard error or no error is applied. Accordingly, 17%, 11%, or 5% of screened chemicals are judged to require a more sophisticated B assessment. In other words, tolerating a potential for a larger number of false negatives decreases the number of chemicals requiring higher-tier assessments by ~6% (both when moving from $PI_{95}$ to SE and from SE to no error). By categorizing approximately four-fifths of compounds as being of no bioaccumulation concern, the procedure functions as a reasonably effective filter; comparison of the three different panels in Figure 3 indicates that this effectiveness as a filter is not unduly dependent on the parameter chosen to represent the prediction error.

Not too much emphasis should be placed on the numerical value of the percentages shown in Figure 3 (and also Figure 4), as they will clearly depend on the specific compound data set being assessed. Because of a relatively large number of polychlorinated biphenyls in our test set, the proportion of chemicals assigned to the “not B” category would likely be even larger for a data set that is truly representative of the universe of commercial chemicals and not biased by a high proportion of highly halogenated compounds. By using a reasonably diverse and large data set, this analysis should at least be broadly suggestive of the effectiveness of the proposed procedure, and especially allow for the comparison of the effectiveness of different variations (e.g., including biotransformation or not,
including log D or not, using different types of error), as will be done in the next section.

A valid question is whether chemicals are correctly categorized by the procedure. Because the stated goal is de-prioritization, success can only be assessed by confirming that de-prioritized chemicals are indeed not B. In other words, whether the chemicals earmarked for higher-tier assessment are in fact B cannot be a measure of success; it is sufficient for them to be plausibly B. The Microsoft® Excel sheet in the supporting information identifies all of the chemicals that are judged not B and potentially B. While it is difficult to state affirmatively that all de-prioritized chemicals cannot bioaccumulate in air-breathing organisms, we can note the absence of substances among them that are known to bioaccumulate. Furthermore, compounds with known bioaccumulation potential in air-breathing organisms, in particular, organochlorinated pesticides (e.g., endosulfan, β-hexachlorocyclohexane) and a large number of polychlorinated biphenyls and benzenes (e.g., 1,2,4,5-tetrachlorobenzene), are identified as requiring higher-tier assessment (see Table S3). Food web modeling had identified trifluralin as potentially bioaccumulative in air-breathing organisms (see Tab. 1 in Kelly et al., 2007), whereas it was de-prioritized here based on rapid biotransformation. Field studies examining the potential bioaccumulation of trifluralin in food webs composed of air-breathing organisms have found no evidence of bioaccumulation (e.g., Morris et al., 2016; Vorkamp & Rigét, 2014).

**Processes responsible for eliminating nonbioaccumulative substances**

Figure 4 summarizes the categorization of 1157 chemicals by the proposed procedure. There are three versions depending on the values that were used for the prediction error. The largest fractions of chemicals are declared not B because of rapid elimination by respiration (yellow bars).

A screening method based solely on partitioning properties would be quite inefficient, as only between a third (PI95) and half (no E) of the chemicals would be classified as not B. If the screening method were to rely only on log K OW and log KOA as elimination criteria, that is, if dissociation and log D were not taken into account, a mere 21%–37% of chemicals would be classified as not B. In other words, the consideration of biotransformation as an elimination process is important, as it accounts for 60%–50% of the decisions to declare a chemical not B.

The inefficiency of screening based on log KOA and log K OW can be illustrated by placing the 1157 chemicals in a two-dimensional partitioning space defined by these two partition ratios (Figure 5). The whiskers on each datapoint represent the PI95, either being the same for all chemicals with experimental SDs (orange markers) or with compound-specific prediction errors for chemicals with predicted SDs (blue markers). The thresholds of log K OW 2 and log KOA 5 are shown by colored shading. Only for those chemicals for which at least one of the two whiskers falls fully within the green shaded area of the plot is a judgment of not B reached based on ppLFER-predicted log KOA values.

Very few chemicals fall in the green area at the bottom of the plots in Figure 5, which indicates that the number of chemicals classified as not B based on the volatility threshold is small (Figure 4). Among the multitude of organic chemicals, only a very small fraction is sufficiently volatile for respiration to be an efficient elimination pathway. For example, among mono-functional linear alkyl compounds, the C6 alcohol, the C8 primary amine, and the C9 aldehyde and ketones have a log KOA above 5. Very few compounds with two functional groups have a log KOA below that threshold. For example, 1,2-ethanediol, 2-nitropropanol, 3-aminopropanol, and 5-aminopentanal all have log KOA above 5, as do aniline, nitro-toluene, and
para-benzoquinone. Despite the small number of chemicals eliminated by the volatility criterion, we believe it nevertheless advisable to include it in a screening assessment, as experiments that may be required at a higher tier of assessment (e.g., an in vitro assay to determine the bio-transformation kinetics of a compound) could be particularly challenging to conduct for such volatile compounds.

Considerably more chemicals fall in the green part to the left of those plots (Figure 5), which is consistent with a larger fraction of substances being sufficiently water soluble for efficient urinary elimination. However, the bulk of the chemicals fall in the upper right of the plots and therefore cannot be judged not B based on partitioning properties alone. The plots also make it apparent that screening for B in fish based solely on partitioning (i.e., based on a log $K_{OW}$ threshold of 4.5 [ECHA, 2017b] or 5 [UNEP, 2017]) is more effective than screening for B in air-breathing organisms based solely on partitioning because of the large number of chemicals with a log $K_{OW}$ between 2 and 4.5 or 5. This had already been highlighted by Kelly et al. (2003), who noted that 36% of the chemicals on the Canadian domestic substances list are deemed not B based on a log $K_{OW}$ threshold of 5, but could conceivably be B in air-breathing organisms.

If there is a large increase in the number of chemicals being judged not B upon relaxing the error stringency (i.e., moving from $PI_{95}$ to SE to no error in Figure 4), this implies that there is a potential for a relatively large number of false negatives, if uncertainty is not considered. However, it also implies that the efficiency of the procedure would potentially benefit from a prediction technique with a smaller error. This is apparent for the ppLFER prediction of log $K_{OW}$ with estimated SDs (intermediate blue bars in Figure 4). The number of chemicals that can be excluded using the log $K_{OW}$ threshold increases from 91 to 163 and 241 if $PI_{95}$, SE, and no error are applied. The fairly large error of a log $K_{OW}$ predicted with estimated SDs is also apparent from the large blue whiskers in the right panel of Figure 5.

If we compare which elimination processes are responsible for a chemical being classified as not B, we can observe a shift of increasing importance of urinary excretion relative to the importance of bio-transformation with decreasing assumed error (Figure 4). For example, 36% and 60% of the not B decisions are based on urinary excretion and bio-transformation, respectively, when $PI_{95}$ is applied, but these shift to 50% and 47% when the error of the predicted properties is ignored. The reason is simply the sequence of the various comparisons in the procedure. If a chemical is already sorted out in an earlier step of the procedure (e.g., based on a comparison of the ppLFER-predicted log $K_{OW}$), the relative importance of subsequent steps in identifying not B chemicals is obviously diminished.

**DISCUSSION**

*Substituting the prediction methods proposed here*

While we relied entirely on predicted property values, it is obviously possible to substitute a measured value in any of
the decisions 2–5, if such a value is available. However, measured property values also have an uncertainty and it would be appropriate to apply an equivalent expression of measurement uncertainty, such as a 95% confidence interval or a standard error. Importantly, precision as obtained by replicate measurements often underestimates the true experimental uncertainty because of unknown measurement bias. Also, not every measured value is necessarily better than a prediction. For example, only high-quality measurements will be able to match or reduce the uncertainty of a log $K_{OA}$ prediction with Equation (9) and experimental SDs.

Similarly, it is possible to substitute any of the prediction methods proposed here with one perceived to be superior or with one that is more easily accessible. As mentioned earlier, the log $K_{OW}$ prediction with pplFERs incurs quite high uncertainties when estimated SDs are used, in particular, if a molecule does not fall within the AD of one of the IFS-QSPR used for the prediction of the SDs. It is conceivable that methods exist that are capable of predicting log $K_{OW}$ from molecular structures with a smaller prediction error, but they may often be commercial and/or fail to provide a quantitative estimate of prediction errors. The latter is a key prerequisite for the use of a prediction method in the proposed procedure.

A different method obviously could also substitute for a prediction that failed. For example, in contrast to the IFS-QSPR, the commercial software ACD/LABS ABSOLV succeeds in predicting SDs for paraquat. Equations (6) and (7) can be used to estimate log $K_{OW}$ and log $K_{OA}$ values for the neutral form of 4.70 and 5.99, respectively, but no information on the error of these predictions is available. In fact, a measured value is almost 10 orders of magnitudes lower (log $K_{OW} = -5$, Platford, 1983), presumably because of the two positive charges in the molecule.

In the future, it is likely that QSPRs will be developed that predict the biotransformation half-life of organic chemicals in air-breathing organisms, for example, based on the extensive data sets that have been generated on the toxicokinetics of organic compounds in rats and mice. It may then be advisable to substitute, or rather complement, predictions of $H_L$ in such organisms for the $H_L$ in humans used here. This seems prudent because while all air-breathing organisms share the basic pattern of metabolism, large variations between species can arise within this pattern (Williams, 1974). In this case, it is imperative that the threshold value $T_{H_L}$ is adjusted to the body size of the organism.

Here, we proposed a procedure that uses $K_{OA}$ and $K_{OW}$ to judge a chemical’s propensity for elimination, which implies the assumption that n-octanol is a suitable surrogate for describing the solvation properties of the tissues of air-breathing organisms. There are pplFERs available that predict solvation of neutral organic substances in different biological materials such as storage and membrane lipids and proteins (Endo et al., 2011, 2012; Geisler et al., 2012), which allows for the estimation of an entire organism’s uptake capacity from an organism’s biological composition and a chemical’s SDs (Endo et al., 2013), that is, the same data that we propose to use for $K_{OW}$ and $K_{OA}$ predictions. It should be feasible to formulate comparisons 2 and 3 in terms of partitioning ratios between the whole organism and the gas and aqueous phases, and predict these partition ratios with the same approaches used here and thus avoid the use of octanol as a surrogate entirely. While such an approach would be more mechanistically realistic, thresholds in $K_{organism/OA}$ and $K_{organism/OA}$ that correspond to rapid elimination would still need to be established (Goss et al., 2013). Also, current regulatory thresholds tend to be based on parameters that not only can be predicted but also are readily accessible to reproducible experimental determination.

**Strength of the proposed procedure**

Of course, one could imagine many different approaches for a tier 1 assessment of bioaccumulation in air-breathing animals, including many based on the same predicted properties as used here. For example, one could envisage using all of the comparisons of a chemical’s log $K_{OW}$ and log $K_{OA}$ and $H_L$ with the threshold values, that is, to not impose a sequence to the comparisons and not disregard a comparison, if an earlier one had already classified a chemical as not B. One could also imagine using the predicted chemical properties as input to a mass balance model of contaminant fate in a representative air-breathing organism and use a model-derived parameter as a criterion for assessing a chemical as potentially B or not, for example, a total elimination half-life (Goss et al., 2013) or a BMF.

Among the advantages of the proposed procedure are:

- Its conceptual simplicity and transparency.
- Its reliance on partition ratio thresholds that have already been enshrined in chemical regulation (e.g., ECHA, 2017a).
- Its use of properties that can be predicted for a very wide range of organic chemicals.
- The potential to flexibly exclude certain threshold comparisons, if the prediction is not accessible (e.g., because of the need to access a commercial software) or because of conceptual preferences (e.g., not wanting to include the potential for biotransformation in a tier 1 assessment).
- Its tolerance for missing property values, a often a decision of not B can be rendered based on predicted values that are available. For example, although no IFS-based prediction could be obtained for paraquat, QSARIN predicted an $H_{OA}$ of 12 days and an $H_{OA}$ of 3 days, rendering it not B.
- The relative ease of considering the prediction uncertainty in the assessment process.
- The possibility to account for dissociation, if relevant, under physiological conditions.

In particular, the latter two aspects are more difficult to implement when using an integrative bioaccumulation model. For example, the consideration of parameter
uncertainty presumably would require a full Monte Carlo analysis to propagate the uncertainty of the chemical properties to the model output that is to serve as the bioaccumulation criterion. While this is feasible, it is operationally difficult to implement when screening large numbers of chemicals. Bioaccumulation models treating dissociating organic compounds are only in their infancy (Armitage et al., 2017) and may not even exist yet for air-breathing organisms. Also, consensus would have to be developed on (i) the organism to be modeled, (ii) the model approach as well as the model parameterization, (iii) the model-generated output parameter to serve as the B criterion, and (iv) the threshold value to identify potentially bioaccumulative substances.

**Weaknesses of the proposed procedure**

There are also shortcomings of our approach: By using separate thresholds for efficient elimination by respiration, urination, and biotransformation, our scheme classifies a chemical as potentially B if none of the three processes by itself is fast enough to prevent bioaccumulation. If such a chemical is efficiently eliminated by a combination of more than one elimination process, the approach would thus lead to a false-positive categorization. Whereas an approach using an integrated bioaccumulation model that considers the various processes working in concert would avoid such errors, false positives are not a major concern in a screening aimed at de-prioritization. Furthermore, the number of chemicals affected is likely to be very small. For example, only for chemicals with partitioning properties falling into the small triangular area close to the threshold values in Figure 5 would the combination of exhalation and urination prevent bioaccumulation.

The proposed approach would also categorize some perfluorinated acids as not B based on their very low log D values, which would suggest that they are rapidly eliminated by urinary excretion. Our procedure does not consider the phenomenon of renal reabsorption, which essentially prevents the efficient urinary excretion of some such substances that therefore are potentially highly bioaccumulative. Examples are perfluorooctanesulfonate (CAS 2795-39-3), perfluorobutane-sulfonate (CAS 29420-49-3), perfluorohexanoic acid (CAS 307-24-4), perfluoropentanoic acid (2706-90-3), and perfluorobutanoic acid (375-22-4), all of which have predicted log D values below 0.

In general, the treatment of dissociating compounds in the procedure proposed here should be considered as preliminary. For example, while we use here the pH of blood to calculate the log D, the pH of urine is often lower, that is, the dissociation equilibrium will be different in blood and urine for compounds with a pK_a or pK_b around 7. As the permeability between blood and urine is different for charged and neutral species, this can lead to ion trap phenomena and, accordingly, to a pH dependence in the rate of urinary excretion (e.g., Kiddie et al., 1972). In fact, it may not be entirely valid to compare a log D value against a threshold for efficient urinary excretion expressed in terms of log K_OW. The assessment of the bioaccumulation potential of dissociating compounds may ultimately need to rely on toxicokinetic models specifically designed for such compounds (Armitage et al., 2017).

The success of the proposed procedure relies very strongly on the ability to predict biotransformation half-lives. The uncertainty of the prediction methods applied here is large and therefore the screening results depend considerably on the error value chosen. The list of chemicals deemed potentially bioaccumulative is strongly dominated by highly halogenated compounds and only includes one substance with less than two halogens (D5, decamethylcyclopentasiloxane, CAS 000541-02-6) (Table S3). It is conceivable that the calibration data set for the QSPRs for HLB was not sufficiently diverse to allow identification of poorly biotransformed structures that do not contain multiple halogens. Of course, it is also possible that our test data set of 1158 chemicals did not contain such substances.

**Higher tiers of assessment**

Chemicals that cannot be excluded as potentially B using a simple tier 1 procedure based on predicted chemical properties, as described here, will need to be subjected to a more sophisticated screening. The current study suggests that the focus of higher tiers should be on better characterization of biotransformation loss, and not necessarily on measured as opposed to predicted partitioning properties. The most promising route for this is the determination of biotransformation rates in in vitro assays using a system (e.g., S9 fraction, hepatocytes) for air-breathing organisms and the application of in vitro in vivo extrapolation methods to estimate the rate of elimination by biotransformation.

**DATA REPOSITORY**

A Microsoft® Excel spreadsheet distributed as supplemental data contains the entire data set of 1158 chemicals, most of the predicted properties and their error, as well as the decisions made on a comparison of the predictions with threshold values.

**ACKNOWLEDGMENT**

We thank Satoshi Endo, Kai-Uwe Goss, Jon Arnot, James Armitage, and Frank Gobas for insightful discussions, and James Franklin, Eleonora Simonini, Heike Laue, and Aaron Redman for meticulous review of an earlier version of the manuscript. We acknowledge funding from project ECO-41 “Improved characterization of partitioning and biotransformation for screening organic compounds for the potential to bioaccumulate in air-breathing species” of the Long-range Research Initiative of the European Chemical Industry Council and a MITACS Elevate postdoctoral fellowship to Alessandro Sangion.

**CONFLICT OF INTEREST**

The authors declare that there are no conflicts of interest.
DATA AVAILABILITY STATEMENT
The final version of the data set of 1158 chemicals and the spreadsheet with the application of the proposed assessment procedure is made available as a Microsoft Excel sheet.

SUPPORTING INFORMATION
Three tables with additional information on the proposed procedure and on the chemicals identified as potentially bioaccumulative.

ORCID
Frank Wania http://orcid.org/0000-0003-3836-0901

REFERENCES
ACIDPercepta. (2021). Version 2019.2.1, Advanced Chemistry Development, Inc. www.acdlabs.com
Armitage, J. M., Erickson, R. J., Luckenbach, T., Ng, C. A., Prosser, R. S., Arnot, J. A., Schirmer, K., & Nichols, J. W. (2017). Assessing the bioaccumulation potential of ionizable organic compounds: Current knowledge and research priorities. Environmental Toxicology and Chemistry, 36(4), 882–897.
Armitage, J. M., & Gobas, F. A. P. C. (2007). A terrestrial food-chain bioaccumulation model for POPs. Environmental Science and Technology, 41(11), 4019–4025.
Arnot, J. A., Brown, T. N., & Wania, F. (2014). Estimating screening-level organic chemical half-lives in humans. Environmental Science and Technology, 48(1), 723–730.
Baskaran, S., Lei, Y. D., & Wania, F. (2021a). Reliable prediction of the octanol-air partition ratio. Environmental Toxicology and Chemistry, 40(11), 3166–3180. https://doi.org/10.1002/etc.5201
Baskaran, S., Lei, Y. D., & Wania, F. (2021b). A database of experimentally derived and estimated octanol-air partition coefficients (KOA). Journal of Physical and Chemical Reference Data, 50, 043101. https://aip.scitation.org/doi/10.1063/5.0059652
Beek, B., Böhlking, S., Bruckmann, U., Franke, C., Jönckhe, U., & Studinger, G. (2000). The assessment of bioaccumulation. In B. Beek (Ed.), The handbook of environmental chemistry, Vol. 2 Part J Bioaccumulation (pp. 235–276). Springer-Verlag. https://doi.org/10.1007/10503050_4
Brown, T. N. (2014). Predicting hexadecane-air equilibrium partition coefficients (L) using a group contribution approach constructed from high quality data. SAR and QSAR in Environmental Research, 25(1), 51–71.
Brown, T. N., Arnot, J. A., & Wania, F. (2012). Iterative fragment selection: A group contribution approach to predicting fish bioaccumulation half-lives. Environmental Science and Technology, 46 (15), 8253–8260.
Cappelli, C. I., Benfenati, E., & Cester, J. (2015). Evaluation of QSAR models for predicting the partition coefficient (log P) of chemicals under the REACH regulation. Environmental Research, 143A, 26–32.
Chiu, C. T., Freed, V. H., Schmedding, D. W., & Kohnert, R. L. (1977). Partition coefficient and bioaccumulation of selected organic chemicals. Environmental Science and Technology, 1(5), 475–478. https://doi.org/10.1021/es60128a001
Czub, G., & McLachlan, M. S. (2004). Bioaccumulation potential of persistent organic chemicals in humans. Environmental Science and Technology, 38(8), 2406–2412.
Endo, S., Bauerfeind, J., & Goss, K. U. (2012). Partitioning of neutral organic compounds to structural proteins. Environmental Science and Technology, 46(22), 12697–12703.
Endo, S., & Goss, K. U. (2013). General model for estimating partition coefficients to organisms and their tissues using the biological compositions and polyparameter linear free energy relationships. Environmental Science and Technology, 47(12), 6630–6639.
Endo, S., Escher, B. I., & Goss, K. U. (2011). Capacities of membrane lipids to accumulate neutral organic chemicals. Environmental Science and Technology, 45(14), 5912–5921.
Endo, S., & Goss, K. U. (2014). Applications of polyparameter linear free energy relationships in environmental chemistry. Environmental Science and Technology, 48(21), 12477–12491.
European Chemicals Agency (ECHA). (2017a). Guidance on information requirements and chemical safety assessment. Chapter R.7c: Endpoint specific guidance. Version 3.0. Retrieved March 08, 2019, from: https://echa.europa.eu/documents/10162/17224/information_requirements_r7c_en.pdf/fe2e23a98-ad2b-4573-b450-c0dafa798b5e?u=1498476107907
European Chemicals Agency (ECHA). (2017b). Guidance on information requirements and chemical safety assessment. Chapter R.11: PBT/vPvB assessment. Version 3.0. Retrieved March 08, 2019, from: https://echa.europa.eu/documents/10162/13632/information_requirements_r11_en.pdf/88ce23f-a56a-46d2-ac8b-92eef19e54f
Franke, C., Studinger, G., Berger, G., Böhlking, S., Bruckmann, U., Cohors-Fresenberg, D., & Jönckhe, U. (1994). The assessment of bioaccumulation. Chemosphere, 29(7), 1501–1514.
Giesler, A., Endo, S., & Goss, K. U. (2012). Partitioning of organic chemicals to storage lipids: elucidating the dependence on fatty acid composition and temperature. Environmental Science and Technology, 46(17), 9519–9524.
Gobas, F. A. P. C., de Wolf, W., Burkhard, L. P., Verbruggen, E., & Plotske, K. (2009). Revisiting bioaccumulation criteria for POPs and PBT assessments. Integrated Environmental Assessment and Management, 5(4), 624–637.
Goss, K. U., Brown, T. N., & Endo, S. (2013). Elimination half-life as a metric for the bioaccumulation potential of chemicals in aquatic and terrestrial food chains. Environmental Toxicology and Chemistry, 32(7), 1663–1671.
Goss, K. U., & Schwarzenbach, R. P. (1999). Empirical prediction of heats of vaporization and heats of adsorption of organic compounds. Environmental Science and Technology, 33(19), 3390–3393. https://doi.org/10.1021/es980812j
Gramatica, P., Chirico, N., Papa, E., Cassani, S., & Kovalar, S. (2013). QSARINS: A new software for the development, analysis, and validation of QSAR MLR models. Journal of Computational Chemistry, 34(24), 2121–2132.
Kelly, B. C., & Gobas, F. A. P. C. (2003). An Arctic terrestrial food-chain model for persistent organic pollutants. Environmental Science and Technology, 37(13), 2966–2974. https://doi.org/10.1021/es021035x
Kelly, B. C., Gobas, F. A. P. C., & Amort, J. A. (2003). Quantitative structure activity relationship for predicting the bioaccumulation of POPs in terrestrial food webs: QSAR & Combinatorial Science, 22(3), 329–336.
Kelly, B. C., Ikonomou, M. G., Blair, J. D., Marin, A. E., & Gobas, F. A. P. C. (2007). Food web-specific biomagnification of persistent organic pollutants. Science, 317(5835), 326–329. https://doi.org/10.1126/science.1138275
Kiddie, M. A., Kaye, C. M., & Turner, P. (1972). The influence of urinary pH on the elimination of mexiletine. British Journal of Clinical Pharmacology, 1(3), 229–232. https://doi.org/10.1111/j.1365-2125.1974.tb00241.x
Mackay, D. (1982). Correlation of bioconcentration factors. Environmental Science and Technology, 17(10), 1365–1369.
Mansouri, K., Grulke, C. M., Judson, R. S., & Williams, A. J. (2018). OPERA models for predicting physicochemical properties and environmental fate endpoints. Journal of Cheminformatics, 10, 10.
McLachlan, M. S., Czub, G., MacLeod, M., & Amort, J. A. (2011). Bioaccumulation of organic contaminants in humans: A multimedia perspective and the importance of biotransformation. Environmental Science and Technology, 45(1), 197–202.
Morris, A. D., Muir, D. C. G., Solomon, K. R., Lether, R. J., McKinney, M. A., Fisk, A. T., McMeans, B. C., Tomy, G. T., Teixeira, C., Wang, X., & Duric, M. (2016). Current-use pesticides in seawater and their bioaccumulation in polar bear–ringed seal food chains of the Canadian Arctic. Environmental Toxicology and Chemistry, 35(7), 1695–1707. https://doi.org/10.1002/etc.3427
O’Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. Journal of Cheminformatics, 3, 33.
Papa, E., Amot, J. A., Sangion, A., & Gramatica, P. (2017). In silico approaches for the prediction of in vivo biotransformation rates. In K. Roy (Ed.), Advances in QSAR modeling. Challenges and advances in computational chemistry and physics (Vol. 24, pp. 425–451). Springer. https://doi.org/10.1007/978-3-319-56850-8_11

Papa, E., Sangion, A., Amot, J. A., & Gramatica, P. (2018). Development of human biotransformation QSARs and application for PBT assessment refinement. Food and Chemical Toxicology, 112, 535–543.

Platford, R. F. (1983). The octanol-water partitioning of some hydrophobic and hydrophilic compounds. Chemosphere, 12(7–8), 1107–1111. https://doi.org/10.1016/0045-6535(83)90264-3

Platts, J. A., Abraham, M. H., Butina, D., & Hersey, A. (2000). Estimation of molecular linear free energy relationship descriptors by a group contribution approach. 2. Prediction of partition coefficients. Journal of Chemical Information and Computer Sciences, 40(1), 71–80.

Platts, J. A., Butina, D., Abraham, M. H., & Hersey, A. (1999). Estimation of molecular linear free energy relation descriptors using a group contribution approach. Journal of Chemical Information and Computer Sciences, 39(5), 835–845.

Rozman, K. (1985). Intestinal excretion of toxic substances. In P. L. Chambers, E. Cholnoky, & C. M. Chambers (Eds.), Receptors and other targets for toxic substances. Archives of toxicology (supplement) (Vol. 8, pp. 87–93). Springer.

Sazonovas, A., Japertas, P., & Didziapetris, R. (2010). Estimation of reliability of predictions and model applicability domain evaluation in the analysis of acute toxicity (LD₅₀). SAR & QSAR in Environmental Research, 21(1–2), 127–148.

Stenzel, A., Goss, K. U., & Endo, S. (2013). Experimental determination of polyparameter linear free energy relationship (ppLFER) substance descriptors for pesticides and other contaminants: New measurements and recommendations. Environmental Science and Technology, 47(24), 14204–14214. https://doi.org/10.1021/es404150e

Szöcs, E., Stirling, T., Scott, E. R., Scharmüller, A., & Schäfer, R. B. (2020). Webchem: An R package to retrieve chemical information from the web. Journal of Statistical Software, 93(13), 1–17. https://doi.org/10.18637/jss.0093.i13

Tülp, H. C., Goss, K. U., Schwarzenbach, R. P., & Fenner, K. (2008). Experimental determination of LSER parameters for a set of 76 diverse pesticides and pharmaceuticals. Environmental Science and Technology, 42(6), 2034–2040.

Ulrich, N., Endo, S., Brown, T. N., Watanabe, N., Bronner, G., Abraham, M. H., & Goss, K. U. (2017). UFZ-LSER database v 3.2.1 [Internet]. Leipzig, Germany, Helmholtz Centre for Environmental Research-UFZ. Retrieved January 02, 2021, from: http://www.ufz.de/lsrd

United Nations Environmental Programme (UNEP). (2017). Stockholm convention on persistent organic pollutants (POPs), Texts and Annexes. Retrieved February 17, 2021, from: http://chm.pops.int/Portals0/download.aspx?id=UNEP-POPS-COP-CONVTEXT-2017.English.pdf

US EPA. (2021). Estimation programs interface suite™ for Microsoft® Windows, v 4.11. United States Environmental Protection Agency, Washington, DC, USA.

Vorkamp, K., & Rigét, F. F. (2014). A review of new and current-use contaminants in the Arctic environment: Evidence of long-range transport and indications of bioaccumulation. Chemosphere, 111, 379–395.

Watkins III, J. B., & Klaasen, C. D. (1986). Xenobiotic biotransformation in livestock: Comparison to other species commonly used in toxicity testing. Journal of Animal Science, 63(3), 933–942. https://doi.org/10.2527/jas1986.633933x

Williams, R. T. (1974). Inter-species variations in the metabolism of xenobiotics. Biochemical Society Transactions, 2(13), 359–377. https://doi.org/10.1042/bst0020359

Zhang, X., Brown, T. N., Wania, F., Heimstad, E. S., & Goss, K. U. (2010). Assessment of chemical screening outcomes based on different partitioning property estimation methods. Environment International, 36(6), 514–520.