A Web-Based Database on Exposure to Persistent Organic Pollutants in China

Zhaomin Dong,¹,² Xiarui Fan,¹ Yao Li,¹ Ziwei Wang,¹ Lili Chen,³ Ying Wang,¹,² Xiaoli Zhao,⁴ Wenhong Fan,¹,² and FengChang Wu⁴

¹School of Space and Environment, Beihang University, Beijing, China
²Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing, China
³Beijing Academy of Edge Computing, Beijing, China
⁴State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, China

https://doi.org/10.1289/EHP8685

Introduction

To comply with the Stockholm Convention, China has initiated a series of research activities to restrict, reduce, and eliminate the production, use, and emission of persistent organic pollutants (POPs). Major findings from these activities have been regularly documented in peer-reviewed journals (Chen et al. 2020; Jiang et al. 2019). Although many studies have been performed, the monitoring results have been reported in different contexts, which typically vary in terms of sample inclusion criteria, field treatments, target POPs or congeners, analytical methods, statistical descriptions, units, and other aspects (Lau et al. 2012). To better understand the nationwide occurrence of POPs, capacity buildings on data collection, harmonization, and analysis are urgently needed.

Toward this goal, we first identified publications related to the monitoring of POPs in China, and then manually extracted data records on the monitoring of POPs in a range of environmental elements (such as air, soil, sediment, dust, water, and food), biota, and human matrices (blood, urine, breast milk, and placenta). In particular, the data extracted from literature were marked and recorded by the first author and then cross-checked by another independent author. After that, we developed an online database (https://pops.hhra.net) that collates literature information, exposure data, spatiotemporal analysis, and risk assessments of POPs in China.

Methods

The establishment of this online database consisted of five steps (Figure 1). First, we created a chemical panel that comprised all POPs listed in the Stockholm Convention. In step 2, we identified relevant literature from four databases, including PubMed, Web of Science, Scopus, and China National Knowledge Infrastructure (CNKI). The titles, authors, publication dates, journal names, and digital object identifiers (DOIs) of the resulting publications were extracted by text mining (Barupal and Fiehn 2019). Briefly, a systematic search was performed with a query based on the combined keywords of “chemical names” and “exposure pathways” and “China.” Multiple chemical names were used for each chemical to retrieve as much of the available literature as possible. We also denoted 11 pathways of exposure as follows: air, soil, dust, sediment, water, food, biological, blood, serum, plasma, and breast milk. Regarding the PubMed database, we used an R-based web crawler to identify target literature. With respect to other three databases, we directly exported the search results based on query sentences. The supplementary methodology, code for web crawler, search strategy, and query sentences for each database were available on GitHub (https://github.com/POPs-EXPs/beta-version).

In step 3, we excluded the irrelevant literature based on title, abstract, and full-text examination, and the data from the remaining literature were manually extracted and cross-checked by two authors. For each record, we included the available items, including chemical name, chemical name, congener, DOIs, sampling location (province, city, and site), sampling time, sample size, pathway, unit and statistics (range, mean, median, and standard deviation), and method detection limit. After standardizing (such as data clean and unit conversion) the extracted exposure data, we initiated the functions of spatiotemporal analysis and risk assessment in step 4. Finally, the online database and data visualization were projected online (https://pops.hhra.net) using (version 1.6.0, R 3.6.0, Windows 10) R “Shiny.”

Results and Discussion

The literature search yielded 9,131 publications that had been published before the end of May 2020. Irrelevant studies (6,332) were excluded based on the examination of titles, abstracts, and full-texts, after which 2,799 studies were kept (Figure 2A). Figure 2B showed that there had been a rapid annual increase in publications on POPs over the last two decades. Figure 2C indicated that the most commonly studied POPs group was organochlorine pesticides (OCPs), followed by dioxins and polychlorinated biphenyls (PCBs), brominated flame retardants (BFRs), per- and polyfluoroalkyl substances (PFAS) and other POPs. The trends in the numbers of environmental matrices associated studies were also illustrated in Figure 2D.

Based on the remaining 2,799 publications, we manually extracted 112,878 records on the exposure of POPs. In particular, there were 21,690 data items for OCPs; 62,495 for dioxins and PCBs; 22,466 for BFRs; 4,534 for PFAS; and 1,693 for other POPs. All records were combined to form our online dashboard of the exposure database. Also, the data analysis and human health risk assessment can be easily performed by setting a range of condition filters (such as chemical panel, pathway, sampling information, unit, and others). In addition, this database will be annually updated by mid-August of each calendar year by integrating exposure data extracted from the latest publications.

The ways in which our database can assist research include at least four aspects: a) the spatiotemporal patterns on the external exposure and risk of POPs, which promise to help evaluate the management of POPs; b) the optimization of pharmacokinetic models and associated parameters, as exemplified by an existing case study that successfully derived the half-lives of short-chain chlorinated paraffins using the simultaneous monitoring of internal and external exposure (Dong et al. 2020); c) the prediction of the internal exposure of POPs using the “forward approach” (Sobus et al. 2015);
d) the reconstruction of external exposure, relying on internal exposure, pharmacokinetic models, and advance statistical techniques (Lyons et al. 2008).

In summary, this study established an open-access database (https://pops.hhra.net) that included 112,878 records extracted from 9,131 publications, serving as, to our knowledge, the first tool illustrating the exposure of POPs in China. This large data set helps reveal the research status and occurrence of POPs in China and also released more research opportunities to fulfill China’s obligations as a signatory to the Stockholm Convention.

Figure 1. The framework used to establish online database of exposure to persistent organic pollutants (POPs) in China. Note: CNKI, China National Knowledge Infrastructure.

Figure 2. The trends for persistent organic pollutants (POPs) exposure research conducted in China: (A) overlap analysis of the 2,799 publications used to extract exposure data; (B) the annual number of publications associated with POP exposure in China; (C) the number of publications on various POPs; and (D) the number of publications reporting POP concentrations in different environmental matrices. Note: BFRs, brominated flame retardants; DPs, dioxins and polychlorinated biphenyls (dioxins and PCBs); OCPs, organochlorine pesticides; PFAS, per- and polyfluoroalkyl substances.
Data Sharing
The code, data, manual, and supplementary methodology are made freely available via GitHub (https://github.com/POPs-EXPs/beta-version).

Acknowledgments
This study was supported by National Natural Science Foundation of China (21707006, 41991315, and 41521003). The ongoing maintenance and update of database will be implemented by the Chinese Research Academy of Environmental Sciences and Beihang University.

References
Barupal DK, Fiehn O. 2019. Generating the blood exposome database using a comprehensive text mining and database fusion approach. Environ Health Perspect 127(9):97008, PMID: 31557052, https://doi.org/10.1289/EHP4713.
Chen Y, Liu Q, Ma J, Yang S, Wu Y, An Y. 2020. A review on organophosphate flame retardants in indoor dust from China: implications for human exposure. Chemosphere 260:127633, PMID: 32683019, https://doi.org/10.1016/j.chemosphere.2020.127633.
Dong Z, Li T, Wan Y, Sun Y, Hu J. 2020. Physiologically based pharmacokinetic modeling for chlorinated paraffins in rats and humans: importance of biliary excretion. Environ Sci Technol 54(2):938–946, PMID: 31736300, https://doi.org/10.1021/acs.est.9b03991.
Jiang Y, Yuan L, Lin Q, Ma S, Yu Y. 2019. Polybrominated diphenyl ethers in the environment and human external and internal exposure in China: a review. Sci Total Environ 696:133902, PMID: 31470322, https://doi.org/10.1016/j.scitotenv.2019.133902.
Lau MHY, Leung KMY, Wong SWY, Wang H, Yan Z-G. 2012. Environmental policy, legislation and management of persistent organic pollutants (POPs) in China. Environ Pollut 165:182–192, PMID: 22285063, https://doi.org/10.1016/j.envpol.2012.01.002.
Lyons MA, Yang RSH, Mayeno AN, Reisfeld B. 2008. Computational toxicology of chloroform: reverse dosimetry using Bayesian inference, Markov chain Monte Carlo simulation, and human biomonitoring data. Environ Health Perspect 116(8):1040–1046, PMID: 18709138, https://doi.org/10.1289/ehp.11079.
Sobus JR, DeWoskin RS, Tan Y-M, Pleil JD, Phillips MB, George BJ, et al. 2015. Uses of NHANES biomarker data for chemical risk assessment: trends, challenges, and opportunities. Environ Health Perspect 123(10):919–927, PMID: 25959901, https://doi.org/10.1289/ehp.1409177.