Health risk assessment with multiple reference indices

Pascal Petit, Dominique J. Bicout

HAL Id: hal-03364381

https://hal.univ-grenoble-alpes.fr/hal-03364381v1

Submitted on 16 Oct 2023

1 **Health risk assessment with multiple reference indices**

2

3 Pascal Petit[a,*], Dominique J. Bicout[a]

4 [a] Univ. Grenoble Alpes, CNRS, Grenoble INP, VetAgro Sup, TIMC, 38000 Grenoble, France

5

6 * Corresponding authors at: Laboratoire TIMC-IMAG, Equipe EPSP, Faculté de Médecine de Grenoble,

7 Domaine de La Merci, 38706 La Tronche Cedex, France.

8 Tel: +33(0) 4 76 63 75 06

9 E-mail address: pascal.petit@univ-grenoble-alpes.fr

10   **ABSTRACT**

11   *Background*: Conducting a risk assessment can be challenging, especially when dealing with several

12   reference indices, which could lead to conflicting conclusions between studies. The common approach

13   is to use a reference index from a single source based on the risk assessor's preference.

14   *Objectives*: To propose an approach for constructing a multi-reference index-based aggregated risk

15   estimate using mathematical objectivity to reflect all of the available information.

16   *Methods*: The aggregated risk estimate based on multiple reference indices (AREMRI) results from

17   the weighted linear combination of risk distributions that were obtained with each reference index

18   available. The weights were calculated using the degree of agreement among the reference index-

19   based risk distributions. The approach is illustrated through three practical cases of benzene inhalation

20   cancer risk assessment using inhalation unit risks (IURs) from six different regulatory agencies.

21   *Results*: The degrees of agreement between the reference index-based risk distribution, obtained with

22   the six IURs, ranged from 0.7 to 92%. The highest weights were attributed to reference index-based

23   risk distributions that had the highest degree of agreement with the maximum number of other

24   reference index-based risk distributions. Regardless of the practical case considered, the AREMRI risk

25   distribution resulted in the third highest risk compared to the six single risk distributions.

26   *Conclusion*: Our approach can be useful in the presence of several reference indices by providing

27   useful insights, consistency and direct comparisons between studies to support better-informed risk

28   assessment and management decisions. This approach can shed some light on some of the

29   uncertainties associated with the toxicological reference values in a risk assessment associated with the

30   toxicological reference values. If the uncertainty is large, more detailed evaluation of the toxicological

31   reference values would be needed.

32

2

1. Introduction

Risk assessments are paramount in ensuring the protection of human health against environmental pollutants (Rotter et al. 2018). Performing a risk assessment requires the use of toxicity reference indices. Reference indices are derived using rigorous methodological weights of evidence approaches by regulatory agencies based on available data generated in experimental toxicology and exposure assessment. For a given absorption pathway (oral, respiratory, cutaneous), dates when the index values were derived, the method of derivation, and knowledge or assumptions about substances' mechanisms of action are often different from one organization/agency/study to another (Petit et al. 2020). These differences arise mainly from the use of different data sources/key studies (e.g., the results from animal studies or epidemiological studies), extrapolation methods, correction factors, and commonly accepted assumptions and uncertainties (ANSES 2015; Beck et al. 2016; Melnick et al. 2008; US EPA 2014). Even when similar studies are considered for the same critical health effects and absorption pathway, organizations/agencies/studies can reach different conclusions due to different policies and/or expert judgments (Beck et al. 2016). Depending on the reference index chosen, the results of the risk assessment can thus differ and lead to different conclusions and decisions, which may have broad and profound implications regarding the prevention of adverse effects and the protection of human health and the environment.

The double question that arises when carrying out a risk assessment is knowing which reference index to use and how to take into account the information provided by the other indices. Indeed, there are no criteria or benchmark/gold standard approaches guiding the choice of an index, especially since, by definition, all reference indices available from health/regulatory agencies are legitimate. In practice, we tend to use either an index from our own home country agency (when available) or an index from an agency to which we subjectively attach a certain importance. Thus, in this paper, we explored and suggest an alternative approach, which would be to construct and use an aggregated risk estimate based on multiple reference indices (AREMRI) created from mathematical objectivity. Such an AREMRI would allow for direct comparisons among studies, promote transparency and ensure a consistency of analysis during risk assessment. The aim here is not to judge how researchers and/or

62 regulatory agencies derive reference indices or to criticize existing reference indices but rather to use

63 them as is, starting from the fact that each of the indices carries a part of the truth.

64 The remainder of this paper is organized as follows. The general approach for constructing an

65 AREMRI risk distribution (uncertain risks) is presented in Section 2. Section 3 is devoted to an

66 illustrative application of the developed approach to inhalation cancer risks. Finally, the last section,

67 Section 4, discusses the proposed approach and considers the practical implications and

68 implementation of the developed analysis in an interactive web application.

69

70 **2. Methods**

71 *2.1 Aggregated risk distribution construction based on multiple reference indices (AREMRI)*

72 Consider a discrete ensemble of n indicators of uncertain risk r, each based on a reference index from

73 a different source (e.g., organization/agency/study), and each being described by a distribution

74 function normalized to one, $R_i(r)$, with $i = 1, 2, \dots, n$. It is assumed that all reference indices from n

75 sources are legitimate and relevant (developed for the same purpose) and that there is no objective

76 and/or impartial reason to choose one over another for the risk assessment. As a first step, the

77 ensemble of n indicators can be characterized by the matrix of degrees of agreement, $\boldsymbol{A}$, given by (Eq.

78 1):

$$
79 \qquad \boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{pmatrix} \qquad (1)
$$

80 where $a_{ij} = a_{ji}$ is the degree of agreement between two reference index-based distributions

81 (indicators) i and j such that $0 \leq a_{ij} = a_{ji} \leq 1$ and $a_{ii} = 1$. $a_{ij} = 0$ refers to no agreement, while

82 $a_{ij} = 1$ indicates total agreement between the two reference index-based risk distributions. A

83 reference index-based distribution is defined as a risk distribution resulting from the use of a

84 given/source-specific reference index (e.g., oral slope factor from US EPA). Mathematically, $a_{ij} =$

85 $\text{overlap}[R_i, R_j]$ is obtained by calculating the overlap between two (normalized to one) distribution

86 functions $R_i(r)$ and $R_j(r)$; a general method for calculating the overlap (or overlapping area) between

87 two functions is described elsewhere (Petit et al. 2020).

88 The content of $\boldsymbol{A}$ is very informative to compare reference index-based risk distribution with each

89 other and also to guide the development of a pragmatic choice of one reference index over the others,

90 in possible combination with other techniques such as stochastic dominance or expected utility

91 (Verteramo Chiu et al. 2020). Whichever approach is used, the result would be to choose the indicator

92 (i.e., reference index) from a single source (e.g., regulatory agency/study). Consequently, we are rather

93 in favor of constructing a composite indicator resulting from the aggregation of all indicators to reflect

94 all of the available information. Our aim is to develop such an aggregation based on the content of $\boldsymbol{A}$

95 and, in this case, there will be no aggregation when $\boldsymbol{A} = \boldsymbol{1}$ (identity matrix), i.e., when none of the

96 reference index-based distributions would agree with the others.

97 To this end, we propose to construct an aggregated risk distribution based on multiple reference

98 indices (AREMRI), $R_a(r)$, as a weighted linear combination of reference index-based risk

99 distributions $R_i(r)$, as follows (Eq. 2):

100
$$R_a(r) = \sum_{i=1}^{n} \omega_i R_i(r) \tag{2}$$

101 where ω_i ($0 \le \omega_i \le 1$ such that $\sum_{i=1}^{n} \omega_i = 1$) is a contributing weight of the risk distribution

102 generated with the reference index "i" and $R_i(r)$ is a risk distribution based on a single given

103 reference index. Next, to determine the ω_i coefficients and obtain an objective aggregated distribution,

104 we require that $R_a(r)$ has a maximum agreement with all risk distributions generated with all

105 reference indices available $R_i(r)$, i.e., to maximize total agreement θ between $R_a(r)$ and all $R_i(r)$.

106 The total agreement θ is related to the agreement matrix $\boldsymbol{A}$ as (Eq. 3):

107
$$\theta = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}^T \underbrace{\begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{pmatrix}}_{\boldsymbol{A}} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{pmatrix} \tag{3}$$

108 The agreement matrix $\boldsymbol{A}$ is symmetric with the trace, $\mathrm{Tr}(\boldsymbol{A}) = \sum_{k=1}^{n} \lambda_k = n$, where λ_k are the

109 eigenvalues associated with (normalized to one) eigenvectors $\boldsymbol{v}_k$ of matrix $\boldsymbol{A}$ such that $0 \le \lambda_1 < \lambda_2 <$

110    $\cdots < \lambda_n \leq n$. When there is no agreement among all reference index-based risk distributions $R_i(r)$, $\mathbf{A}$

111   reduces to the identity matrix with $a_{ij} = 0$ for $i \neq j$, and the total agreement is always $\theta = 1$

112   regardless of the weights ω_i. In this case, no aggregation based on the agreement matrix $\mathbf{A}$ is possible

113   as there is no objective way of choosing one reference index over another. In the opposite situation

114   when all $a_{ij} = 1$, i.e., perfect agreement among all reference index-based risk distributions, the total

115   agreement is maximum, $\theta = n$. We will show below that, in this case, $\omega_i = 1/n$ (arithmetic mean).

116   Finally, when $0 < a_{ij} < 1$, the maximum of the total agreement, θ, is obtained by setting the ω_i

117   coefficients equal to eigenvector $\boldsymbol{v}_n$ corresponding to the largest eigenvalue λ_n as shown in (Eq. 4):

118    $$\begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{pmatrix} = \frac{1}{[\sum_{i=1}^{n} v_{n,i}]} \begin{pmatrix} v_{n,1} \\ v_{n,2} \\ \vdots \\ v_{n,n} \end{pmatrix} \Longrightarrow \theta = \lambda_n \frac{1}{[\sum_{i=1}^{n} v_{n,i}]} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}^T \begin{pmatrix} v_{n,1} \\ v_{n,2} \\ \vdots \\ v_{n,n} \end{pmatrix} = \lambda_n \qquad (4)$$

119   This determines the weights of $R_a(r)$ in Eq. 2. It is paramount to ensure that all $v_{n,i}$ have the same

120   sign. Fig. 1 presents a schematic summary of the different steps of the construction of the aggregated

121   risk distribution based on multiple reference indices.
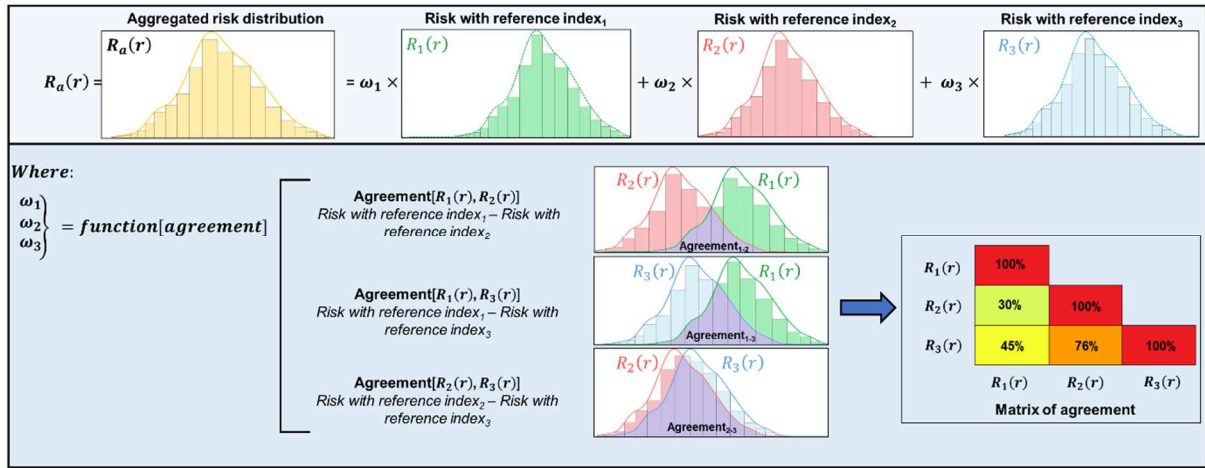
123   **Fig. 1**: Scheme of the general approach to construct an aggregated risk distribution based on multiple

124   reference indices

125   ω_i: contributing weight of agency "i" with $0 \leq \omega_i \leq 1$ such that $\sum_{i=1}^{n} \omega_i = 1$; $R_a(r)$: aggregated risk

126   distribution function based on multiple reference indices; $R_i(r)$: risk distribution function generated

127   with reference index "i".

129 In case of perfect agreement among all reference index-based risk distributions (indicators), i.e., all

130 $a_{ij} = 1$, the highest eigenvalue of the agreement matrix **A** is $\lambda_n = n$, associated with the eigenvector

131 $v_n^T = (1, 1, \cdots, 1)^T$; this leads to $\omega_1 = \omega_2 = \cdots = \omega_n = 1/n$.

132 When dealing with only two reference indices (indicators), i.e., $n = 2$ and $a_{12} = a_{21} \neq 0$, the highest

133 eigenvalue of **A** is $\lambda_2 = 1 + a_{12}$, associated with the eigenvector $v_2^T = (1, 1)^T$. In this case, the

134 weights are $\omega_1 = \omega_2 = 1/2$, independent of a_{12}, provided that $a_{12} \neq 0$.

135 To gain more understanding of these weights, consider an example of $n = 3$ for which weights can be

136 calculated explicitly as:

$$137 \qquad A = \begin{pmatrix} 1 & a_{12} & 0 \\ a_{21} & 1 & a_{23} \\ 0 & a_{32} & 1 \end{pmatrix} \Longrightarrow \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \frac{1}{\left[a_{12} + \sqrt{a_{12}^2 + a_{23}^2} + a_{23}\right]} \begin{pmatrix} a_{12} \\ \sqrt{a_{12}^2 + a_{23}^2} \\ a_{23} \end{pmatrix} (5)$$

138 where $a_{12} = a_{21} < 1$ and $a_{23} = a_{32} < 1$ such that $a_{13} = a_{31} = 0$, and the highest eigenvalue of **A** is

139 $\lambda_3 = 1 + \sqrt{a_{12}^2 + a_{23}^2}$. Clearly, Eq. 5 shows how the weights are "proportional or a function" of the

140 respective agreements. For instance, setting $a_{23} = 0$ leads to $\omega_3 = 0$, i.e., the contribution of a

141 reference index-based risk distribution (indicator) is zero when the reference index-based risk

142 distribution does not agree with any other. In addition, the weight of a reference index-based risk

143 distribution (indicator) increases with its total agreement with the other reference index-based risk

144 distributions (indicators). Indeed, we have $\omega_3 < \omega_1 < \omega_2$ when $a_{23} < a_{12}$, and $\omega_1 < \omega_3 < \omega_2$ for

145 $a_{23} > a_{12}$. These observations remain valid for $n > 3$.

146

147 *2.2 Application to inhalation cancer risks*

148

149 To illustrate the usefulness of using an aggregated risk distribution based on multiple reference

150 indices, we consider three practical cases/scenarios of risk assessment of benzene exposure where each

151 exposure scenario is characterized by a distribution of benzene concentrations assumed constant over

152 time and following a lognormal distribution with the geometric mean (GM) and geometric standard

153 deviation (GSD) parameters given in Table 1 and an exposure duration of 4 hours a day, 120 days a

154   year for 50 years. Regardless of the reference index considered, the concentration distributions used

155   for a given scenario were exactly the same (same geometric mean and geometric standard deviation).

156   The concentration distribution parameters differed from one practical case to another with a GM set to

157   253, 253 and 58.4 µg/m^3 and a GSD set to 2.80, 1.03 and 2.80 for cases 1, 2 and 3, respectively. For

158   all scenarios, both the GM and GSD were set at pragmatic random values generated with R software

159   4.0.5® (R Core Team, Vienna, Austria) for Windows 10©.

160   We used the *InCaRisk* app (Petit et al. 2020) with 1,000 simulations runs to generate all of the

161   reference index-based risk distributions $R_i(r)$. Briefly, the *InCaRisk* app is a free user-friendly and

162   interactive web application that allows for cancer risk estimation following inhalation exposure

163   (https://exporisk-timc.imag.fr/InCaRisk/). A total of 305 substances/chemicals with reference values

164   from up to eight agencies are available. *InCaRisk* requires no programming knowledge to use. This

165   app also incorporates a variety of features and options to make it easy to use, with the possibility of

166   configuring the exposure settings (exposure scenario), the concentration distribution settings (choice

167   of the type of distribution, distribution parameters, uploading an exposure concentration file) and to

168   export the results as a report. This app offers the chance to have an immediate glimpse into the results

169   and to see how the results change according to different setting configurations with interactive easy-to-

170   read graphs, which are downloadable.

171   The reference index-based risk distributions $R_i(r)$ were generated for six sanitary agencies for which

172   the tumor site (leukemia) and tumor type (hematologic/immune) used in the inhalation unit risk (IUR)

173   establishment were the same: ANSES (French Agency for Food, Environmental and Occupational

174   Health & Safety), Health Canada, OEHHA (California Office of Environmental Health Hazard

175   Assessment), RIVM (National Dutch Institute for Public Health and the Environment), US EPA (US

176   Environmental Protection Agency) and WHO (World Health Organization). The IURs for benzene set

177   by these agencies regarding the risk of leukemia for lifetime exposure were 2.6x10^{-5} (ANSES 2019),

178   3.3x10^{-6} (Health Canada 2010), 2.9x10^{-5} (OEHHA 2019), 5x10^{-6} (RIVM 2001), 7.8x10^{-6} (US EPA

179   2018) and 6x10^{-6} m^3/µg (WHO 2000).

180

181   Table 1 here (see Table 1 after the Reference section).

3. Results

For a given practical case, each IUR-based risk distribution resulted in different cancer estimates even though the exposure concentration distributions were the same. These differences came solely from the IUR values that differed from one agency to another (Table 1). Indeed, each IUR-based risk distribution had the same shape (height and width) because of similar concentration parameters (Table 1) but was shifted from one another (different mean) by a factor that depended on the ratio $IUR_{agency\ i}/IUR_{agency\ j}$ (Fig. 2C, 3C and 4C).

As seen in the matrix of agreement between sources (agencies) (Fig. 2A, 3A and 4A), the degrees of agreement were heterogeneous between IUR-based risk distributions, regardless of the practical case considered. The results were similar for cases 1 and 3. Degrees of agreement were lower for the second scenario than for the first and third scenarios because the GSDs were the lowest (narrower distributions) for case 2. For all scenarios, the highest degrees of agreement, 92% for cases 1 and 3 and 81% for case 2, were found between IUR-based risk distributions from RIVM and WHO, while the lowest degrees of agreement were found between IUR-based risk distributions from Health Canada and OEHHA (30% for cases 1 and 3, 0.7% for case 2). Regardless of the practical case considered, IUR-based risk distributions from OEHHA and ANSES had a high degree of agreement with each other (84% for cases 1 & 3, 80% for case 2), a medium degree of agreement with all other IUR-based risk distributions for cases 1 & 3 (30 to 56%), and a low degree of agreement with all other IUR-based risk distributions for case 2 (0.7 to 15%). IUR-based risk distributions from the WHO, US EPA, Health Canada and RIVM had a medium to high degree of agreement with each other (66 to 92% for cases 1 & 3 and 30 to 81% for case 2).

The weights were similar for cases 1 and 3, ranging from 0.136 to 0.189, and they ranged from 0.053 to 0.247 for case 2 (Fig. 2B, 3B and 4B). The highest weights were attributed to IUR-based risk distributions that had the highest degree of agreement with the maximum number of other IUR-based risk distributions. Conversely, IUR-based risk distributions that shared a high degree of agreement with the smallest number of other IUR-based risk distributions had the smallest weights.

9

209

210 For the first and third scenarios, the aggregated multi IUR-based risk distribution (AREMRI) followed

211 a lognormal distribution of parameters GM = 1.65×10^{-4} and GSD = 2.8 for case 1 and GM = 3.82×10^{-5}

212 and GSD = 2.84 for case 3. The AREMRI resulted in the third highest risk, with a leukemia risk of

213 6.56×10^{-4} [IC 95%: 2.31×10^{-5} – 1.32×10^{-3}] for case 1 and 1.53×10^{-4} [IC 95%: 5.30×10^{-6} – 3.07×10^{-4}] for

214 case 3. For both scenarios, these leukemia risks were 2.5 and 2.2 times lower than the cancer risk

215 estimated with the reference indices (IURs) of OEHHA and ANSES, respectively (Fig. 2C and 4C).

216 Conversely, the aggregated multi-IUR-based risks were 1.5, 1.9, 2.3 and 3.5 times higher than the

217 leukemia risk estimated with the reference indices from the US EPA, WHO, RIVM and Health

218 Canada, respectively. The concentration GSDs used for cases 1 & 3 were the same (GSD = 2.80),

219 which explains why the weights and degrees of agreement were similar for both scenarios. The only

220 difference between case 1 and case 3 came from the estimated leukemia risk, which was 4.3 times

221 lower for case 3 than for case 1. This difference corresponded to the ratio of their concentration GM

222 (253 µg/m^3 for case 1 *vs.* 58.4 µg/m^3 for case 3). In other words, for a given concentration GSD,

223 changing the concentration GM will shift the risk estimate toward higher values when the GM is

224 increased and toward lower values when the GM is decreased, with little to no change at all of the

225 degrees of agreement or the weighted values.

226

227 For the second scenario, the aggregated multi-IUR-based risk distribution (AREMRI) followed a

228 lognormal distribution of parameters GM = 1.08×10^{-4} and GSD = 1.52. The AREMRI resulted in the

229 third highest risk, with a leukemia risk of 1.83×10^{-4} [IC 95%: 4.77×10^{-5} – 2.55×10^{-4}], which was 3.5

230 and 3.1 times lower than the cancer risk estimated with the reference indices (IURs) of OEHHA and

231 ANSES, respectively (Fig. 3C). Conversely, the aggregated multi-IUR-based risk was 1.1, 1.4, 1.6 and

232 2.5 times higher than the leukemia risk estimated with the reference indices from the US EPA, WHO,

233 RIVM and Health Canada, respectively. For this scenario, the GM concentration was the same as that

234 in the first scenario (GM = 253 µg/m^3), but the GSD concentration was lower (1.03 for case 2 *vs.* 2.80

235 for case 1). A lower GSD concentration for case 2 resulted in lower degrees of agreement (Fig. 3A)

236 and better risk estimation (narrower 95% confidence interval) than for case 1. Indeed, there was a 5.3

237 factor between the lower and upper bounds of the 95% confidence interval of the risk distribution for

238 case 2 compared to a factor of 57 for case 1.

239
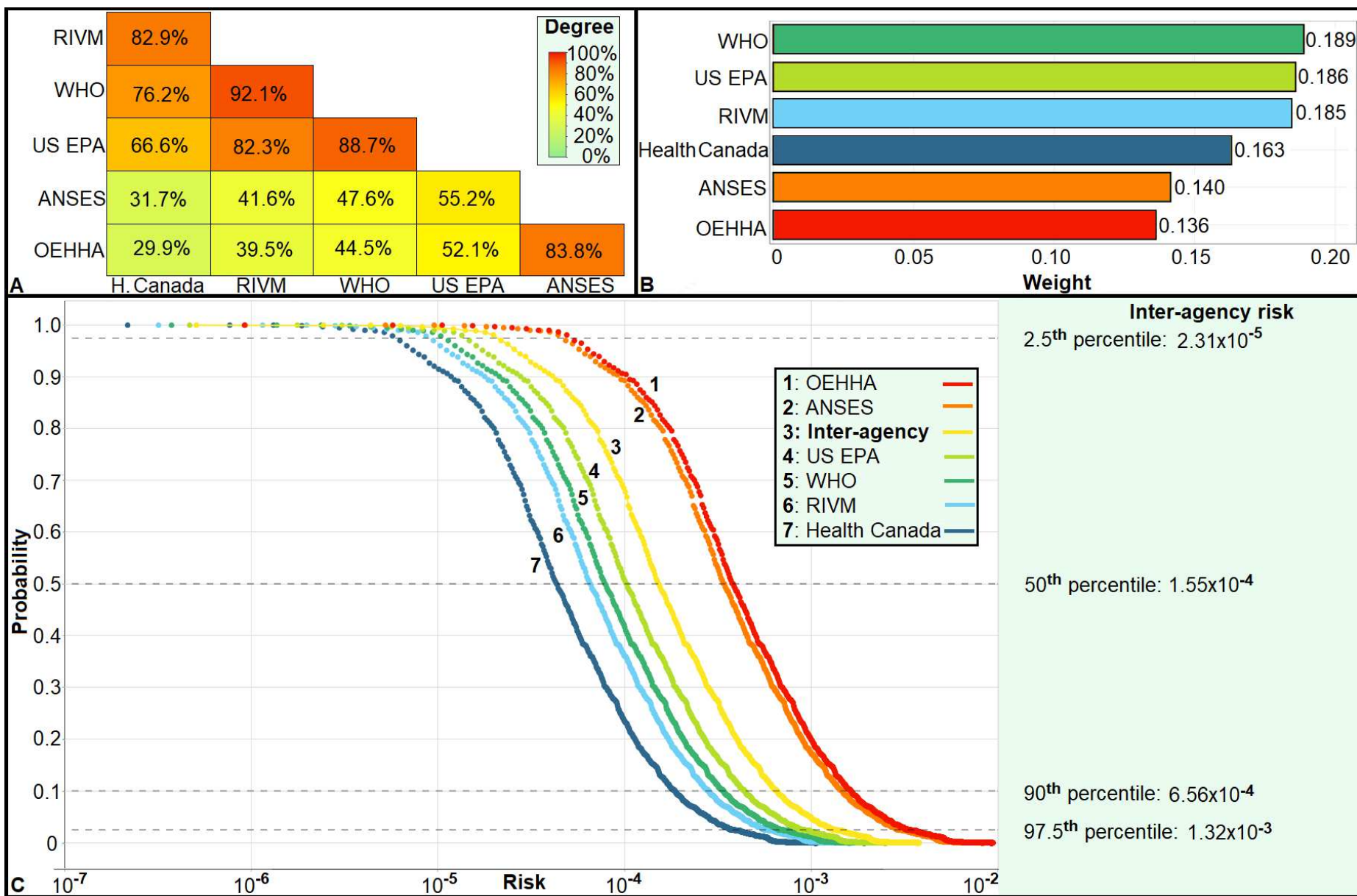
240   **Fig. 2**: Side-by-side comparison of single IUR-based risk distributions for the construction of the aggregated multi-IUR-based risk distribution – case 1

241   A: Matrix of degrees of agreement between a single IUR-based risk distribution, B: Weights attributed to the IUR-based risk distribution for the aggregated

242   multi-IUR-based risk distribution construction, C: IUR-based risk distributions.

243   ANSES: French Agency for Food, Environmental and Occupational Health & Safety, H. Canada: Health Canada, OEHHA: California Office of

244   Environmental Health Hazard Assessment, RIVM: National Dutch Institute for Public Health and the Environment, US EPA: US Environmental Protection

245   Agency, WHO: World Health Organization. Each agency name refers to the risk distribution generated using its benzene IUR and the exposure scenario

246   considered. The term "Inter-agency" refers here to the aggregated risk estimate based on multiple reference indices (AREMRI).
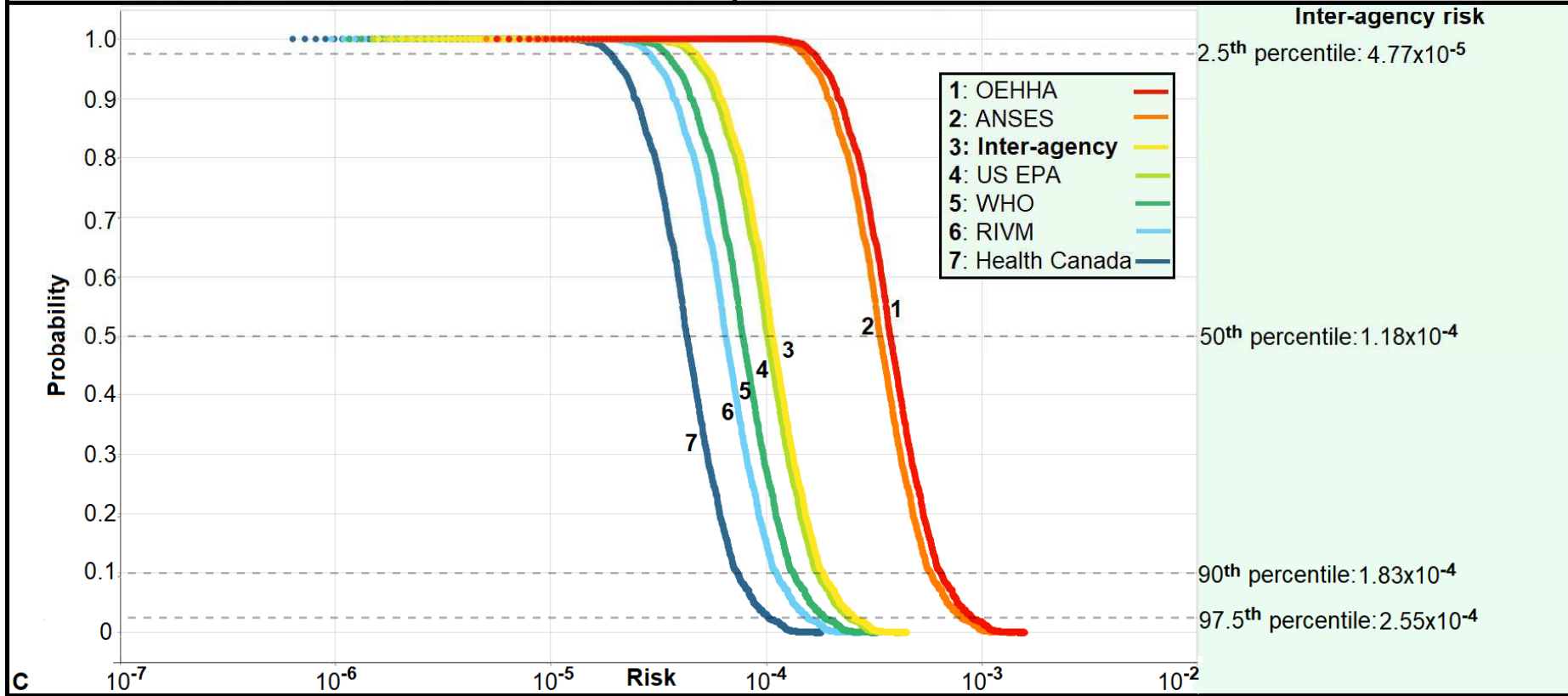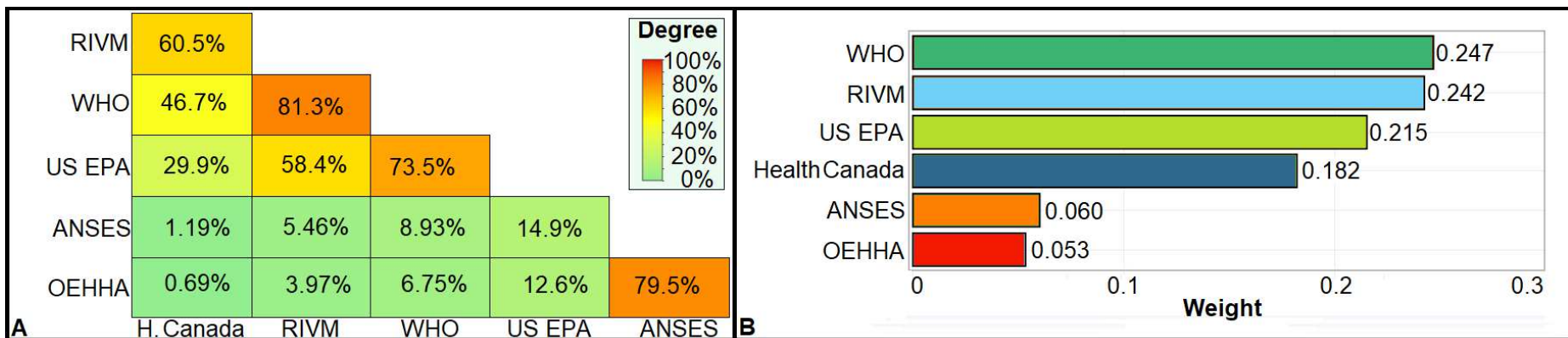
248 **Fig. 3**: Side-by-side comparison of single IUR-based risk distributions for the construction of the aggregated multi-IUR-based risk distribution – case 2

249 A: Matrix of degrees of agreement between a single IUR-based risk distribution, B: Weights attributed to the IUR-based risk distribution for the aggregated

250 multi-IUR-based risk distribution construction, C: IUR-based risk distributions.

251 ANSES: French Agency for Food, Environmental and Occupational Health & Safety, H. Canada: Health Canada, OEHHA: California Office of

252 Environmental Health Hazard Assessment, RIVM: National Dutch Institute for Public Health and the Environment, US EPA: US Environmental Protection

253 Agency, WHO: World Health Organization. Each agency name refers to the risk distribution generated using its benzene IUR and the exposure scenario

254 considered. The term "Inter-agency" refers here to the aggregated risk estimate based on multiple reference indices (AREMRI).
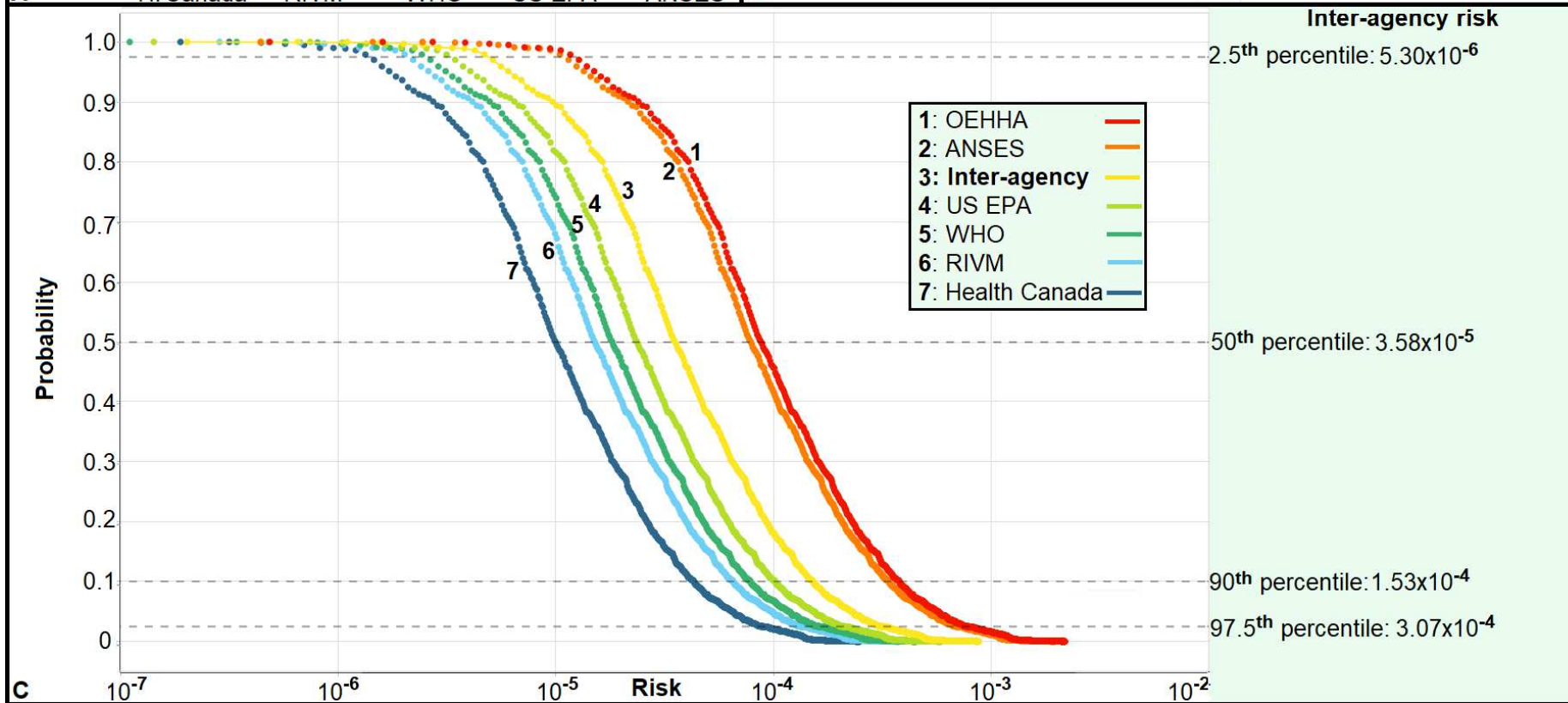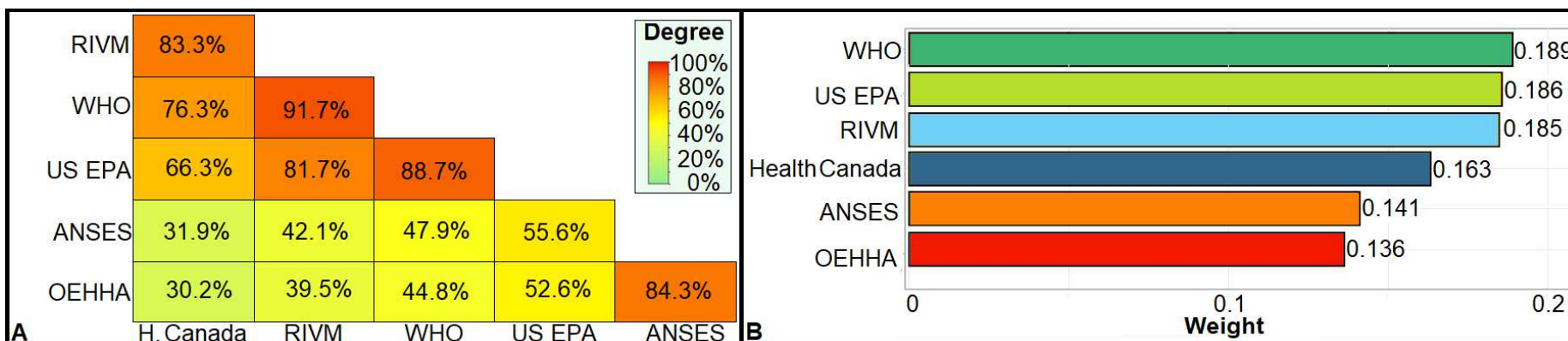
256 **Fig. 4**: Side-by-side comparison of single IUR-based risk distributions for the construction of the aggregated multi-IUR-based risk distribution – case 3

257 A: Matrix of degrees of agreement between a single IUR-based risk distribution, B: Weights attributed to the IUR-based risk distribution for the aggregated

258 multi-IUR-based risk distribution construction, C: IUR-based risk distributions.

259 ANSES: French Agency for Food, Environmental and Occupational Health & Safety, H. Canada: Health Canada, OEHHA: California Office of

260 Environmental Health Hazard Assessment, RIVM: National Dutch Institute for Public Health and the Environment, US EPA: US Environmental Protection

261 Agency, WHO: World Health Organization. Each agency name refers to the risk distribution generated using its benzene IUR and the exposure scenario

262 considered. The term "Inter-agency" refers here to the aggregated risk estimate based on multiple reference indices (AREMRI).

4. Discussion

Performing a risk assessment requires the use of toxicity reference indices. When dealing with several reference indices, risk assessment can be challenging, as there is no objective way to choose one reference index over the others. By definition, all indices available from sanitary/regulatory agencies are legitimate, but in practice, we tend to use an index from a source or agency to which we subjectively attach a certain importance. In this paper, we propose a novel approach consisting of constructing and using an aggregated multi-reference index-based risk estimate constructed only on mathematical objectivity, assuming that the reference indices considered are all relevant and suitable for the risk assessment to be conducted. An R code allowing for the estimation of risks (AREMRI) by inputting multiple reference indices is provided in the Supplemental Material.

The construction of the aggregated risk distribution was based on the weighted linear aggregation (arithmetic mean) of the probability risk distribution generated using a given reference index and using the degree of agreement among the reference index-based risk distributions for the assignment of weight. Other approaches could be used, such as stochastic dominance, which is a form of stochastic ordering that would favor one reference index-based risk distribution over the others based on a specific criterion. For example, using stochastic dominance, it would be possible to favor only the risk distribution resulting from the source or agency with the highest reference index, that is to say the most protective, or with the lowest reference index (weight of 1 while a weight of 0 for all the other reference index-based risk distribution) (Verteramo Chiu et al. 2020). That kind of approach is, however, different from our goal, which was not to judge or provide arguments for choosing a reference index over the others but rather to objectively construct an aggregated risk indicator that takes into account several reference indices from different sources (e.g., sanitary agencies), as all reference indices are legitimate and carry part of the truth.

The proposed approach was illustrated using three practical cases where the exposure concentration distributions were assumed to follow lognormal distributions. Our approach is not limited to lognormal distributions and can work with any type of distribution. Two additional examples are available in the Supplemental Material (Fig. S1 and S2), with exposure concentrations following normal and Poisson distributions, respectively.

291   The approach consisting of constructing and using an aggregated multi-reference index-based risk

292   estimate was illustrated for one kind of stochastic reference value (inhalation cancer unit risk). The

293   presented approach could be used with other types of toxicological reference values, such as oral slope

294   factors, drinking water unit risks or threshold/deterministic values. For some toxicological reference

295   values, the distribution of the underlying toxicity data is available, and the reference value itself can

296   sometimes correspond to a certain point on that distribution (e.g., for BMDL values). In that particular

297   case, these reference values could potentially be set as distributions instead of single values using the

298   distribution of the underlying toxicity data when available. While the presented approach can still be

299   used with reference value distributions, it should be noted that the choice of the reference value

300   distribution (e.g., gamma, lognormal, Poisson, Weibull) and the associated uncertainties (confidence

301   intervals) will both have an important impact on the resulting risk estimate. In particular, in the

302   presence of large confidence intervals (high data dispersion), the degrees of agreement (overlaps

303   between distributions of risks) are so high that all of the reference indices contribute almost equally to

304   the aggregate distribution and the notion of multi-reference indices becomes less relevant. This issue is

305   particularly true with deterministic values (threshold values), for which the choice of the derivation

306   method and, in particular, the choice and use of large safety factors may dominate the differences

307   between the toxicological reference values. One possible solution to address this issue would be to

308   replace traditional noncancer reference doses (threshold values) with probabilistic estimates (Chiu et

309   al. 2018).

310   Regardless of the practical case considered, IUR-based risk distributions from OEHHA and ANSES

311   had a high degree of agreement with each other but a medium to low degree of agreement with other

312   IUR-based risk distributions. The six benzene reference values used in this illustration were derived

313   for the same purpose (lifetime benzene exposure), all based on cohort studies of benzene-exposed

314   workers, but using different key studies and extrapolation/derivation methods among agencies, which

315   may explain the observed differences (ANSES 2019; Health Canada 2010; OEHHA 2019; RIVM

316   2001; US EPA 2018; WHO 2000). OEHHA and ANSES used the same extrapolation method, a

317   relatively risk procedure, but chose two different key studies, Rinsky et al. (1981) for OEHHA and

318   Richardson (2008) for ANSES. RIVM chose the same key study as OEHHA but used a nonthreshold

319 extrapolation method. The WHO used an average relative risk model with Crump (1994) as a key

320 study. The US EPA applied a low-dose linearity utilizing the maximum likelihood estimates method

321 based on six key studies (Crump and Allen 1984; Crump 1994; Paustenbach et al. 1993; Rinsky et al.

322 1981, 1987; US EPA 1998). Health Canada applied a linear quadratic model of the exposure response

323 relationship based on four key studies (Bond et al. 1986; Rinsky et al. 1987; Wong 1987a,b).

324 Differences in policies and/or expert judgments could also have played a role in the observed

325 differences (Beck et al. 2016).

326 The aggregated multi-reference index-based distribution (the degrees of agreement and weights)

327 depends on the number of reference indices included in the aggregation as well as the concentration

328 distribution parameters. The narrower the concentration distributions (lower GSD), the lower the

329 degrees of agreement and the better the risk distribution estimation (narrower confidence interval). If

330 the GSD concentration tends to 1, as in the second scenario, the AREMRI tends to the mean risk value

331 and also tends to approach and equal a single reference index-based risk distribution (e.g., the risk

332 distribution using the reference value from the US EPA for case 2).

333 Regarding the reference indices, we found, for our practical cases, six agencies that set/proposed an

334 IUR for benzene and the risk of leukemia. There are most likely other agencies and studies around the

335 world which we are not aware of that may also have proposed an IUR for benzene and the risk of

336 leukemia for lifetime exposure. Integrating these potential other IURs into the aggregated multi-

337 reference index-based risk distribution construction would change the leukemia risk obtained with the

338 aggregated risk indicator. However, the approach presented in this paper does not remove the

339 requirement of a prior screening of reference values to be included in the integrated approach. This

340 screening should always be made to determine how reference indices were derived, which data they

341 are based on, and their intended use (e.g., assessment of occupational exposure, exposure via

342 foodstuffs and unintentional exposure). Indeed, it is paramount that the risk assessor has some

343 knowledge regarding the reference values used in an assessment and understands the way the values

344 were derived, their associated uncertainties and their relevance and validity for the risk assessment

345 being carried out. If the approach presented in this paper is used to interpret the results of a risk

346 assessment, then only reference values derived for the same purpose(s) should be included in the

347 analysis.

348

349 **5. Conclusions**

350 The three practical cases highlighted the usefulness of comparing risk estimates generated using

351 multiple reference indices from several sources (e.g., regulatory agencies) when assessing the human

352 health risk associated with exposure to environmental pollutants.

353 Therefore, the proposed approach hinges on its ability to provide useful information when conducting

354 risk assessment in the presence of several reference indices from various sources. These results are

355 essential for risk managers, practitioners, and decision-makers by providing guidance and by

356 supporting direct comparisons between studies. Indeed, our approach can provide a more complete,

357 consistent and transparent risk assessment of pollutants that pose a threat to human health and support

358 better-informed risk management decisions by giving some insight into the uncertainties associated

359 with the choice of the reference index in the risk assessment. Our approach could also serve to

360 integrate risk assessments, such as by obtaining an integrated view of risk assessments when trying to

361 prioritize among different cases for remedial action.

362 To facilitate the use of our approach, the matrix of degrees of agreement, side-by-side comparison of

363 reference index-based risk distributions, the choice of the reference values to consider and the

364 construction of the aggregated multi-reference index-based risk distribution will be implemented in the

365 *InCaRisk* app for inhalation cancer risk assessment. In addition to making information more readily

366 understood and retained in a quicker time, the visual aids and representations provided by *InCaRisk*

367 could be useful for improving risk communication and promoting transparency among studies when

368 conducting risk assessments (Beck et al. 2016).

369 Regardless of the type of exposure (e.g., acute, subchronic, or chronic), absorption pathways, exposure

370 setting (environmental or occupational), stages of the life cycle of a chemical (e.g., manufacture, use,

371 disposal, consumer products or waste) and stages of the life of an individual (e.g., adulthood), the

372 approach presented in this paper can be used. However, this approach can only be used with reference

373 indices intended for the same purpose (e.g., the same absorption pathway, same type of exposure and

21

374 same exposure setting), and some prior screening of reference indices has to be done to ensure that the

375 aggregated approach does not look at a mixture of reference values that should not be integrated.

376 Finally, the proposed approach has applicability in other fields and studies with similar problematics,

377 as long as there is some constraint on the choice of reference values to be included.

378

379 *Author contributions*

380 **Pascal Petit**: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data

381 Curation, Writing - Original Draft, Visualization. **Dominique J. Bicout**: Conceptualization,

382 Methodology, Validation, Writing - Review & Editing.

383

392   **REFERENCES**

393   ANSES, 2015. Valeurs toxicologiques de référence (VTR). Guide d'élaboration de VTR. Rapport

394   d'expertise collective. Février 2010 - Edition scientifique. Seconde édition - mise à jour septembre

395   2015 [in French]. ANSES, Maisons-Alfort (France).

396   ANSES, 2019. List of toxicity reference values (TRVs) established by ANSES.

397   https://www.anses.fr/en/content/list-toxicity-reference-values-trvs-established-anses   (accessed   06

398   February 2019).

399   Beck, N.B., Becker, R.A., Erraguntla, N., Farland, W.H., Grant, R.L., Gray, G. et al., 2016.

400   Approaches for describing and communicating overall uncertainty in toxicity characterizations: U.S.

401   Environmental Protection Agency's Integrated Risk Information System (IRIS) as a case study**.**

402   **Environ. Int.** 89-90, 110-128. PMID: 26827183. DOI: http://dx.doi.org/10.1016/j.envint.2015.12.031.

403   Bond, G.G., McLaren, E.A., Baldwin, C.L., Cook, R.R., 1986. An update of mortality among

404   chemical workers exposed to benzene. **Br. J. Ind. Med.** 43, 685-691. PMID: 3465366. DOI:

405   https://dx.doi.org/10.1136%2Foem.43.10.685.

406   Chiu WA, Axelrad DA, Dalaijamts C, Dockins C, Shao K, Shapiro AJ et al., 2018. Beyond the RfD:

407   Broad Application of a Probabilistic Approach to Improve Chemical Dose-Response Assessments for

408   Noncancer Effects. **Environ. Health. Perspect.** 126(6), 067009. PMID: 29968566.   DOI:

409   https://doi.org/10.1289/ehp3368.

410   Crump, K.S., 1994. Risk of benzene-induced leukemia: a sensitivity analysis of the pliofilm cohort

411   with additional follow-up and new exposure estimates. **J. Toxicol. Environ. Health.** 42, 219–242.

412   PMID: 8207757. DOI: https://doi.org/10.1080/15287399409531875.

413   Crump, K.S., Allen, B.C., 1984. Quantitative estimates of risk of leukemia from occupational

414   exposure to benzene. Prepared for the Occupational Safety and Health Administration by

415   Science Research Systems, Inc., Ruston, LA. Unpublished. In: US EPA, 1998.

416   Health Canada, 2010. Federal Contaminated Site Risk Assessment in Canada. Part II: Health Canada

417   Toxicological Reference Values (TRVs). Version 2.0. Health Canada, Ottawa.

418 Melnick, R.L., Thayer, K.A., Bucher, J.R., 2008. Conflicting views on chemical carcinogenesis arising

419 from the design and evaluation of rodent carcinogenicity studies. **Environ. Health. Perspect.** 116(1),

420 130-135. PMID: 18197312. DOI: https://doi.org/10.1289/ehp.9989.

421 OEHHA, 2019. Consolidated table of OEHHA/ARB approved risk assessment health values.

422 https://ww3.arb.ca.gov/toxics/healthval/contable.pdf (accessed 06 February 2019).

423 Paustenbach, D., Bass, R., Price, P., 1993. Benzene toxicity and risk assessment, 1972-1992:

424 implications for future regulation. **Environ. Health. Perspect.** 101 (Suppl 6), 177-200. PMID:

425 8020442. DOI: https://doi.org/10.1289/ehp.93101s6177.

426 Petit, P., Maître, A., Bicout, D. J., 2021. A consensus approach for estimating health risk: application

427 to inhalation cancer risks. **Environ. Res.** 196, 110436. PMID: 33166535. DOI:

428 https://doi.org/10.1016/j.envres.2020.110436.

429 Richardson, D.B., 2008. Temporal variation in the association between benzene and leukemia

430 mortality. **Environ. Health. Perspect.** 116(3), 370-374. PMID: 18335105. DOI:

431 https://doi.org/10.1289/ehp.10841.

432 Rinsky, R.A., Young, R.J., Smith, A.B., 1981. Leukemia in benzene workers. **Am. J. Ind. Med.** 2,

433 217-245. PMID: 7345926. DOI: https://doi.org/10.1002/ajim.4700020305.

434 Rinsky, R.A., Smith, A.B., Hornung, R., Filloon, T.G., Young, R.J., Okun, A.H., et al., 1987. Benzene

435 and leukemia - An epidemiologic risk assessment. **N. Engl. J. Med.** 316, 1044-1050. PMID: 3561457.

436 DOI: https://doi.org/10.1056/nejm198704233161702.

437 RIVM, 2001. RIVM report 711701 025. Re-evaluation of human-toxicological maximum permissible

438 risk levels. https://www.rivm.nl/bibliotheek/rapporten/711701025.pdf (accessed 10 September 2020).

439 Rotter, S., Beronius, A., Boobis, A.R., Hanberg, A., van Klaveren, J., Luijten, M. et al., 2018.

440 Overview on legislation and scientific approaches for risk assessment of combined exposure to

441 multiple chemicals: the potential EuroMix contribution. **Crit. Rev. Toxicol.** 48(9), 796-814. PMID:

442 30632445. DOI: https://doi.org/10.1080/10408444.2018.1541964.

443 U.S. EPA, 1998. Carcinogenic effects of benzene: an update. Prepared by the

444 National Center for Environmental Health, Office of Research and Development. US EPA,

445 Washington DC. EPA/600/P-97/001F.

446 U.S. EPA, 2014. Guidance for applying quantitative data to develop data-derived extrapolation factors

447 for interspecies and intraspecies extrapolation. EPA/R-14/002. Risk Assessment Forum, Office of the

448 Science Advisor, Washington, DC.

449 U.S. EPA, 2018. Table 1. Prioritized Chronic Dose-Response Values for Screening Risk Assessments.

450 U.S. EPA, Washington (DC). https://www.epa.gov/sites/production/files/2014-

451 05/documents/table1.pdf (accessed 06 February 2019).

452 Verteramo Chiu, L.J., Tauer, L.W., Gröhn, Y.T., Smith, R.L., 2020. Ranking disease control strategies

453 with stochastic outcomes. **Prev. Vet. Med.** PMID: 32014682. DOI:

454 https://doi.org/10.1016/j.prevetmed.2020.104906.

455 WHO, 2000. Air Quality Guidelines for Europe. Second Edition. WHO Regional Publications,

456 European Series, No. 91. https://www.euro.who.int/data/assets/pdf_file/0005/74732/E71922.pdf

457 (accessed 10 September 2020).

458 Wong, O., 1987a. An industry wide mortality study of chemical workers occupational exposed to

459 benzene. I - General results. **Br. J. Ind. Med.** 44, 365-381. PMID: 3606966. DOI:

460 https://doi.org/10.1136/oem.44.6.365.

461 Wong, O., 1987b. An industry wide mortality study of chemical workers occupational exposed to

462 benzene. II - Dose response analyses. **Br. J. Ind. Med.** 44, 382-395. PMID: 3606967. DOI:

463 https://doi.org/10.1136/oem.44.6.382.

464 **Tables**

465

466 **Table 1**: Summary of the lognormal exposure concentration distribution parameters used for three

467 practical cases

Reference index source	Reference index	Case 1	Case 2	Case 3
	IUR (m^3/µg)	GM; GSD	GM; GSD	GM; GSD
Using the reference index from ANSES	2.6×10^{-5}	253 µg/m^3; 2.80	253 µg/m^3; 1.03	58.4 µg/m^3; 2.80
Using the reference index from Health Canada	3.3×10^{-6}	253 µg/m^3; 2.80	253 µg/m^3; 1.03	58.4 µg/m^3; 2.80
Using the reference index from OEHHA	2.9×10^{-5}	253 µg/m^3; 2.80	253 µg/m^3; 1.03	58.4 µg/m^3; 2.80
Using the reference index from RIVM	5.0×10^{-6}	253 µg/m^3; 2.80	253 µg/m^3; 1.03	58.4 µg/m^3; 2.80
Using the reference index from US EPA	7.8×10^{-6}	253 µg/m^3; 2.80	253 µg/m^3; 1.03	58.4 µg/m^3; 2.80
Using the reference index from WHO	6.0×10^{-6}	253 µg/m^3; 2.80	253 µg/m^3; 1.03	58.4 µg/m^3; 2.80

468 *Note*: GM: geometric mean; GSD: geometric standard deviation, IUR: inhalation unit risk, ANSES:

469 French Agency for Food, Environmental and Occupational Health & Safety, H. Canada: Health

470 Canada, OEHHA: California Office of Environmental Health Hazard Assessment, RIVM: National

471 Dutch Institute for Public Health and the Environment, US EPA: US Environmental Protection

472 Agency, WHO: World Health Organization.

Fig. 1: Scheme of the general approach to construct an aggregated risk distribution based on multiple reference indices

ω_i: contributing weight of agency "i" with $0 \leq \omega_i \leq 1$ such that $\sum_{i=1}^{n} \omega_i = 1$; $R_a(r)$: aggregated risk distribution function based on multiple reference indices; $R_i(r)$: risk distribution function of agency "i".

Fig. 2: Side-by-side comparison of single IUR-based risk distributions for the construction of the aggregated multi-IUR-based risk distribution – case 1

A: Matrix of degrees of agreement between a single IUR-based risk distribution, B: Weights attributed to the IUR-based risk distribution for the aggregated multi-IUR-based risk distribution construction, C: IUR-based risk distributions.

ANSES: French Agency for Food, Environmental and Occupational Health & Safety, H. Canada: Health Canada, OEHHA: California Office of Environmental Health Hazard Assessment, RIVM: National Dutch Institute for Public Health and the Environment, US EPA: US Environmental Protection Agency, WHO: World Health Organization. Each agency name refers to the risk distribution generated using its benzene IUR and the exposure scenario considered. The term "Inter-agency" refers here to the aggregated risk estimate based on multiple reference indices (AREMRI).

Fig. 3: Side-by-side comparison of single IUR-based risk distributions for the construction of the aggregated multi-IUR-based risk distribution – case 2

A: Matrix of degrees of agreement between a single IUR-based risk distribution, B: Weights attributed to the IUR-based risk distribution for the aggregated multi-IUR-based risk distribution construction, C: IUR-based risk distributions.

ANSES: French Agency for Food, Environmental and Occupational Health & Safety, H. Canada: Health Canada, OEHHA: California Office of Environmental Health Hazard Assessment, RIVM: National Dutch Institute for Public Health and the Environment, US EPA: US Environmental

501 Protection Agency, WHO: World Health Organization. Each agency name refers to the risk

502 distribution generated using its benzene IUR and the exposure scenario considered. The term "Inter-

503 agency" refers here to the aggregated risk estimate based on multiple reference indices (AREMRI).

504

505 **Fig. 4**: Side-by-side comparison of single IUR-based risk distributions for the construction of the

506 aggregated multi-IUR-based risk distribution – case 3

507 A: Matrix of degrees of agreement between a single IUR-based risk distribution, B: Weights attributed

508 to the IUR-based risk distribution for the aggregated multi-IUR-based risk distribution construction,

509 C: IUR-based risk distributions.

510 ANSES: French Agency for Food, Environmental and Occupational Health & Safety, H. Canada:

511 Health Canada, OEHHA: California Office of Environmental Health Hazard Assessment, RIVM:

512 National Dutch Institute for Public Health and the Environment, US EPA: US Environmental

513 Protection Agency, WHO: World Health Organization. Each agency name refers to the risk

514 distribution generated using its benzene IUR and the exposure scenario considered. The term "Inter-

515 agency" refers here to the aggregated risk estimate based on multiple reference indices (AREMRI).