Food and Chemical Toxicology 74 (2014) 279–288

Contents lists available at ScienceDirect

Food and Chemical Toxicology

journal homepage: www.elsevier.com/locate/foodchemtox

Computational tool for usual intake modelling workable at the European level

Polly E. Boon [a,*], Hilko van der Voet [b], Jiri Ruprich [c], Aida Turrini [d], Salomon Sand [e], Jacob D. van Klaveren [a]

[a] National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[b] Biometris, Wageningen University and Research Centre (WUR), Wageningen, The Netherlands
[c] National Institute of Public Health (NIPH), Prague, CHNF Brno, Czech Republic
[d] Agricultural Research Council – Research Centre for Food and Nutrition (CRA-NUT), Rome, Italy
[e] Risk Benefit Assessment Department, National Food Agency, Uppsala, Sweden

ARTICLE INFO

Article history:
Received 18 July 2014
Accepted 15 October 2014
Available online 22 October 2014

Keywords:
Usual intake
Lead
FoodEx1
EFSA
Statistical model

ABSTRACT

In this paper two models present in the computational tool Monte Carlo Risk Assessment (MCRA) were compared for assessing the usual intake of lead in five countries. For this, we used national food consumption data organised according to the format of the European Food Safety Authority (EFSA) Comprehensive database and a single lead concentration database in which analysed commodities were organised according to EFSA's Standard Sampling Description (SSD) system. This meant that both input data were coded according to the hierarchical FoodEx1 classification system. We demonstrate that the naïve Observed Individual Means model resulted in more conservative estimates of the exposure in the right tail of the exposure distribution compared to a refined usual intake model, the LogisticNormal–Normal model. With MCRA, the usual intake could be estimated with both models using food consumption and concentration data that were coded according to the hierarchical FoodEx1 classification system demonstrating that this tool can be used in EFSA's data environment. Additionally, the computational tool has functionalities 1) to check the input data quality by presenting detailed information about these data around a specified percentile of exposure and 2) to decide whether the use of a more refined usual intake model is appropriate.

1. Introduction

Different model types to assess long-term (usual) exposure to chemical substances can be distinguished. In the first model type (so-called point estimate), a single input parameter for consumption (often a mean consumption level) is combined with a single concentration value (e.g. a mean or high value). If exposure to more than one food occurs, the calculated exposures per food are added to assess the total usual exposure. This approach is for example used to assess the usual exposure to pesticides and additives within Europe (EFSA, 2007, 2012b). In the second model type, variation in amounts of food consumed in a population is included in the exposure assessment. Information on consumption per day per individual is linked to a mean concentration per food to obtain the expected exposure per food, and subsequently added to obtain the expected exposure per day per individual. Typically, information on food consumption is available over 2–7 days per individual. By subsequently averaging the exposure over the available days per individual a distribution of average exposure per individual is generated. From this distribution the mean exposure and upper percentiles (e.g. P95 or P97.5) can be obtained. This last approach is currently used to assess the exposure to food contaminants (e.g. EFSA, 2012a, 2012c)), and is also known as the Observed Individual Mean (OIM) approach (EFSA Panel on Plant Protection Products and their Residues (PPR), 2012). Even though OIM results in more refined usual exposure estimates by addressing variations in dietary patterns between individuals, this approach still results in conservative estimates of usual exposures in the tail of the exposure distribution (Boon et al., 2011; Goedhart et al., 2012). The reason for this is that the exposure estimates contain both the variation between and within individuals, whereas for a long-term exposure only the variation between individuals is relevant, since day-to-day variations level out in the long run. To remove the within individual variation from the exposure distribution variance component models have been developed, including the Iowa State University Foods Model (ISUF) (Nusser et al., 1996, 1997), the

* Corresponding author. National Institute for Public Health and the Environment (RIVM), P.O. Box 1, NL-3720 BA Bilthoven, The Netherlands. Tel.: +31 30 274 7049; fax: +31 30 274 4475.
E-mail address: polly.boon@rivm.nl (P.E. Boon).

http://dx.doi.org/10.1016/j.fct.2014.10.019

Betabinomial–Normal model (BBN) (de Boer et al., 2009), the LogisticNormal–Normal model (LLN, NCI) (Dodd, 2011; Goedhart et al., 2012), the Statistical Program to Assess Dietary Exposure (SPADE) (van Rossum et al., 2011) and the Multiple Source Method (MSM) (Harttig et al., 2011). For an overview of these models, we refer to Goedhart et al. (2012).

The use of variance component models to assess the usual exposure to chemicals is not (yet) officially accepted by the European Food Safety Authority (EFSA), and therefore not (yet) included in risk management decisions by the European Commission. Exposure assessments using these models have up to date only been performed at the national level (e.g. Boon et al., 2009; Cimenci et al., 2013; Hamborg Jensen et al., 2008; Sioen et al., 2012). In 2009, EFSA commissioned the project 'European Tool Usual Intake' (ETUI) (CFP/EFSA/DATEX/2009/03). Within this project a computational tool was developed containing several models to estimate usual intake distributions, including OIM, LNN, BBN and ISUF (de Boer and van der Voet, 2011). Based on a validation study performed within this project, the advocated model for practical use was the LNN model (Goedhart et al., 2012).

In the present paper we describe a case study to 1) demonstrate the potential of the LNN model to assess long-term dietary exposure to an adverse chemical substance compared to the OIM model, and 2) test the computational tool in EFSA's data environment (van Klaveren et al., 2012). With this environment, we mean the input data that are available to EFSA to perform European risk assessments: individual food consumption data of different European Member States present in the Comprehensive database and chemical concentration data as submitted to EFSA by the Member States. Within this environment, the hierarchical FoodEx1 classification system is used to code foods consumed (EFSA, 2011a) and commodities analysed (EFSA, 2010). To address both goals, the dietary exposure to lead was calculated using OIM and LNN as implemented in the computational tool (de Boer and van der Voet, 2011) using food consumption and lead concentration data coded according to FoodEx1. The exposure results are discussed regarding the usefulness of the computational tool in relation to dietary exposure assessments to chemical substances as presently performed by EFSA (e.g. EFSA, 2012a, 2012c).

2. Materials and methods

2.1. Food consumption data

Exposure calculations were performed using individual food consumption data of five countries: Czech Republic (CZ), France (F), Italy (IT), Netherlands (NL) and Sweden (SE). For comparability of the exposure results, calculations were performed using the consumption data of the adult age group (18–64 years), since this age group was the only one covered by the Dutch food consumption data. The food consumption data used in this paper are described in the Appendix (see also Table 1). For more detailed descriptions references per food consumption survey are given.

The consumed foods per country were all classified according to the FoodEx1 coding system as used in the EFSA Comprehensive database (EFSA, 2011a). FoodEx1 is a hierarchical system based on 20 main food categories that are further divided into subgroups up to a maximum of four levels. Level 4 is the most refined (e.g. bread) and level 1 is the least refined (e.g. grains and grain based products) hierarchical level of FoodEx1. Depending on the details available in the original national food consumption databases, foods consumed were assigned to one of the four hierarchical levels of FoodEx1. For more details, see EFSA (2011a).

Table 1
Characteristics of national consumption data from Czech Republic (CZ), France (F), Italy (IT), The Netherlands (NL) and Sweden (SE).

Country	Year	Method of consumption data collection	Adults	
			Age (years)	Number
CZ	2003–2004	2 × 24 h recall	18–64	1666
F	2005–2007	7-d record	18–64	2276
IT	2005–2006	3-d record	18–64	2313
NL	2003	2 × 24 h recall	19–30	750
SE	1997–1998	7-d record	18–64	1081

2.2. Lead concentration data

Lead concentration data as used in the 2010 EFSA lead opinion were used to assess the dietary exposure to lead (EFSA Panel on Contaminants in the Food Chain (CONTAM), 2010). These data were supplied by EFSA as part of the EFSA ETUI project, without reference to the countries in which the samples were analysed. The lead concentration data covered the period 2003–2009. Data providers authorised the use of their data. In total, 94,126 lead concentrations in various food commodities and tap water were available. These data were submitted to EFSA by 14 Member States and Norway. Germany was the major contributor providing 44% of the data, followed by France (15%), Czech Republic (9.7%) and Romania (9.6%). The samples covered all kinds of commodities that may contain lead, including milk, vegetables, fruits, cereals, fish and meat. The analysed commodities were also classified according to the FoodEx1 coding system (EFSA, 2010).

The EFSA concentration database contained many samples that were reported to contain lead at a concentration below the limit of detection (LOD) or quantification (LOQ), the so-called non-detect samples. In the case study reported here it was assumed that these samples were true zeros, containing no lead (lower bound (LB) scenario). For a more detailed description of the concentration data we refer to the 2010 EFSA opinion on lead (EFSA Panel on Contaminants in the Food Chain (CONTAM), 2010).

2.3. Linking food consumption and concentration data

Both foods consumed and commodities analysed were classified according to FoodEx1, making it possible to link the majority of the analysed commodities directly to the food consumption data. An exception was the lead concentration data of coffee, tea and cocoa. These were measured as purchased (undiluted/unprepared), whereas the consumption data refer to the amounts as consumed (diluted/prepared). Dilution factors of 18, 60 and 10 were therefore applied to the concentration values analysed in coffee, tea and cocoa, respectively, to obtain the concentrations in the consumed products (van Klaveren et al., 2012).

In the concentration database, not all foods that may contain lead were available. These data gaps were addressed by assigning mean lead concentrations to comparable foods or by assigning a mean lead concentration to a hierarchical level of the FoodEx1 classification based on the available data at the next more refined hierarchical level. For example, the mean lead concentration analysed in individual root vegetables (level 3 of FoodEx1) was assigned to the FoodEx1 food group 'Root vegetables' (level 2 of FoodEx1). In this way, all individual vegetables belonging to this food group with no lead concentration were linked to the mean lead concentration for the total group. The mean concentrations were not weighed based on consumption. FoodEx1 food groups with a mean lead concentration based on less than 15 samples were deleted from the database, when a linkage to a less refined hierarchical level with a mean lead concentration based on more than 15 samples was available. For example, for FoodEx1 food group 'Rice, brown' (level 4) only eight samples were present in the concentration database, whereas at the next less refined hierarchical level (FoodEx1 food group 'Rice', level 3) concentration data of 565 analysed samples were available. The number of at least 15 samples was arbitrarily chosen.

With the computational tool, the concentrations and consumed amounts were linked as follows. The computational tool tries first to match the food groups consumed to concentration data at the same hierarchical level. If no match is available, the tool moves up to the next less refined hierarchical level to establish a link between concentration and consumption. It stops when a match is found. If no match is found, this food was not included in the exposure assessment.

2.4. Exposure calculations

The dietary exposure was calculated using the Monte Carlo Risk Assessment (MCRA) software, release 7.1. This software contains the computational tool developed in the EFSA ETUI project (van Klaveren et al., 2012). The usual exposure was assessed using the OIM and LNN models. For these models, daily consumption patterns of individuals were multiplied with the mean lead concentration per consumed FoodEx1 food group, and summed over groups per day per individual. All daily estimated exposures were adjusted for individual body weight. With OIM these daily exposure levels were averaged over the days present in the consumption databases per individual, resulting in a distribution of individual average exposures. These exposure estimates are then taken as a proxy for the usual intake. To refine these estimates, they were additionally corrected for the day-to-day variation in exposure by applying the LNN approach (Goedhart et al., 2012).

LNN models exposure frequencies and exposure amounts separately, followed by an integration step. For the consumption frequencies, LNN fits a logistic regression model to the number of days with consumption per individual, providing both an estimate of the mean consumption frequency and of the variation between individuals in this frequency (dispersion factor). For the modelling of the positive