Superfund Research Center Faculty Publications

Superfund Research Center

3-1-2017

An Open-Sourced Statistical Application for Identifying Complex Toxicological Interactions of Environmental Pollutants

Jordan T. Perkins
University of Kentucky

Michael C. Petriello
University of Kentucky, michaelcpetriello@gmail.com

Li Xu
University of Kentucky

Arnold J. Stromberg
University of Kentucky, astro11@uky.edu

Bernhard Hennig
University of Kentucky, bhennig@uky.edu

Click here to let us know how access to this document benefits you.

Follow this and additional works at: https://uknowledge.uky.edu/superfund_facpub

Part of the Environmental Health and Protection Commons, and the Toxicology Commons

Repository Citation

Mini Review

Jordan T. Perkins, Michael C. Petriello, Li Xu, Arnold Stromberg and Bernhard Hennig*

An open-sourced statistical application for identifying complex toxicological interactions of environmental pollutants

Abstract: The rising number of chemicals that humans are exposed to on a daily basis, as well as advances in biomonitoring and detection technologies have highlighted the diversity of individual exposure profiles (complex body burdens). To address this, the toxicological sciences have begun to shift away from examining toxic agents or stressors individually to focusing on more complex models with multiple agents or stressors present. Literature on interactions between chemicals is fairly limited in comparison with dose-response studies on individual toxicants, which is largely due to experimental and statistical challenges. Experimental designs capable of identifying these complex interactions are often avoided or not evaluated to their fullest potential because of the difficulty associated with appropriate analysis as well as logistical factors. To assist with statistical analysis of these types of experiments, an online, open-sourced statistical application was created for investigators to use to analyze and interpret potential toxicant interactions in laboratory experimental data using a full-factorial three-way analysis of variance (ANOVA). This model utilizes backward selection on interaction terms to model main effects and interactions.

Keywords: environment; pollutants; toxicological interactions.

*Corresponding author: Bernhard Hennig, 900 S. Limestone Street, Superfund Research Center, University of Kentucky, Lexington, KY 40536-0200, USA, Tel.: +1 859-218-1343; Fax: +1 859-257-1811; E-mail: bhennig@uky.edu; and Department of Animal and Food Sciences, University of Kentucky, Lexington, KY 40536, USA
Jordan T. Perkins and Michael C. Petriello: Superfund Research Center, University of Kentucky, Lexington, KY 40536-0200, USA; and Department of Animal and Food Sciences, University of Kentucky, Lexington, KY 40536, USA
Li Xu: Department of Statistics, University of Kentucky, Lexington, KY 40506, USA
Arnold Stromberg: Superfund Research Center, University of Kentucky, Lexington, KY 40536-0200, USA; and Department of Statistics, University of Kentucky, Lexington, KY 40506, USA

Introduction

Environmental pollutants in water, air, soil, and food, in addition to social factors like lifestyle, occupational hazard, and geography all contribute to an individual's toxicant exposure-response profile. The culmination of these factors means humans are concomitantly exposed to many different types of toxic agents at any given time [1]. This fact is further highlighted by the Chemical Substance Inventory, which currently contains over 80,000 different chemicals that are used in manufacturing [2].

Toxicology has historically focused on research examining the toxicity of individual agents, largely due to technological, statistical, and logistical limitations of proper study designs [3]. While these types of studies have been invaluable, especially in toxicology, and serve as a basis for setting exposure standards to protect human health, the ultimate application of these findings to realistic human exposures is limited by the attention to individual compounds. It has therefore been recognized that attention needs to be given to the effects of mixtures of agents. The National Institutes of Health, and in particular the National Institute of Environmental Health Sciences (NIEHS), have continually acknowledged the importance of mixture research and have emphasized the necessity of evaluating mixtures of exposures [3–5]. This national focus on mixture research, as well as improvements in analytical capabilities, has spurred emerging areas of research shifting away from single-agent exposures, to more complicated experimental designs involving multiple exposures.

As the necessity increases for toxicologists to shift away from implementing studies examining agents on an individual basis, several methods of risk assessment for multiple-pollutant exposures have been proposed and used. For example, dioxins, furans, and polychlorinated biphenyls, have historically had their toxicity expressed in relation to 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) using the toxic equivalency factor (TEF), which can be summed and used to express the toxicity of a mixture using the toxic equivalency (TEQ) [6]. Multiple other approaches for risk assessment, such as the maximum

cumulative ratio (MCR) [7], concentration-addition (CA) model [8, 9], and the independent-action model [10, 11] have been developed to analyze large mixtures of dissimilar compounds. These approaches are largely geared toward risk assessment in epidemiological settings or theoretical models, and are less appropriate for laboratory research.

In order to study the mechanisms of toxicity of multiple exposures in a laboratory setting, appropriate statistical tests must be used on data obtained from suitable experimental designs, specifically a full-factorial experimental design. The analysis and interpretation of these types of experiments often rely on statistical knowledge that requires additional training or support from a statistician. Frequently, these types of experiments are avoided, improperly analyzed, or not evaluated to their fullest potential because of a lack of familiarity with the proper statistical tests or the absence of statistical support. This can lead to missed opportunities at identifying complex and novel interactions that are occurring when multiple agents are present that can enhance or attenuate the toxicity of the agents. Therefore, the statistical application presented herein is intended to make these complex statistical analyses more approachable and efficient by providing a tool that allows for proper interpretation and analyses without the use of intricate statistical software.

Materials and methods

Required experimental design

In order to utilize the full-functionality model, a full-factorial experimental design must be used, where all the combinations of variables being present or absent are studied. For example, an experiment studying three agents would therefore have eight groups: a control, three individual exposures, three binary mixtures of exposures, and a tertiary mixture of all exposures. The model is capable of running analyses of two-way factorial as well as one-way designs as long as the higher order terms are omitted from the analyses. These experiments still require a full-factorial experimental design in order for the model to be utilized.

Statistical analysis

The model presented employs a full-factorial three-way analysis of variance (ANOVA), which is the most powerful design available for examining three treatment effects with possible interactions. The interactions are identified using this test and iterations of backwards elimination when appropriate. A knowledge or understanding of backwards elimination is not needed to utilize the open-sourced app; one needs to just follow the instructions on the Shiny App website (explained below). In general, the backwards elimination process is a step-wise progression that involves eliminations of interactions terms with non-significant p-values. It begins by the elimination of the third-order term (the interaction term that contains all three of
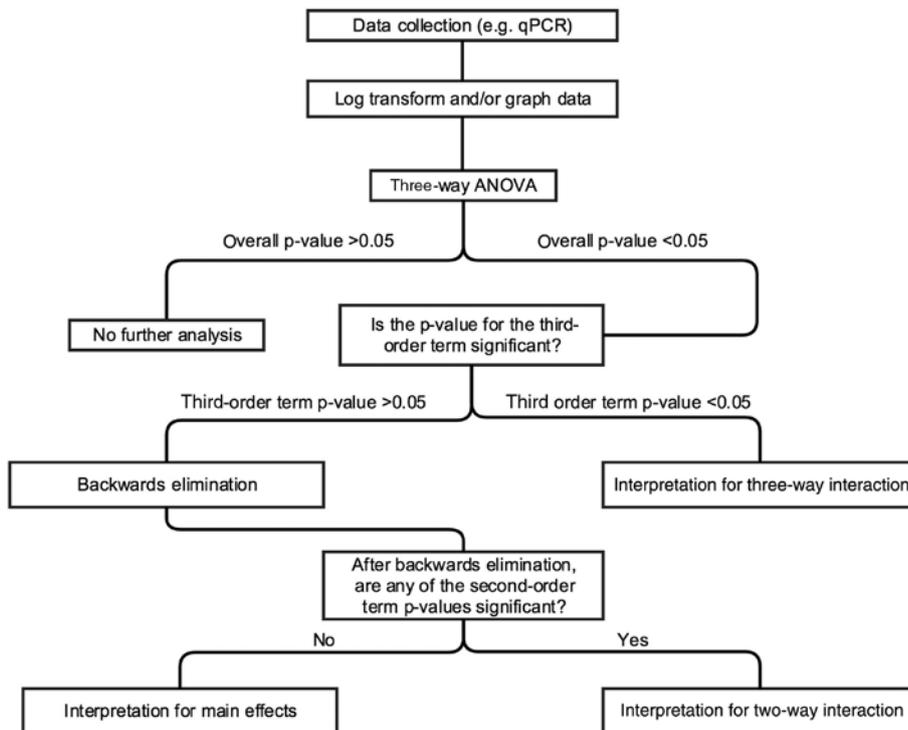

Figure 1: Flowchart describing the progression of analysis of a full-factorial experimental design using the proposed statistical application.

the variables of interest) from any subsequent statistical testing and rerunning the model. If any of the second-order terms (the interaction terms that contain any combination of two of the variables of interest) have a non-significant p-value, then the term with the largest p-value is removed, just as the third-order term was, and the model is run again. The process is repeated until all the second-order terms have significant p-values or no second-order terms remain in the model, leaving only the main effects (the terms containing only one of the variables of interest). This progression is summarized in Figure 1.

All statistical analyses and programming were carried out using the R language and environment for statistical computing (Vienna, Austria) [12]. It is suggested that, prior to any statistical analyses, the data be graphically represented and evaluated for normality. If normality issues are observed or suspected, then a $\log(x+1)$ transformation is recommended.

Shiny application

The open-source application can be accessed by visiting https://shiny.as.uky.edu/li/ and is capable of accepting the comma-separated values' file type (.csv) for file upload. Specific instructions for the use and interpretation of the findings using the program can be found on the linked PDF file on the application's interface.

Discussion

Large quantities of pollutants and therefore an equally large number of potential interactions ensure that mixture modeling and the subsequent statistical analysis will continue to prove challenging. To combat some of this difficulty, this proposed statistical application provides a resource for investigators who are interested in mixture research, especially in a laboratory setting with minimal statistical support to accurately analyze and assess interactions data. Specifically, it employs the three-way ANOVA to analyze the presence or absence of statistically significant interactions occurring in an accompanying full-factorial experimental setting. Most importantly, it is designed in a way that is easily usable and readily interpretable. As described above, this application can be accessed online and used at the discretion of the user. Additionally, the application also contains instructions for use, proper analysis, and interpretations of the findings, as well as sample data. Taken together, these documents should enable researchers to analyze their data for potential interactions.

Conclusion

The statistical application described above will attempt to address inadequate statistical support that may limit complete analysis and interpretation of experimental data that could potentially yield important information concerning interactions between toxicants or related agents. It is suggested that users utilize the documentation provided with the application as a guide to ensure proper interpretation of their results. Ultimately, this model should be used as a statistical tool, leaving the final explanation of the results to the discretion of the researchers. In summary, this application provides an open-sourced application capable of performing the statistical analysis required for mixture experiments, paired with accompanying documentation that explains the interpretation of the results.

References

1. Churchill JE, Ashley DL, Kaye WE. Recent chemical exposures and blood volatile organic compound levels in a large population-based sample. Arch Environ Health 2001;56:157–66.
2. US Environmental Protection Agency, Chemical Substance Inventory, U.E.P. Agency, Editor.
3. Suk WA, Olden K, Yang RS. Chemical mixtures research: significance and future perspectives. Environ Health Perspect 2002;110:(Suppl 6):891–2.
4. Carlin DJ, Rider CV, Woychik R, Birnbaum LS. Unraveling the health effects of environmental mixtures: an NIEHS priority. Environ Health Perspect 2013;121:A6–8.
5. Rider CV, Carlin DJ, Devito MJ, Thompson CL, Walker NJ. Mixtures research at NIEHS: an evolving program. Toxicology 2013;313:94–102.
6. Walker NJ, Crockett PW, Nyska A, Brix AE, Jokinen MP. Dose-additive carcinogenicity of a defined mixture of "dioxin-like compounds". Environ Health Perspect 2005;113:43–8.
7. Price PS, Han X. Maximum cumulative ratio (MCR) as a tool for assessing the value of performing a cumulative risk assessment. Int J Environ Res Public Health 2011;8:2212–25.
8. Berenbaum MC. The expected effect of a combination of agents: the general solution. J Theor Biol 1985;114:413–31.
9. Backhaus T, Faust M, Kortenkamp A. Cumulative risk assessment: a European perspective on the state of the art

and the necessary next steps forward. Integr Environ Assess Manag 2013;9:547–8.

10. Lam GK. A general formulation of the concept of independent action for the combined effects of agents. Bull Math Biol 1994;56:959–80.

11. Hadrup N, Taxvig C, Pedersen M, Nellemann C, Hass U, et al. Concentration addition, independent action and generalized concentration addition models for mixture effect prediction of sex hormone synthesis in vitro. PLoS One 2013;8:e70490.

12. R Core Development Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2016. ISBN 3-9000051-07-0, Available at: http://www.R-project.org/.