Chapter 2

Computational Toxicology: Application in Environmental Chemicals

Yu-Mei Tan, Rory Conolly, Daniel T. Chang, Rogelio Tornero-Velez, Michael R. Goldsmith, Shane D. Peterson, and Curtis C. Dary

Abstract

This chapter provides an overview of computational models that describe various aspects of the source-to-health effect continuum. Fate and transport models describe the release, transportation, and transformation of chemicals from sources of emission throughout the general environment. Exposure models integrate the microenvironmental concentrations with the amount of time an individual spends in these microenvironments to estimate the intensity, frequency, and duration of contact with environmental chemicals. Physiologically based pharmacokinetic (PBPK) models incorporate mechanistic biological information to predict chemical-specific absorption, distribution, metabolism, and excretion. Values of parameters in PBPK models can be measured in vitro, in vivo, or estimated using computational molecular modeling. Computational modeling is also used to predict the respiratory tract dosimetry of inhaled gases and particulates [computational fluid dynamics (CFD) models], to describe the normal and xenobiotic-perturbed behaviors of signaling pathways, and to analyze the growth kinetics of preneoplastic lesions and predict tumor incidence (clonal growth models).

Key words: Computational toxicology, Source-to-effect continuum, Fate and transport, Dosimetry, Signaling pathway, Physiologically based pharmacokinetic model, Biologically based dose response model, Clonal growth model, Virtual tissue

1. Overview

Computational toxicology involves a variety of computational tools including databases, statistical analysis packages, and predictive models. In this chapter, we focus on computational models that describe various aspects of the source-to-health effect continuum (Fig. 1). Literature on the application of computational models across the continuum has been expanding rapidly in recent years. Using the Web of Science portal, we conducted a
bibliometric analysis of publications that appeared between 1970 and 2009. Using the search structure \[ \text{TS} = (\text{computational OR "in silico" OR predictive OR model* OR virtual}) \text{ AND TS} = (\text{toxicology}) \text{ AND TS} = (\text{environment*}) \] (TS: Topic), a total of 397 articles were found. Adding “NOT pharmaceutic*” to the search structure above, found 371 articles, indicating only a small fraction of the 397 deal with aspects of drug development. A PubMed search (Feb 17, 2011) on “physiologically based pharmacokinetic (PBPK) modeling” found 769 articles, indicating that our search, which focused on computational modeling specifically in environmental toxicology, was quite restrictive.

Literature searches using specific terminology were performed to understand the publication frequency of some of the most common types of modeling used in computational toxicology, including fate and transport, exposure, PBPK, computational fluid dynamic (CFD), signaling pathway, biologically based dose–response (BBDR), and clonal growth modeling. Searches were restricted to original scientific publications only (i.e., reviews were excluded) and fields of science were restricted (e.g., “NOT eco*”) in order to focus on applications relevant to human health effects. A yearly breakdown showing publication frequency over time is presented in Fig. 2. The data show a rapid increase in publication frequency for many of the modeling types beginning in the early 1990s and that PBPK, fate and transport and signaling pathways are the most common. BBDR and clonal growth modeling have received considerably less attention.

Fig. 1. Major components of the source-to-effect continuum.
2. Computational Models Along the Source-to-Health Effect Continuum

2.1. Fate and Transport

Fate and transport models describe the release, transportation, and transformation of chemicals from sources of emission throughout the general environment. Fate addresses persistence, dissipation, and loss of chemical mass along the migration pathway; and transport addresses mobility of a chemical along the migration pathway (1). Based on their complexity, models of fate and transport can be used for either “screening-level” or “higher-tiered” applications (2). Screening-level models often use default input parameters that tend to over-predict exposures (the preferred default approach used in the absence of data). These models are suitable for obtaining a first approximation or to screen out exposures that are not likely to be of concern (3). Screening-level models have limited spatial and temporal scope. Higher-tiered models are needed when analyses require greater temporal and spatial resolution, but much more information is required, such as site-specific data.

The processes that can be described in fate and transport models include advection, dispersion, diffusion, equilibrium partitioning between solid and fluid, biodegradation, and phase separation of immiscible liquids (1). In general, fate and transport models require information on physicochemical properties; mechanisms of release
of chemicals to environmental media; physical, chemical, and biological properties of the media though which migration occurs; and interactions between the chemical and medium (1). For example, typical inputs to an air quality and dispersion model are source data (e.g., emission rates), meteorological data (e.g., temperature), and physicochemical properties of the chemical. Inputs to a surface water model, in addition to source data and physicochemical properties, may include water flows, soil properties and topography, and advective/dispersive movement (2).

2.2. Exposure

The outputs of a fate and transport model are concentrations to which humans may be exposed. These predicted concentrations are then used, in some cases, as surrogates for actual exposure (2). Since these provisional estimates do not provide sufficient resolution about variation of exposure among individuals and by time and location, they can also be used as inputs to exposure models. Exposure models integrate the microenvironmental concentrations with the amount of time an individual spends in these microenvironments to provide qualitative and quantitative evaluations of the intensity, frequency, and duration of contact with chemicals, and sometimes, the resulting amount of chemicals that is actually absorbed into the exposed organism. Exposure models vary considerably in their complexity. Some models are deterministic and generate site of contact-specific point estimates (e.g., dermal concentration × contact time). Others are probabilistic, describing spatial and temporal profiles of chemical concentrations in microenvironments. Both deterministic and probabilistic models may aggregate some or all of the major exposure pathways.

Probabilistic models can also be used to describe variability in human behavior. Human activities contribute to exposure variability, and at first glance appear to be arbitrary, yet patterns of behavior are known to be representative of different age groups (e.g., hand-to-mouth behavior among 3–5 year olds) and this information can be used to better inform stochastic exposure models (4). A major challenge in characterizing human activity is overcoming the cost of collecting information. For example, food consumption questionnaires are important in dietary modeling (e.g., estimating chronic arsenic exposure by shellfish consumption); however the accuracy in assessing chronic exposure is limited by the lack of longitudinal survey information in the surveys such as Continuing Survey of Food Intake by Individuals (CSFII) and National Health and Nutrition Examination Survey (NHANES) (5, 6). The recent study of Song et al. (7) examined how much information is needed in order to predict human behavior. The authors examined the predictability of macro-scale human mobility over a span of 3 months based on cell phone use—comparing a continuous record (e.g., hourly) of a user’s momentary location with a less expensive measure of mobility. The authors found that there is a potential
93% average predictability in user mobility. This predictability reflects the inherent regularity of human behavior (7) and exemplifies an approach that holds promise for examining aspects of human mobility, thereby reducing the cost of exposure modeling.

The degree of complexity needed in an exposure model depends on (1) the nature of the chemical (e.g., volatility) and (2) the number and complexity of the most common exposure scenarios that the model is required to describe. The number of parameters in the model and their corresponding data needs are functions of model complexity. The first choice for obtaining input parameter data is direct measurement of the environment concentrations and observations of human activity patterns. When these specific data are not available, inputs may be obtained from population-based surveys, such as NHANES or the Exposure Factors Handbook (8). The outputs of fate, transport, and exposure models can serve as inputs to pharmacokinetic models for estimating internal tissue dosimetry.

### 2.3. Dosimetry

Pharmacokinetic processes translate the exposure or applied dose into a delivered dose at an internal site. Internal doses often correlate better with apical effects than do the external doses due to nonlinear pharmacokinetics (9). Pharmacokinetic data can be obtained from studies using laboratory animals (10) or from controlled human exposures (11, 12). Controlled human exposures are largely reserved for evaluating the safety and efficacy of drugs or therapies not for environmental chemicals. Human observational studies (13–15) may provide some insight into the disposition of environmental chemicals, but the relationship between exposure and systemic levels is obscured because of the complexity of exposure. Furthermore, the lack of control with regards to chemical co-occurrence may confound interpretation of the exposure–dose relationship.

The relationship between exposure to a chemical and its dose at an internal target site is determined by a set of chemical structure-dependent properties (e.g., solubility in water, blood, and tissues, volatility, susceptibility to biotransformation) and corresponding properties of the biological system (e.g., tissue volumes, blood flows, metabolic capabilities). Computational models that describe the minimum set of these characteristics needed to predict chemical-specific absorption, distribution, metabolism, excretion (ADME) are commonly referred to as PBPK models though PBTK, were the T stands for toxicokinetic, is also used. Because models of this type describe the relevant biology that determines ADME, they are useful not only for predicting pharmacokinetic behavior within the dose range and time course of available data but also for extrapolation outside these ranges. These characteristics make these models particularly useful in risk assessments, where extrapolation to doses well below those for which data are available is often necessary (16).
Many of the parameters used in PBPK models can be measured in vitro (17). Obach and colleagues (18, 19) observed that scaling in vitro metabolism data from human liver microsomes to in vivo clearance values yielded predictions that were within 70–80% of actual values. They also found that the clearance predictions were improved by accounting for plasma and microsomal protein binding. Tornero-Velez and colleagues (20) applied the same approach to account for deltamethrin’s age-dependent pharmacokinetics in the maturing Sprague-Dawley rats using in vitro parameters for hepatic and plasma metabolic clearance of deltamethrin. Finding agreement between in vitro parameter values and in vivo parameter estimates is one way to explore pharmacokinetic mechanisms and reduce pharmacokinetic data gaps. In the absence of data, however, which may often be the case for new chemicals, the exposure–dose modeler may turn to the emerging field of molecular modeling and chemoinformatics to obtain provisional pharmacokinetic values.

Molecular modeling makes use of a wide variety of techniques to predict or understand chemical behavior at the atomic level. Modeling chemical interactions is an important step in understanding the molecular events encountered in both biological and environmental systems (21–23). These methods have the potential to explain the underlying molecular processes of chemical interactions and transformations in the source–exposure–dose–response continuum. Here, the primary use of such tools is to provide in silico predictions of relevant data where little or no actual data exist. Provisional estimates derived from structure–activity relationships may then be tested using focused methods to validate or augment parameter values.

The field of molecular modeling comprises a wide variety of tools from chemoinformatics-based disciplines [e.g., quantitative structure–activity relationships (QSAR)] and graph network theory (e.g., two-dimensional topological molecular descriptors) to detailed atomistic simulations (e.g., molecular dynamics) and quantum mechanical simulations of the electron distributions of a molecule. Chemoinformatic techniques have a long history in promoting simple concepts such as lipophilicity and partitioning (24) as indicators of persistence and toxicity within the environment (i.e., fate and transport) (25). These techniques are also used to obtain indicators of chemical disposition (26) and pharmacodynamics (27) within biological organisms (28). Many software packages exist whereby one can develop, augment, and utilize new or existing QSARs for parameters such as blood–brain barrier transfer coefficients, dermal permeation rates, cell line permeability, and octanol–water partition coefficient [e.g., molecular operating environment (MOE), QiKProP (https://www.schrodinger.com/products/14/17/), and OpenEye (http://www.eyesopen.com)]. These QSAR packages are generally confined within biological system analysis as seen on the right side of the source–exposure–dose–response continuum (Fig. 1).
For environmental fate and transport models, QSAR can be used to estimate the values of the physicochemical parameters describing the partitioning and transfer processes among air, water, and soil. For example, the US EPAs SPARC predictive modeling system is able to calculate large numbers of physical/chemical parameters from molecular structure and basic information about the environment (e.g., media, temperature, pressure, PH). These parameters are used in fate and transport modeling of organic pollutants, nutrients, and other stressors.

Techniques such as QSAR are ideally suited for rapid evaluation of parameters for pharmacokinetic and fate and transport models. However, development of these techniques is data intensive, requiring training sets with well-defined endpoints to develop the relationship between chemical structure and observed activity. In addition, QSAR models are fitted to specific molecular subsets (training set) and it is difficult to apply them to compounds outside the chemical space represented in the training set.

While QSAR is the more known molecular modeling technique within computational toxicology, there are other tools, such as classical force-field docking techniques, that can aid in understanding the biological processes which involve chemical interactions with biomolecular targets. Inter-molecular interactions between ligands and biomolecular targets determine binding mechanics that ultimately lead to altered physiological responses and potential toxicological effects. Thus, an understanding of the relevant binding interactions can lead to a better understanding of chemical function and provide a visual representation of chemical binding and mechanisms of toxicity. For example, estimating the relative binding affinities of 281 chemicals to a surrogate rat estrogen receptor, Rabinowitz et al. (29) utilized docking techniques to screen out 16 actives (“true competitive inhibitors”) from nonactive substrates with no false negatives and eight misclassified false positives. Molecular dynamics (17, 30) or ab initio molecular dynamics (31) can be used to simulate time-evolving processes such as diffusion through environmental media, solvation effects, and “classical” kinetic rate constants (e.g., solvent-mediated hydrolysis, oxidation, and hydrogen abstraction rates). This information can be used as chemical-specific inputs to pharmacokinetic and environmental fate and transport models (32–37).

Computational models are also used to predict the respiratory tract dosimetry of inhaled gases and particulates. These models are needed because the complex shapes of the nasal airways and the branching pattern of the airways leading from the trachea to the alveoli often result in nonuniform deposition of inhaled materials. Models of the respiratory tract incorporate varying degrees of anatomical realism. CFD models of the nasal airways use accurate, three-dimensional reconstructions of the airways (38), while one-dimensional reconstructions have been more commonly used for the pulmonary airways (39).
2.4. Signaling Pathways

Signaling pathways such as the mitogen-activated protein kinase (MAPK) pathway (40) consist of one or more receptors at the cell surface that, when activated by their cognate ligands, transmit signals to cytosolic effectors and also to the genome. The cytosolic effects are rapid, occurring within seconds or minutes of receptor activation, while the effects on gene expression take longer, with changes in the associated protein levels typically occurring after one or more hours. A number of computational models of signaling pathways have been described (e.g., (41, 42).

The National Research Council (NRC) report, Toxicity Testing in the twenty-first century (18) introduced the concept of “toxicity pathways.” Toxicity pathways were defined by the NAS as “interconnected pathways composed of complex biochemical interactions of genes, proteins, and small molecules that maintain normal cellular function, control communication between cells, and allow cells to adapt to changes in their environment” and which, “when sufficiently perturbed, are expected to result in adverse health effects are termed toxicity pathways” (43). The adverse effect is the clinically evident effect on health and is often referred to as the apical effect, denoting its placement at the terminal end of the toxicity pathways. Although not much work has been done to date, computational models of signaling pathways are expected to be integral components of toxicity pathway models.

2.5. BBDR/Clonal Growth

Cancer is a disease of cell division. In healthy tissue, the respective rates of cellular division and death are tightly regulated, allowing for either controlled growth or the maintenance of tissue size in adulthood. When regulation of division and death rates is disrupted, tumors can develop. (It should also be noted that embryonic development depends on tight regulation of division and death rates, where dysregulation can result in malformations). A number of computational models have been developed to describe tumor incidence and the growth kinetics of preneoplastic lesions. These vary from purely statistical models fit to incidence data (44) to models that track time-dependent division and death rates of cells in the various stages of multi-stage carcinogenesis (45). These latter kinds of models provide insights into how different kinds of toxic effects—e.g., direct reactivity with DNA versus cytotoxicity—can differentially affect tumor development.

BBDR models represent the entire exposure—target site dose—apical response continuum. These kinds of models require large amounts of supporting data but have the capability to predict both dose–response and time course for development of apical effects as well as for some intermediate effects (e.g., (46). This latter capability is important as it provides the opportunity to use data on biomarkers in support of model development. The resources needed to develop such models are, unfortunately,
seldom available. In some cases, however, where the economic importance or the degree of human exposure is sufficient, development of BBDR models can be justified.

3. Virtual Tissues

The computational models described above incorporate varying degrees of biological detail. Over time, these models will be refined as new data and new degrees of understanding of the relevant biological processes become available. Taking a long-term view, the iterative refinement of such models will lead asymptotically to the development of virtual tissues, where multiple scales of biology—molecular, macromolecular, organelle, tissue—are described in a spatially and temporally realistic manner. Numerous efforts that are self-described as virtual tissues are underway (47–49). While important and useful, these are, however, preliminary steps toward actual development of virtual tissues, and we do not expect that this goal will be realized for some time. However, while a long-term goal of computational toxicology, virtual tissues and, by extension, virtual organisms, have the potential to eventually reduce and perhaps even eliminate the use of laboratory animals, thereby revolutionizing toxicity testing.

Disclaimer

The United States Environmental Protection Agency through its Office of Research and Development funded and managed the research described here. It has been subjected to Agency’s administrative review and approved for publication.

References


Computational Toxicology
Volume I
Reisfeld, B.; Mayeno, A.N. (Eds.)
2013, XII, 612 p., Hardcover
ISBN: 978-1-62703-049-6
A product of Humana Press