Environment International 106 (2017) 105–118

Contents lists available at ScienceDirect

Environment International

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

Identifying populations sensitive to environmental chemicals by simulating toxicokinetic variability

Caroline L. Ring[a,b,1], Robert G. Pearce[a,b], R. Woodrow Setzer[b], Barbara A. Wetmore[c,d], John F. Wambaugh[b,*]

[a] Oak Ridge Institute for Science and Education, Oak Ridge, TN 37831, United States
[b] National Center for Computational Toxicology, Office of Research and Development, United States Environmental Protection Agency, Research Triangle Park, NC 27711, United States
[c] ScitoVation, LLC, Research Triangle Park, NC, United States
[d] National Exposure Research Laboratory, Office of Research and Development, United States Environmental Protection Agency, Research Triangle Park, NC 27711, United States

ARTICLE INFO

Keywords:
Risk assessment
Toxicokinetics
High throughput
Environmental chemicals
IVIVE

ABSTRACT

The thousands of chemicals present in the environment (USGAO, 2013) must be triaged to identify priority chemicals for human health risk research. Most chemicals have little of the toxicokinetic (TK) data that are necessary for relating exposures to tissue concentrations that are believed to be toxic. Ongoing efforts have collected limited, *in vitro* TK data for a few hundred chemicals. These data have been combined with biomonitoring data to estimate an approximate margin between potential hazard and exposure. The most "at risk" 95th percentile of adults have been identified from simulated populations that are generated either using standard "average" adult human parameters or very specific cohorts such as Northern Europeans. To better reflect the modern U.S. population, we developed a population simulation using physiologies based on distributions of demographic and anthropometric quantities from the most recent U.S. Centers for Disease Control and Prevention National Health and Nutrition Examination Survey (NHANES) data. This allowed incorporation of inter-individual variability, including variability across relevant demographic subgroups. Variability was analyzed with a Monte Carlo approach that accounted for the correlation structure in physiological parameters. To identify portions of the U.S. population that are more at risk for specific chemicals, physiologic variability was incorporated within an open-source high-throughput (HT) TK modeling framework. We prioritized 50 chemicals based on estimates of both potential hazard and exposure. Potential hazard was estimated from *in vitro* HT screening assays (*i.e.*, the Tox21 and ToxCast programs). Bioactive *in vitro* concentrations were extrapolated to doses that produce equivalent concentrations in body tissues using a reverse dosimetry approach in which generic TK models are parameterized with: 1) chemical-specific parameters derived from *in vitro* measurements and predicted from chemical structure; and 2) with physiological parameters for a virtual population. For risk-based prioritization of chemicals, predicted bioactive equivalent doses were compared to demographic-specific inferences of exposure rates that were based on NHANES urinary analyte biomonitoring data. The inclusion of NHANES-derived inter-individual variability decreased predicted bioactive equivalent doses by 12% on average for the total population when compared to previous methods. However, for some combinations of chemical and demographic groups the margin was reduced by as much as three quarters. This TK modeling framework allows targeted risk prioritization of chemicals for demographic groups of interest, including potentially sensitive life stages and subpopulations.

Abbreviations: EPA, Environmental Protection Agency; HTS, High-Throughput Screening; IVIVE, *in vivo-in vitro* extrapolation; OED, Oral equivalent dose; AER, Activity:Exposure ratio; TK, Toxicokinetics; HTTK, High Throughput TK; NHANES, National Health and Nutrition Examination Survey; HTTK-package, Open source, public R tool for HTTK; HTTK-Pop, HTTK-package with human variability informed by NHANES; MC, Monte Carlo; GFR, Glomerular filtration rate (kidney); F_{up}, F_{ub}, Fraction of chemical unbound in plasma or blood; CL_{int}, intrinsic chemical clearance by hepatocytes; C_{ss}, The plasma concentration resulting from steady-state exposure

* Corresponding author at: 109 T.W. Alexander Dr, NC 27711, United States.
E-mail address: Wambaugh.John@epa.gov (J.F. Wambaugh).

[1] Current Address: Caroline L. Ring is now employed by ToxStrategies, Inc., a scientific consulting firm whose clients include private industry, trade associations, and governmental entities. However, Dr. Ring received no funding from ToxStrategies or any of its clients for this project, and neither ToxStrategies nor any of its clients was involved in the development or approval of this research or this report.

http://dx.doi.org/10.1016/j.envint.2017.06.004
Received 18 January 2017; Received in revised form 1 June 2017; Accepted 2 June 2017
0160-4120/ Published by Elsevier Ltd.

1. Introduction

The U.S. Environmental Protection Agency (EPA) is responsible for determining risks associated with chemicals in the environment. In order to address the thousands of man-made chemicals present (USGAO, 2013) in environmental samples (Rager et al., 2016) and human blood (Park et al., 2012), the EPA requires efficient risk-based methods to prioritize, screen, and evaluate chemicals. A promising framework for prioritization (Thomas et al., 2013; Wetmore et al., 2015) identifies chemicals with greater putative risk by combining data from high-throughput *in vitro* toxicity screening (HTS) assays, such as those developed and used by the US EPA ToxCast program (Kavlock et al., 2012), with data from high-throughput exposure modeling frameworks, such as US EPA ExpoCast (Wambaugh et al., 2013; Wambaugh et al., 2014). Those chemicals more likely to pose a risk become targets for further investigation.

In vitro bioactivity HTS has been conducted for thousands of chemicals to date (Kavlock et al., 2012). To use *in vitro* HTS for chemical prioritization, methods for *in vivo-in vitro* extrapolation (IVIVE) have been developed that relate external chemical exposures to internal tissue concentrations (Aylward and Hays, 2011; Rotroff et al., 2010). A primary application of IVIVE in risk prioritization has been reverse dosimetry, which uses toxicokinetic (TK) modeling to predict the oral equivalent dose (OED) of a chemical needed to produce an internal (*e.g.*, plasma) concentration equal to a bioactive *in vitro* concentration (Rotroff et al., 2010; Tan et al., 2007; Tan et al., 2006). The ratio of OED to estimated human exposure is a measure of potential risk (Judson et al., 2011; Thomas et al., 2013); this ratio is known as the activity:exposure ratio, or AER.

Unfortunately, unlike the thousands of chemicals with predicted estimates for toxicity and exposure, TK data from traditional methods are available for only a few dozen chemicals (Wetmore et al., 2015; Wetmore et al., 2012). Alternative *in vitro* methods for TK have allowed the development of very simple prototype TK models for many hundreds of chemicals (Rotroff et al., 2010; Wetmore et al., 2015; Wetmore et al., 2012). These simple "high throughput" TK (HTTK) models are useful precisely because they are simple – they can be rapidly parameterized using *in vitro* measurements of chemical clearance by hepatocytes and plasma protein binding (Rotroff et al., 2010) as well as bioavailability data (Wetmore et al., 2012). With additional data, more elaborate models have been generated (Wambaugh et al., 2015), including simulation of population variability in metabolizing enzymes (Wetmore et al., 2014). However, these data cannot currently be rapidly generated for large numbers of chemicals. The minimal HTTK data used for characterizing chemicals constrains models built on these data in how they may describe variability between individuals. However, these HTTK data and models have been useful for identifying those chemicals that are more likely to pose a human health risk (Thomas et al., 2013; Tonnelier et al., 2012; USEPA, 2014; Wetmore et al., 2015; Wetmore et al., 2012).

TK, hazard, and exposure are known to vary between individuals, life stages, and populations with varying genetics, ontogeny, and physiology (Belle and Singh, 2008; Hines, 2007; Jamei et al., 2009a; Lipscomb and Kedderis, 2002; McNally et al., 2014; Wambaugh et al., 2014). Therefore, to better describe vulnerable life stages and populations, HT risk-based chemical prioritization needs to incorporate inter-individual variability in predictions of both hazard and exposure. The open-source, publicly available R package 'httk' (hereafter referred to as "HTTK-package"), was developed to facilitate HTTK modeling for IVIVE (Pearce et al., 2016). Currently, the HTTK-package incorporates the ability to simulate inter-individual physiological variability by Monte Carlo (MC) sampling of the HTTK model parameters using uncorrelated normal distributions characterized by means and coefficients of variation (Wambaugh et al., 2015); these distributions typically reflect the physiology of a healthy young Caucasian adult (Birnbaum et al., 1994; Valentin, 2002).

Here, we incorporate the inter-individual variability of the modern U.S. population into high-throughput risk-based chemical prioritization. To do this, we develop a population physiology simulation that makes demographic-specific predictions of chemical risk, using data collected as part of the ongoing National Health and Nutrition Examination Survey (NHANES) performed by the Centers for Disease Control (CDC) (http://www.cdc.gov/nchs/nhanes.htm) (Johnson et al., 2014). This new tool, which we refer to here as "HTTK-Pop," has been publicly released for use by the TK and risk assessment communities.

HTTK−Pop uses a correlated MC approach to simulate inter-individual physiological variability across demographic groups. We evaluate our approach using predictions of steady-state plasma concentrations (C_{ss}) derived from *in vivo* measurements for 95 pharmaceutical (Obach et al., 2008) and other (Wetmore et al., 2012) compounds. We further assess predictions of inter-individual variability in C_{ss} using published *in vivo* measurements from 86 studies of 14 compounds (Howgate et al., 2006; Johnson et al., 2006).

We demonstrate the impact of human variability using an example framework for chemical risk prioritization (Wetmore et al., 2015). We simulate ten important demographic groups of the U.S. population informed by the NHANES. For each demographic group, we use chemical bioactivity HTS data and our new description of physiology for individuals within each demographic to predict the population distributions of doses needed to cause bioactivity for 50 chemicals. We then estimate the group-specific AERs as a measure of potential risk. Finally, we identify chemicals and groups within the U.S. population with greater predicted risk of chemical exposure-induced bioactivity, as characterized by high throughput methods.

2. Methods

2.1. High throughput toxicokinetics data

All the TK data and models for HTTK−Pop used in this analysis are open source and publicly available as R package "httk" (Pearce et al., 2016) v1.5. This version included literature data on 543 chemicals, but our analysis was limited to 50 chemicals for which both HTS bioactivity data (Section 2.8) and exposure inferences (Section 2.9) were available. In the HTTK-package the chemicals are described by physico-chemical properties (molecular weight and hydrophobicity) obtained from EPI Suite (USEPA, 2015), as well as ionization equilibria (Strope et al., 2015). The chemicals are further described by *in vitro* measurements of pooled human hepatocyte clearance and plasma protein binding, as in Wetmore et al. (2015, 2012). Absorption is assumed to be fast (1/h) and bioavailability is assumed to be 100%. The MC population simulation methods included in previous versions of the HTTK-package were replaced in version 1.5 by the HTTK−Pop methods described here.

2.2. Model used for reverse TK

The HTTK-package includes several TK models. For our analysis, we used a general TK model from previous HT risk prioritization studies (Wetmore et al., 2014; Wetmore et al., 2015; Wetmore et al., 2012) to predict steady-state plasma concentrations (C_{ss}):

$$C_{ss} = \frac{k_o}{(GFR \times F_{ub}) + \frac{Q_{liver} \times F_{ub} \times CL_{int,h}}{Q_{liver} + F_{ub} \times CL_{int,h}}} \quad (1)$$

In Eq. (1), k_o represents the dose rate (mg/kg/h); F_{ub}, the fraction of parent compound unbound in blood; Q_{liver}, the hepatic portal vein blood flow per kg body weight (L/h/kg); GFR, the glomerular filtration rate per kg body weight (L/h/kg); $CL_{int,h}$, the whole-liver intrinsic clearance rate per kg body weight under first-order metabolism conditions (L/h/kg). The model is equivalent to the steady-state concentration in a three-compartment (liver, gut, and body blood) model. It assumes zero-order uptake of a daily dose from the gut with 100%