



Published in final edited form as:

Curr Environ Health Rep. 2018 March ; 5(1): 88–99. doi:10.1007/s40572-018-0180-5.

Cumulative Risk and Impact Modeling on Environmental Chemical and Social Stressors

Hongtai Huang, PhD, MSE^{1,2,§}, Aolin Wang, PhD^{1,2}, Rachel Morello-Frosch, PhD, MPH^{1,3}, Juleen Lam, PhD, MH, MHS¹, Marina Sirota, PhD^{2,4}, Amy Padula, PhD, MSc^{1,*}, and Tracey J. Woodruff, PhD, MPH^{1,*}

¹Program on Reproductive Health and the Environment, Department of Obstetrics, Gynecology & Reproductive Sciences, University of California, San Francisco, California, USA

²Institute for Computational Health Sciences, University of California, San Francisco, California, USA

³Department of Environmental Science, Policy, and Management, and the School of Public Health; University of California, Berkeley, California, USA

⁴Department of Pediatrics, University of California, San Francisco, California, USA

Abstract

Purpose of review—The goal of this review is to identify cumulative modeling methods used to evaluate combined effects of exposures to environmental chemicals and social stressors. The specific review question is: What are the existing quantitative methods used to examine the cumulative impacts of exposures to environmental chemical and social stressors on health?

Recent findings—There has been an increase in literature that evaluates combined effects of exposures to environmental chemicals and social stressors on health using regression models; very few studies applied other data mining and machine learning techniques to this problem.

[§]Corresponding author.

*These authors contributed equally to this work.

Hongtai Huang^{1, 2§}, PhD, MSE.

550 16th. Street, 7341, San Francisco, CA 94158. Hongtai.Huang@ucsf.edu

Aolin Wang^{1, 2}, PhD.

550 16th. Street, 7338, San Francisco, CA 94158. Aolin.Wang@ucsf.edu

Rachel Morello-Frosch^{1, 3}, PhD, MPH.

128B Giannini Hall, Berkeley, CA 94720. rmf@berkeley.edu

Juleen Lam¹, PhD, MS, MHS.

550 16th. Street, 7336, San Francisco, CA 94158. Juleen.Lam@ucsf.edu

Marina Sirota^{2, 4}, PhD.

550 16th. Street, 4761, San Francisco, CA 94158. Marina.Sirota@ucsf.edu

Amy Padula^{1,*}, PhD, MSc.

550 16th. Street, 7333, San Francisco, CA 94158. Amy.Padula@ucsf.edu

Tracey J. Woodruff^{1,*}, PhD, MPH.

550 16th. Street, 7330, San Francisco, CA 94158. Tracey.Woodruff@ucsf.edu

Compliance with Ethical Standards

Conflict of Interest

Hongtai Huang, Aolin Wang, Rachel Morello-Frosch, Juleen Lam, Marina Sirota, Amy Padula, and Tracey J. Woodruff declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent

This article does not contain any studies with human or animal subjects performed by any of the authors.

Summary—The majority of studies we identified used regression models to evaluate combined effects of multiple environmental and social stressors. With proper study design and appropriate modeling assumptions, additional data mining methods may be useful to examine combined effects of environmental and social stressors.

Keywords

Cumulative Risk; Combined Effects; Environmental Stressors; Non-chemical Stressors; Social Stressors; Quantitative Modeling

Introduction

Individuals are exposed to multiple environmental chemicals (both natural and synthetic) via different environmental media such as air, water and soil [1]. For instance, studies find that U.S. pregnant women are exposed to multiple chemicals including polychlorinated biphenyls (PCBs), organochlorine pesticides, perfluoroalkyl and polyfluoroalkyl substances (PFAS), phenols, polybrominated diphenyl ethers (PBDEs), phthalates, polycyclic aromatic hydrocarbons (PAHs), and perchlorate [2]. Human biomonitoring – measuring the concentration of chemicals in body fluids (blood, urine, and breast milk) or tissues (hair, nails, fat, and bone) – is often used to assess chemical burden as it provides evaluation of the internal doses reflective of exposures via multiple pathways [3]. Since 2000, the U.S. Center for Disease Control and Prevention (CDC) has been biomonitoring about 300 chemicals, using the National Health and Nutrition Examination Survey (NHANES). These chemicals include metals, pesticides, PCBs, PBDEs, PFAS, volatile organic compounds (VOCs), tobacco smoke, PAH metabolites, phthalate and metabolites, among many others [4]. Certain classes are frequently measured simultaneously and in a single matrix (maternal urine or maternal serum): non-persistent phenols and phthalates are often examined in urine while persistent chemicals such as PFAS, PBDEs, PCBs and organochlorine pesticides are commonly measured in serum. Multiple chemical exposures can result in higher risks than exposures to individual pollutants. The NAS concluded that ‘combinations of phthalates and of other antiandrogens generate combined effect at doses that when administered alone do not have significant effects’ and recommend that risk assessment should account for cumulative risk to chemicals that affect the same adverse health endpoint [5].

The concept of cumulative impacts, the focus of this review, refers to potential adverse human health effects resulting from combined exposures to multiple environmental and social stressors [6–8]. ‘Cumulative risk’ aims to quantify to the extent possible ‘combined risks from aggregate exposures to multiple [environmental] agents or stressors’ [1]. Statistical methods used to characterize and model the combined and potential interactive effects of multiple environmental hazard and social stressor exposures are referred to as cumulative risk and impact modeling.

Recognizing the public health impacts of exposure to multiple environmental chemicals, an increasing number of studies have assessed the cumulative impact of exposures to chemical mixtures or multiple air pollutants simultaneously. Dose-addition based methods including relative potency factors and toxic equivalency factors have been used to examine the

cumulative risk of chemicals from a single class such as organophosphate pesticides [9], PCBs [10], and phthalates [11].

In addition to environmental chemical pollutants, non-chemical stressors – particularly psychological and social stressors (e.g. such as poverty, lack of social support and chronic discrimination due to race/ethnicity) – can independently influence health and have been considered in cumulative risk and cumulative impact studies [12]. Psychological stress and socioeconomic factors in recent years have been identified as critical non-chemical stressors that could increase the adverse health effects of chemical exposures. For example, biomarkers of chronic stress response, such as ‘allostatic load’, were found to amplify the risk of increased blood pressure associated with lead exposure in adults [13]. Urban children exposed to violence had higher risks of developing asthma in the presence of traffic-related air pollution [14]. Social stressors, measured by indicators such as poverty and race/ethnicity, have been often included as one of the key effect modifiers in environmental health research to address disparities [15–22]. Social stressors such as educational attainment level and population density were examined previously as well [20]. To evaluate cumulative health risks from both chemical and non-chemical stressors, U.S. EPA proposed a framework guidance for cumulative risk assessment in 2003 [1] and subsequently provided a technical resource document in 2007 [23], acknowledging the challenges of incorporating non-chemical stressors in risk assessment. Although it is known that humans are exposed to multiple chemical and social stressors which are likely to cumulatively impact health, cumulative risk and impact modeling methods have not yet been fully developed to evaluate the joint exposures.

Overall, the main categories of established approaches used to evaluate aspects of cumulative impacts of multiple stressors are either quantitative or semi-quantitative methods, including biomonitoring, health risk assessment, ecological risk assessment, health impact assessment, burden of disease, and mapping of cumulative impacts [24]. The majority of the established approaches involve elements of quantitative analysis, but these vary to a large degree [24]. For example, when little or no mechanistic data are available, the hazard index was used to assess the cumulative non-cancer risks for chemicals that have an established chronic reference dose or reference concentration [25]. More established modeling methods have been applied in recent years to attempt to address cumulative impacts. Air dispersion and exposure models were employed to examine cumulative diesel particulate matter emission in Southern California for five traffic/mobile sources comparing four environmental goals including impact, efficiency, quality and justice [26]. To account for non-chemical stressors in the process of exposure and dose estimates, the Average Daily Dose model [27], was linked to multiple social indicators and applied to examine dose estimates on both the U.S. nationwide census tract-level and community-wise local scale [28]. In addition, association rule mining [29], an unsupervised machine learning method, was also utilized to evaluate associations between social factors and environmental chemical concentrations relevant to cumulative impacts in the U.S. [30].

To our knowledge, no review has yet been performed to identify the statistical models used to evaluate the combined effects of multiple factors of environmental chemical and social stressors. A review of cumulative risk and impact modeling techniques can fill this scientific

gap and provide useful modeling reference for future cumulative risk and impact studies. This review identifies and evaluates the types of statistical models used to quantify the cumulative effects of multiple environmental and social stressors to provide modeling suggestions. The purpose of this review is not to provide a framework for modeling selection in designing a scientific study, but to summarize what modeling techniques have been considered after research questions, study design and data were determined. Many factors including research questions, study design and data availability are important in model selection. However, there are often multiple choices for statistical modeling and there may be opportunities to bring in new types of models into the field to evaluate cumulative risk and impacts. This review was conducted from this perspective.

Methods

Given the diversity of chemical and non-chemical factors involved in cumulative impacts and their various possible combinations and relevant health effects, different studies used a variety of statistical approaches to model and evaluate the health effects from multiple stressors. The two major categories of statistical models are supervised and unsupervised modeling methods. The former predefines response and explanatory variables, and evaluates their statistical relationships, while the latter has no such predetermined condition but instead examines and identifies potential associations or hidden statistical structure among different input variables. Supervised methods include both regression models (e.g. Cox's regression model [31]) and classification models (e.g. Classification and Regression Trees [32]). Unsupervised approaches encompass cluster analysis [33] and association rule mining or frequent itemset mining.

In this study, we review the statistical models used in studies whose primary objective was to analyze chemical and non-chemical stressors collectively. Many exposure studies have varying interpretations of the concept of 'environment'— for instance, characterizing exposures in the home, work, or neighborhood environments. Similar to the definition presented in a previous review [34], the universe of exogenous chemical exposures in this review is referred to 'those that are generally addressed by U.S. EPA, and include manufactured chemicals and chemical byproducts (e.g. air pollution)' [34] except smoking. In this review, we did not evaluate studies that were specific to home or work environments. Also, no restriction was imposed upon our search based on the type of data used.

We also utilized searching terms related to environmental justice to broadly capture articles that evaluated the health effects of multiple chemical and non-chemical stressors in a cumulative manner, in that many environmental justice studies emphasized the combined effects of multiple stressors. However, environmental justice is not the main focus of this review.

We searched articles published 2012/01/01 – 2017/06/21 in English and indexed in PubMed with the following four groups of searching terms and identified articles that had to meet each of the four criteria:

- #1 (cumulative, multiple, aggregated, joint, combined) AND (risk, impact, exposure)

- #2 (Environmental) AND (justice, injustice, equality, inequality, equity, inequity, disparity)
- #3 Environmental/Chemical exposure*
- #4 Non-chemical stressors*

*Search terms for #3 and #4 were adopted from protocols developed by [34] that evaluated the combined effects of prenatal-exposure to both chemicals and psychological stress on fetal growth.

Inclusion criteria

- I. Original peer-reviewed research articles that evaluate both environmental and social stressors, and analyzed their health effects (excluding home or work environment)
- II. Human subject studies
- III. Articles published during 2012/01/01 and 2017/06/21
- IV. Articles that included quantitative method information

One reviewer (HH) was responsible for screening and identifying relevant studies.

Results

HH identified 376 articles with full text availability and found 79 eligible articles based on initial title and abstract screening. After full-text review of these eligible articles, HH identified 31 relevant articles. The excluded 345 references consisted of: 1) articles that did not involve both social stressors and environmental/chemical exposures (n=241); 2) studies that did not analyze health effects of multiple stressors (n=69); 3) non-original research articles (n=24); 4) articles that did not use quantitative methods (n=9); and 5) studies that were not based on human subjects (n=2).

Supervised methods were divided into regression and classification (Figure 1). Currently, most of the modeling techniques utilized to examine cumulative impact are supervised regression models. We considered commonly used regression methods such as multivariable linear/non-linear regression and logistic regression models as simple regression techniques. Other regression models, such as generalized linear model (GLM), multilevel model and spatial regression model, were classified as complex regression techniques. None of the studies identified in this review used supervised classification models.

As shown in Table 1, among the 31 articles identified [35–70], 10 studies [37, 42, 43, 46, 52, 55, 61, 66, 68, 69] used multivariable linear/non-linear regression models to evaluate the combined effects of multiple chemical and non-chemical stressors, and 7 studies [36, 38, 40, 45, 47, 59, 68] used logistic regression models. In addition, we found 5 studies [49, 56, 57, 63, 67] that used hierarchical/multilevel regression models. All studies used supervised techniques, specifically regression models, but several of them also used unsupervised methods such as hierarchical cluster analysis, factor analysis and principal component analysis (PCA), in addition to regression models. Air pollutants were the environmental

chemical exposure most often modeled: 20 out of the 31 studies evaluated air pollutants [35, 38, 40, 42, 43, 45, 49–53, 56, 58, 60–64, 66, 69], especially the criteria pollutants such as particulate matter (PM) and nitrogen dioxide (NO₂). For instance, the joint effects of exposure to PM_{2.5} and O₃ and socioeconomic status measures upon pregnancy outcomes including low birth weight, preterm birth and small for gestational age were evaluated based on a linear or logistic mixed regression model [49]. Other chemical exposures evaluated cumulatively include industrial cadmium [55], lead [46], zinc [67], Bisphenol A (BPA) [47] and PAH [68]. Socioeconomic factors, especially race/ethnicity and income level, were among the most frequent non-chemical stressors modeled. As to the health endpoints evaluated, mortality rate and cancer risks followed by pregnancy outcomes were considered more frequently than others (7 and 5 out of 31) (Table 1).

We evaluated in more detail the statistical modeling technique and combination of exposures and outcomes identified in Table 1 to provide examples of model selection given different stressors focused (Table 2). We found that some complex regression-based modeling approaches have been used to evaluate the joint effects of multiple stressors. For example, the combined effects of exposure to arsenic contamination in drinking water and health intervention programs on child mortality from acute lower respiratory infections were modeled by a zero-inflated negative binomial regression [54]. Both simple regression and Bayesian sparse spatial multilevel models were utilized to evaluate the relationship between lead exposure and both gonorrhoea and chlamydia, accounting for other non-chemical stressors such as index of concentrated disadvantage [57]. Negative binomial regression models were used to examine associations between mortality and environmental factors including air pollution and drinking water quality with consideration to socioeconomic deprivation [62].

Discussion

Our review provided a summary of statistical modeling methods considered in studies to quantify potential combined effects of multiple environmental and social stressors. It was not our intent to provide a framework for modeling selection in designing a scientific study, which is beyond the scope of our review. The selection of modeling technique involves consideration of many important factors, including the research question(s), study design, and data availability. The combination of these factors may lead an investigator to choose one method over another.

For example, to answer respiratory health inequalities questions concerning relationships between respiratory health situations across different cities and their medical amenities, socioeconomic and physical features (e.g. air pollution), Aschan-Leygonie *et al.* analyzed health data describing hospitalizations of chronic obstructive pulmonary disease and a large set of different indicator variables of both social and environmental stressors using linear correlations and multiple linear regression models in an ecological case study [37]. They found that socioeconomic features may be the major drivers for inequities of respiratory health status in urban units and concluded that better understanding of ‘differences among cities in their entirety’ is essential to develop effective urban policies. Multiple linear regression models were an appropriate choice to answer the research questions of interest

given their study design and data availability. However, if the intent of the study was not to understand risk factors for a specific type of disease or outcome, but instead to identify all the possible associations among the variables with different combination to provide guidance for further analysis, then the association rule mining method could have been more useful in that setting, assuming the modeling assumptions were met and data requirements satisfied.

We found that a large number of more complex statistical models, supervised or unsupervised, have been utilized in other scientific domains but less commonly applied in cumulative impact studies. A relevant example is the Random Forest model [71] that is a supervised classification tool widely employed in various applications such as compound classification and quantitative structure-activity relationship (QSAR) modeling for predicting categorical biological activity [72], land-cover classification [73] and gene selection and classification [74], but was not used to understand the joint effects of multiple stressors. When multiple exposures including social stressors were considered, Random Forest model can be rather useful to determine variable importance and potentially separate those of more importance from a larger set of variables. Another supervised technique example is neural network ensembles [75], which serves as the foundation of deep learning [76] and has also been extensively employed in numerous scientific disciplines [77–81], but not been used in cumulative impact studies. Provided that there are known associations between certain health outcome and multiple stressors, this type of models can be very powerful in predicting occurrences of the health outcome of interest within the context of examining multiple exposures.

Models Comparison

Simple Regression vs. Complex Regression—The advantages of simple regression models include: 1. straightforward model execution; 2. easy interpretation of model techniques; 3. model outputs accessible and understandable to a larger group of audiences, which can be conducive to risk communication and community engagement. For example, Vishnevetsky *et al.* evaluated cumulative effects of low socioeconomic status and PAH air pollution exposure in children. They found that children with socioeconomic disadvantages as measured by recurrent material hardship and high level of prenatal exposure to PAH had 5.81 point lower full scale intellectual quotient (IQ) score than those experienced material hardship but with low PAH exposure. The same significant association was not observed within the low material hardship group [68]. The numeric construct of the findings is informative and understandable to public audiences and will benefit future community-based risk assessment and communication.

One of the challenges of using simple regression models is that normal or multi-normal distribution assumptions may not always hold for environmental exposures or social factors. Although in some cases log-normal transformation of response variables can address this problem, it cannot be used to account for data that has other types of distributions, such as negative binomial distribution, Poisson distribution and gamma distribution.

Complex regression models have the following benefits: 1. higher level of flexibility regarding distribution assumptions made; 2. the ability to account for inherent data issues

(e.g. spatial autocorrelation); 3. potentially better predictive power. For instance, GLM allows flexibility regarding data distribution assumption which is the advantage of this kind of modeling [33], but similar to other supervised methods, it also requires specification of both response and explanatory variables.

When modeling the joint effects of multiple environmental chemical and social stressors, using multivariable linear regression or logistic regression models may not be appropriate, especially if the social stressors are based on placed-based measurement and these can be spatially autocorrelated which can introduce biased estimates. In these cases, more complex regression modeling method such as spatial error regression models [82] can be useful in addressing such issue, because they do not assume independent and identically distributed errors at the census tract level, but rather allow errors distributed by a spatial autoregressive process. This type of model can account for residual spatial autocorrelation when units of observation are located proximally, and thus non-independently, in space. For example, simultaneous autoregressive models were utilized to examine the health impacts of NO₂ and several community-level social stressors such as violent crime and physical disorder, crowding and poor access to resources across New York communities, accounting for spatial relationship between air pollution and social stressors [64].

Simple regression models such as multivariable linear or logistic regression models are special cases of GLM. In this review, we distinguish between multivariable linear or logistic regression and other uses of GLM models beyond linear/logistic regression models. Multilevel/hierarchical modeling [83] is another example of a modeling approach that permits examination of the effects of stimulus variables upon response variable on the local versus global scale, accounting for variance among variables at different levels, but it requires a sufficient sample size for unbiased estimation [84]. In this review, we found that multi-level models were a popular option for modeling cumulative effects of various stressors that have nested effects and frequently considered in longitudinal studies. For example, association between PM_{2.5} and O₃, socioeconomic factors and birth outcomes were modeled using North Carolina birth data from 2002 to 2006 and multi-level models in which census tract was specified as a random effect to account for neighborhood-level correlation [49].

Although unsupervised methods such as PCA was employed occasionally, coupled with other supervised techniques (e.g. generalized estimating equations, zero-inflated negative binomial regression), other complex unsupervised machine learning methods have not yet been explored. In recent years, association rule mining modeling was used to identify and prioritize relations between environmental stressors and negative human health effects [85] and discern prevalent chemical combinations in the U.S. population [86]. Provided that population-based health outcome information is available on a small geographic unit such as census tract and can be linked to social and environmental data, this unsupervised model can also be applied to evaluate the synergistic health impacts of multiple chemical and non-chemical stressors in the future.

It should be acknowledged that using complex data mining techniques are not necessarily always a better option. For example, more complex modeling methods may have higher data

input requirement such as larger sample size to have enough statistical power for generating reliable results. However, with proper study design and data availability, outputs from certain data mining techniques can be useful to address certain research questions, provided proper assumptions are met.

Bayesian vs. Non-Bayesian—We also found that several studies adopted a Bayesian approach to examine combined effects of multiple stressors. The advantage of Bayesian methods is that they can incorporate qualitative information or *a priori* knowledge to improve model fitting and predictions. Incorporating feedback from local residents or experts is a critical component in local scale cumulative risk assessment, and Bayesian statistical models could play a key role in connecting qualitative information to quantitative calculation. However, both the amount of non-quantitative data needed to integrate into the model and the degree to which information will be applied are subjective, which may potentially introduce bias. Non-Bayesian approaches predominantly utilize quantitative information without qualitative inputs, and therefore could avoid subjective bias embedded in qualitative data. However, these approaches have less flexibility in integrating non-quantitative information such as expert opinions that could be potentially useful in situations where quantitative data alone are not sufficient for conclusive analysis.

Current and Emerging Exposures

More than two thirds of the studies identified focused on air pollutants, which is largely driven by data availability and response to policies developed as part of implementation of the Clean Air Act and subsequent regulations. Importantly, spatial studies, such as those that use air pollution, make it feasible to analyze place-based exposures to environmental exposure and social stressors because these data can be accessed via publicly available data sources, such as the National Air Toxics Assessment database (<https://www.epa.gov/national-air-toxics-assessment/2011-nata-assessment-results>). This facilitates data analysis with a larger geographic and population scope, which provides sufficient power to observe signals for multiple types of groups, each of which themselves may be relatively small. There were fewer studies that used biomonitoring data, in part because this type of data is less available, though more is being generated with investments from the research community and government. Thus, studies that characterize individual-level exposures to both multiple environmental chemicals (using targeted and non-targeted approaches) and social stressors (biomarkers of chronic stress response, perceptual and place-based stressors) are also becoming more viable.

There exist many emerging exposures that warrant researchers' and policy-makers' attention such as heat exposure [87], multimedia screen light exposure [88], nanomaterial exposure [89], chemicals not bio-monitored previously [4, 90] and poor access to resources [64]. Analyses based on place-based measures allow researchers to take advantage of data sets that are likely to have wide variability in exposures to chemical and non-chemical stressors, and facilitate research with a wide geographic scope. However, it can narrow the scope of the kinds of exposures that one can analyze due to constraints of data availability. This makes it more challenging to evaluate these emerging exposures. Provided relevant data sets are available, future research can build upon the methods and findings from spatial studies

and apply them to evaluating combined effects of some of these emerging exposures with known pollutants, to better characterize the effects of multiple chemical exposures that individuals experience [90] along with social stressors.

Limitations

There are several limitations of this review that warrant further consideration. First, this review mainly focused on the modeling aspects of research studies and less so regarding specific research questions evaluated, study design and data available. However, due to the high degree of dimensions regarding different possible research settings that may call for use of distinct modeling methods, covering all the combinations of different research questions, study designs and data, and then proposing modeling suggestions accordingly is beyond scope of this review. As a starting point towards promoting use of appropriate statistical methods in examining cumulative risk, we focus on the modeling perspective. Therefore, this manuscript reviews statistical models used to examine the combined effects of both environmental chemical and social stressors in recent studies. Lastly, although our intent was to focus on studies whose primary objective was to investigate cumulative exposure, our title and abstract searching mechanism might have missed studies that found no positive results regarding the combined effects of multiple environmental chemical and social stressors. Such negative findings would be important in a systematic review and meta-analysis. However, it should be recognized that this review is not a systematic review and no quantitative synthesis was performed across different studies. Therefore, we may have potentially excluded references containing other useful data mining techniques not mentioned in this review, but we estimate that the negative consequence is not substantial.

Regression models have been often applied to evaluate the potential adverse human health effects from combined exposures to multiple environmental chemical and social stressors. With proper study design and appropriate modeling assumptions, additional data mining methods may be useful in the evaluation of cumulative health impacts of multiple chemical exposures and social stressors.

Conclusion

The importance of understanding joint effects of environmental chemical and social stressors has been recognized. There is growing literature to evaluate the combined effects of multiple stressors on health with the majority of them using regression models. With increasing knowledge in exposure science and the advent of more quantitative tools in the era of ‘big data’, we recommend that additional data mining techniques are considered in certain appropriate research settings and potentially incorporated in the analytical procedure to better characterize chemical and non-chemical stressors for risk assessment to identify potential health risks and to provide public health protection, particularly to the vulnerable and susceptible populations.

Acknowledgments

This work is supported in part by the NIEHS grants R00ES021470 [AP, HH], P01ES022841 and R01ES027051, the U.S. EPA grants RD-83564301 and RD-83543301 [TJW, RMF, AW], NLM grant K01LM012381 [MS, HH and

AW], Preterm Birth Initiative at UCSF [TJW, AP, MS and HH], the March of Dimes Prematurity Research Center at Stanford [MS and AW] and Burroughs Wellcome Fund [MS].

We thank Drs. Marc Weisskopf and Zeyan Liew for their comments and suggestions.

References

Papers of particular interest, published recently, have been highlighted as:

* Of importance

** Of major importance

1. EPA, US., Environmental Protection Agency. Framework for Cumulative Risk Assessment. US EPA, National Center for Environmental Assessment; Washington, DC: 2003. EPA/600/P-02/001FAvailable: https://www.epa.gov/sites/production/files/2014-11/documents/frmwrk_cum_risk_assmnt.pdf
2. Woodruff TJ, Zota AR, Schwartz JM. Environmental Chemicals in Pregnant Women in the United States: NHANES 2003-2004. *Environmental Health Perspectives*. 2011; 119(6):878–85. [PubMed: 21233055]
3. National Research Council (U.S.). Human biomonitoring for environmental chemicals. Washington, DC: National Academies Press; 2006. Committee on Human Biomonitoring for Environmental Toxicants; p. xxix. 291
4. U.S. CDC (Centers for Disease Control and Prevention). The Fourth National Report on Human Exposure to Environmental Chemicals: Updated Tables, January 2017. Available: <https://www.cdc.gov/exposurereport/index.html2017>
5. National Research Council (U.S.). Committee on the Health Risks of Phthalates., National Academies Press (U.S.). Phthalates and cumulative risk assessment : the task ahead. Washington, D.C.: National Academies Press; 2008. p. xixp. 188
- 6*. McHale CM, Osborne G, Morello-Frosch R, Salmon AG, Sandy MS, Solomon G, et al. Assessing health risks from multiple environmental stressors: Moving from $G \times E$ to $I \times E$. *Mutation Research/Reviews in Mutation Research*. 2018; 775:11–20. January–March 2018.
7. deFur PL, Evans GW, Hubal EAC, Kyle AD, Morello-Frosch RA, Williams DR. Vulnerability as a function of individual and group resources in cumulative risk assessment. *Environ Health Persp*. 2007; 115(5):817–24.
8. Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding The Cumulative Impacts Of Inequalities In Environmental Health: Implications For Policy. *Health Affair*. 2011; 30(5):879–87.
9. U.S. EPA (Environmental Protection Agency). Organophosphorus Cumulative Risk Assessment (2006 Update). Available: http://www.epa.gov/pesticides/cumulative/pr_a_op_methods.htm2006
10. Gennings C, Sabo R, Carney E. Identifying Subsets of Complex Mixtures Most Associated With Complex Diseases. *Epidemiology*. 2010; 21:S77–S84. [PubMed: 21422968]
11. Varshavsky JR, Zota AR, Woodruff TJ. A Novel Method for Calculating Potency-Weighted Cumulative Phthalates Exposure with Implications for Identifying Racial/Ethnic Disparities among U.S. Reproductive-Aged Women in NHANES 2001-2012. *Environmental Science and Technology*. 2016; 50:10616–24. [PubMed: 27579903]
12. Sexton K, Linder SH. Cumulative risk assessment for combined health effects from chemical and nonchemical stressors. *American journal of public health*. 2011; 101(S1):81–8.
13. Zota AR, Shenassa ED, Morello-Frosch R. Allostatic load amplifies the effect of blood lead levels on elevated blood pressure among middle-aged US adults: a cross-sectional study. *Environ Health-Glob*. 2013:12.
14. Clougherty JE, Levy JI, Kubzansky LD, Ryan PB, Suglia SF, Canner MJ, et al. Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology. *Environmental health perspectives*. 2007; 115(8):1140–6. [PubMed: 17687439]

15. Evans AM, Rice GE, Teuschler LK, Wright JM. Joint exposure to chemical and nonchemical neurodevelopmental stressors in U.S. women of reproductive age in NHANES. *Int J Environ Res Public Health*. 2014; 11(4):4384–401. [PubMed: 24758893]
16. Perera FP, Rauh V, Whyatt RM, Tsai W-Y, Bernert JT, Tu Y-H, et al. Molecular evidence of an interaction between prenatal environmental exposures and birth outcomes in a multiethnic population. *Environmental Health Perspectives*. 2004; 112(5):626–30. [PubMed: 15064172]
17. Apelberg BJ, Buckley TJ, White RH. Socioeconomic and racial disparities in cancer risk from air toxics in Maryland. *Environmental Health Perspectives*. 2005; 113(6):693–9. [PubMed: 15929891]
18. Bell ML, Ebisu K. Environmental inequality in exposures to airborne particulate matter components in the United States. *Environmental health perspectives*. 2012; 120(12):1699–704. [PubMed: 22889745]
19. Downey L, Hawkins B. Race, income, and environmental inequality in the United States. *Sociol Perspect*. 2008; 51(4):759–81. [PubMed: 19578560]
20. Pastor M, Morello-Frosch R, Sadd JL. The air is always cleaner on the other side: Race, space, and ambient air toxics exposures in California. *Journal of Urban Affairs*. 2005; 27:127–48.
21. Perlin SA, Wong D, Sexton K. Residential proximity to industrial sources of air pollution: Interrelationships among race, poverty, and age. *Journal of the Air & Waste Management Association*. 2001; 51(3):406–21. [PubMed: 11266104]
22. Woodruff TJ, Parker JD, Kyle AD, Schoendorf KC. Disparities in exposure to air pollution during pregnancy. *Environmental Health Perspectives*. 2003; 111(7):942–6. [PubMed: 12782496]
23. U.S. EPA (Environmental Protection Agency). *Concepts, Methods and Data Sources for Cumulative Health Risk Assessment of Multiple Chemicals, Exposures and Effects: A Resource Document*. U.S. EPA, National Center for Environmental Assessment; Cincinnati, OH: 2007. EPA/600/R-06/013F
- 24*. Solomon GM, Morello-Frosch R, Zeise L, Faust JB. *Cumulative Environmental Impacts: Science and Policy to Protect Communities*. *Annu Rev Publ Health*. 2016; 37:83–96.
25. Callahan MA, Sexton K. If cumulative risk assessment is the answer, what is the question? *Environ Health Persp*. 2007; 115:799–806.
26. Marshall JD, Swor KR, Nguyen NP. Prioritizing environmental justice and equality: diesel emissions in southern California. *Environ Sci Technol*. 2014; 48(7):4063–8. [PubMed: 24559220]
27. U.S. EPA (Environmental Protection Agency). *Exposure Factors Handbook*. U.S. EPA, National Center for Environmental Assessment; Washington, DC: 2011. EPA/600/R-09/052FAvailable: <http://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=236252> [accessed 29 December 2015] 2011
28. Huang H, Barzyk TM. Connecting the Dots: Linking Environmental Justice Indicators to Daily Dose Model Estimates. *Int J Environ Res Public Health*. 2016; 14(1)doi: 10.3390/ijerph14010024
29. Borgelt C. Frequent item set mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2012; 2(6):437–56.
30. Huang H, Tornero-Velez R, Barzyk TM. Associations between socio-demographic characteristics and chemical concentrations contributing to cumulative exposures in the United States. *J Expo Sci Environ Epidemiol*. 2017; doi: 10.1038/jes.2017.15
31. Andersen PK, Gill RD. Cox Regression-Model for Counting-Processes - a Large Sample Study. *Ann Stat*. 1982; 10(4):1100–20.
32. Breiman, L. *Classification and regression trees*. Belmont, Calif: Wadsworth International Group; 1984. p. xpi. 358
33. Hastie, T., Tibshirani, R., Friedman, JH. *The elements of statistical learning : data mining, inference, and prediction*. 2nd. New York, NY: Springer; 2009. p. xxii. 745
- 34*. Vesterinen HM, Morello-Frosch R, Sen S, Zeise L, Woodruff TJ. Cumulative effects of prenatal-exposure to exogenous chemicals and psychosocial stress on fetal growth: Systematic-review of the human and animal evidence. *PloS one*. 2017; 12(7):e0176331. [PubMed: 28700705]
35. Al-Wahaibi A, Zeka A. Health impacts from living near a major industrial park in Oman. *BMC public health*. 2015; 15:524. [PubMed: 26032629]
36. Allen EM, Alexander BH, MacLehose RF, Nelson HH, Ryan AD, Ramachandran G, et al. Occupational exposures and lung cancer risk among Minnesota taconite mining workers. *Occupational and environmental medicine*. 2015; 72(9):633–9. [PubMed: 25977445]

37. Aschan-Leygonie C, Baudet-Michel S, Mathian H, Sanders L. Gaining a better understanding of respiratory health inequalities among cities: an ecological case study on elderly males in the larger French cities. *International journal of health geographics*. 2013; 12:19. [PubMed: 23575258]
38. Balmes JR, Cisternas M, Quinlan PJ, Trupin L, Lurmann FW, Katz PP, et al. Annual average ambient particulate matter exposure estimates, measured home particulate matter, and hair nicotine are associated with respiratory outcomes in adults with asthma. *Environmental research*. 2014; 129:1–10. [PubMed: 24528996]
39. Best EA, Juarez-Colunga E, James K, LeBlanc WG, Serdar B. Biomarkers of Exposure to Polycyclic Aromatic Hydrocarbons and Cognitive Function among Elderly in the United States (National Health and Nutrition Examination Survey: 2001-2002). *PloS one*. 2016; 11(2):e0147632. [PubMed: 26849365]
40. Bravo MA, Son J, de Freitas CU, Gouveia N, Bell ML. Air pollution and mortality in Sao Paulo, Brazil: Effects of multiple pollutants and analysis of susceptible populations. *Journal of exposure science & environmental epidemiology*. 2016; 26(2):150–61. [PubMed: 25586330]
41. Candido da Silva AM, Moi GP, Mattos IE, Hacon Sde S. Low birth weight at term and the presence of fine particulate matter and carbon monoxide in the Brazilian Amazon: a population-based retrospective cohort study. *BMC pregnancy and childbirth*. 2014; 14:309. [PubMed: 25193316]
42. Chen JC, Wang X, Wellenius GA, Serre ML, Driscoll I, Casanova R, et al. Ambient air pollution and neurotoxicity on brain structure: Evidence from women's health initiative memory study. *Annals of neurology*. 2015; 78(3):466–76. [PubMed: 26075655]
43. Deguen S, Petit C, Delbarre A, Kihal W, Padilla C, Benmarhnia T, et al. Neighbourhood Characteristics and Long-Term Air Pollution Levels Modify the Association between the Short-Term Nitrogen Dioxide Concentrations and All-Cause Mortality in Paris. *PloS one*. 2015; 10(7):e0131463. [PubMed: 26197409]
44. Domazet SL, Grontved A, Timmermann AG, Nielsen F, Jensen TK. Longitudinal Associations of Exposure to Perfluoroalkylated Substances in Childhood and Adolescence and Indicators of Adiposity and Glucose Metabolism 6 and 12 Years Later: The European Youth Heart Study. *Diabetes care*. 2016; 39(10):1745–51. [PubMed: 27489335]
45. Dominguez-Cortinas G, Cifuentes E, Escobar ER, Martinez FD. Assessment of environmental health children's population living in environmental injustice scenarios. *Journal of community health*. 2012; 37(6):1199–207. [PubMed: 22418761]
46. Filigrana PA, Mendez F. Blood lead levels in schoolchildren living near an industrial zone in Cali, Colombia: the role of socioeconomic condition. *Biological trace element research*. 2012; 149(3): 299–306. [PubMed: 22547322]
47. Findlay LC, Kohen DE. Bisphenol A and child and youth behaviour: Canadian Health Measures Survey 2007 to 2011. *Health reports*. 2015; 26(8):3–9.
48. Genowska A, Jamiolkowski J, Szafraniec K, Stepniak U, Szpak A, Pajak A. Environmental and socio-economic determinants of infant mortality in Poland: an ecological study. *Environmental health : a global access science source*. 2015; 14:61. [PubMed: 26195213]
49. Gray SC, Edwards SE, Schultz BD, Miranda ML. Assessing the impact of race, social factors and air pollution on birth outcomes: a population-based study. *Environmental health : a global access science source*. 2014; 13(1):4. [PubMed: 24476365]
50. Grineski SE, Collins TW, Morales DX. Asian Americans and disproportionate exposure to carcinogenic hazardous air pollutants: A national study. *Soc Sci Med*. 2017; 185:71–80. [PubMed: 28554161]
51. Hart JE, Kallberg H, Laden F, Costenbader KH, Yanosky JD, Klareskog L, et al. Ambient air pollution exposures and risk of rheumatoid arthritis. *Arthritis care & research*. 2013; 65(7):1190–6. [PubMed: 23401426]
52. Hicken MT, Adar SD, Diez Roux AV, O'Neill MS, Magzamen S, Auchincloss AH, et al. Do psychosocial stress and social disadvantage modify the association between air pollution and blood pressure?: the multi-ethnic study of atherosclerosis. *American journal of epidemiology*. 2013; 178(10):1550–62. [PubMed: 24064742]

53. James P, Hart JE, Banay RF, Laden F. Exposure to Greenness and Mortality in a Nationwide Prospective Cohort Study of Women. *Environmental health perspectives*. 2016; 124(9):1344–52. [PubMed: 27074702]
54. Jochem WC, Razzaque A, Root ED. Effects of health intervention programs and arsenic exposure on child mortality from acute lower respiratory infections in rural Bangladesh. *International journal of health geographics*. 2016; 15(1):32. [PubMed: 27586497]
55. Kippler M, Tofail F, Hamadani JD, Gardner RM, Grantham-McGregor SM, Bottai M, et al. Early-life cadmium exposure and child development in 5-year-old girls and boys: a cohort study in rural Bangladesh. *Environmental health perspectives*. 2012; 120(10):1462–8. [PubMed: 22759600]
56. Li L, Laurent O, Wu J. Spatial variability of the effect of air pollution on term birth weight: evaluating influential factors using Bayesian hierarchical models. *Environmental health : a global access science source*. 2016; 15:14. [PubMed: 26850268]
57. Nelson EJ, Shacham E, Boutwell BB, Rosenfeld R, Schootman M, Vaughn M, et al. Childhood lead exposure and sexually transmitted infections: New evidence. *Environ Res*. 2015; 143(Pt A): 131–7. [PubMed: 26479187]
58. Osiecki KM, Kim S, Chukwudozie IB, Calhoun EA. Utilizing Exploratory Spatial Data Analysis to Examine Health and Environmental Disparities in Disadvantaged Neighborhoods. *Environmental justice (Print)*. 2013; 6(3):81–7. [PubMed: 26594302]
59. Ou Y, Mai J, Zhuang J, Liu X, Wu Y, Gao X, et al. Risk factors of different congenital heart defects in Guangdong, China. *Pediatric research*. 2016; 79(4):549–58. [PubMed: 26679154]
60. Padilla CM, Deguen S, Lalloue B, Blanchard O, Beaugard C, Troude F, et al. Cluster analysis of social and environment inequalities of infant mortality. A spatial study in small areas revealed by local disease mapping in France. *The Science of the total environment*. 2013:454–455. 433–41.
61. Pratt GC, Vadali ML, Kvale DL, Ellickson KM. Traffic, air pollution, minority and socioeconomic status: addressing inequities in exposure and risk. *International journal of environmental research and public health*. 2015; 12(5):5355–72. [PubMed: 25996888]
62. Ribeiro AI, de Pina Mde F, Mitchell R. Development of a measure of multiple physical environmental deprivation. After United Kingdom and New Zealand, Portugal. *European journal of public health*. 2015; 25(4):610–7. [PubMed: 25653297]
63. Richmond-Bryant J, Meng Q, Cohen J, Davis JA, Svendsgaard D, Brown JS, et al. Effect measure modification of blood lead-air lead slope factors. *Journal of exposure science & environmental epidemiology*. 2015; 25(4):411–6. [PubMed: 24961837]
64. Shmool JL, Kubzansky LD, Newman OD, Spengler J, Shepard P, Clougherty JE. Social stressors and air pollution across New York City communities: a spatial approach for assessing correlations among multiple exposures. *Environmental health : a global access science source*. 2014; 13:91. [PubMed: 25374310]
65. Skroder HM, Hamadani JD, Tofail F, Persson LA, Vahter ME, Kippler MJ. Selenium status in pregnancy influences children’s cognitive function at 1.5 years of age. *Clinical nutrition (Edinburgh, Scotland)*. 2015; 34(5):923–30.
66. Solimini AG, D’Addario M, Villari P. Ecological correlation between diabetes hospitalizations and fine particulate matter in Italian provinces. *BMC public health*. 2015; 15:708. [PubMed: 26208978]
67. Vaghri Z, Guhn M, Weinberg J, Grunau RE, Yu W, Hertzman C. Hair cortisol reflects socioeconomic factors and hair zinc in preschoolers. *Psychoneuroendocrinology*. 2013; 38(3):331–40. [PubMed: 22809790]
68. Vishnevetsky J, Tang D, Chang HW, Roen EL, Wang Y, Rauh V, et al. Combined effects of prenatal polycyclic aromatic hydrocarbons and material hardship on child IQ. *Neurotoxicology and teratology*. 2015; 49:74–80. [PubMed: 25912623]
69. Wilson S, Burwell-Naney K, Jiang C, Zhang H, Samantapudi A, Murray R, et al. Assessment of sociodemographic and geographic disparities in cancer risk from air toxics in South Carolina. *Environmental research*. 2015; 140:562–8. [PubMed: 26037107]
70. Winquist A, Steenland K. Modeled PFOA exposure and coronary artery disease, hypertension, and high cholesterol in community and worker cohorts. *Environmental health perspectives*. 2014; 122(12):1299–305. [PubMed: 25260175]

71. Breiman L. Random forests. *Mach Learn.* 2001; 45(1):5–32.
72. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comp Sci.* 2003; 43(6):1947–58.
73. Rodriguez-Galiano VF, Ghimire B, Rogan J, Chica-Olmo M, Rigol-Sanchez JP. An assessment of the effectiveness of a random forest classifier for land-cover classification. *Isprs J Photogramm.* 2012; 67:93–104.
74. Diaz-Uriarte R, de Andres SA. Gene selection and classification of microarray data using random forest. *Bmc Bioinformatics.* 2006; 7doi: 10.1186/1471-2105-7-3
75. Hansen LK, Salamon P. Neural Network Ensembles. *Ieee T Pattern Anal.* 1990; 12(10):993–1001.
76. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015; 521(7553):436–44. [PubMed: 26017442]
77. Cangelosi D, Pelassa S, Morini M, Conte M, Bosco MC, Eva A, et al. Artificial neural network classifier predicts neuroblastoma patients' outcome. *Bmc Bioinformatics.* 2016; 17(Suppl 12):347. [PubMed: 28185577]
78. Inthachot M, Boonjing V, Intakosum S. Artificial Neural Network and Genetic Algorithm Hybrid Intelligence for Predicting Thai Stock Price Index Trend. *Comput Intell Neurosci.* 2016; 2016:3045254. [PubMed: 27974883]
79. Rowley HA, Baluja S, Kanade T. Neural network-based face detection. *Ieee T Pattern Anal.* 1998; 20(1):23–38.
80. Zhang Q. Credit Risk Model Based on Artificial Neural Network for Financial Market. *J Invest Med.* 2014; 62(8):S110–S1.
81. Lugade V, Lin V, Farley A, Chou LS. An artificial neural network estimation of gait balance control in the elderly using clinical evaluations. *Plos One.* 2014; 9(5):e97595. [PubMed: 24836062]
82. Bivand, R., Pebesma, EJ., Gómez-Rubio, V. *Applied spatial data analysis with R.* Second. New York: Springer; 2013. p. xviii. 405
83. Gelman, A., Hill, J. *Data analysis using regression and multilevel/hierarchical models.* Cambridge; New York: Cambridge University Press; 2007. p. xxi. 625
84. Maas C, Hox J. Sufficient Sample Sizes for Multilevel Modeling. *Methodology.* 2005; 1(3):86–92.
85. Bell SM, Edwards SW. Identification and Prioritization of Relationships between Environmental Stressors and Adverse Human Health Impacts. *Environ Health Persp.* 2015; 123(11):1193–9.
86. Kapraun DF, Wambaugh JF, Ring CL, Tornero-Velez R, Setzer RW. A Method for Identifying Prevalent Chemical Combinations in the U.S. Population. *Environ Health Persp.* 2017; 125(8): 087017.
87. Gronlund CJ, Berrocal VJ, White-Newsome JL, Conlon KC, O'Neill MS. Vulnerability to extreme heat by socio-demographic characteristics and area green space among the elderly in Michigan, 1990-2007. *Environmental research.* 2015; 136:449–61. [PubMed: 25460667]
88. van der Lely S, Frey S, Garbazza C, Wirz-Justice A, Jenni OG, Steiner R, et al. Blue blocker glasses as a countermeasure for alerting effects of evening light-emitting diode screen exposure in male teenagers. *J Adolesc Health.* 2015; 56(1):113–9. [PubMed: 25287985]
89. Dumont E, Johnson AC, Keller VDJ, Williams RJ. Nano silver and nano zinc-oxide in surface waters - Exposure estimation for Europe at high spatial and temporal resolution. *Environ Pollut.* 2015; 196:341–9. [PubMed: 25463731]
90. Wang A, Padula A, Sirota M, Woodruff TJ. Environmental influences on reproductive health: the importance of chemical exposures. *Fertil Steril.* 2016; 106(4):905–29. [PubMed: 27513554]

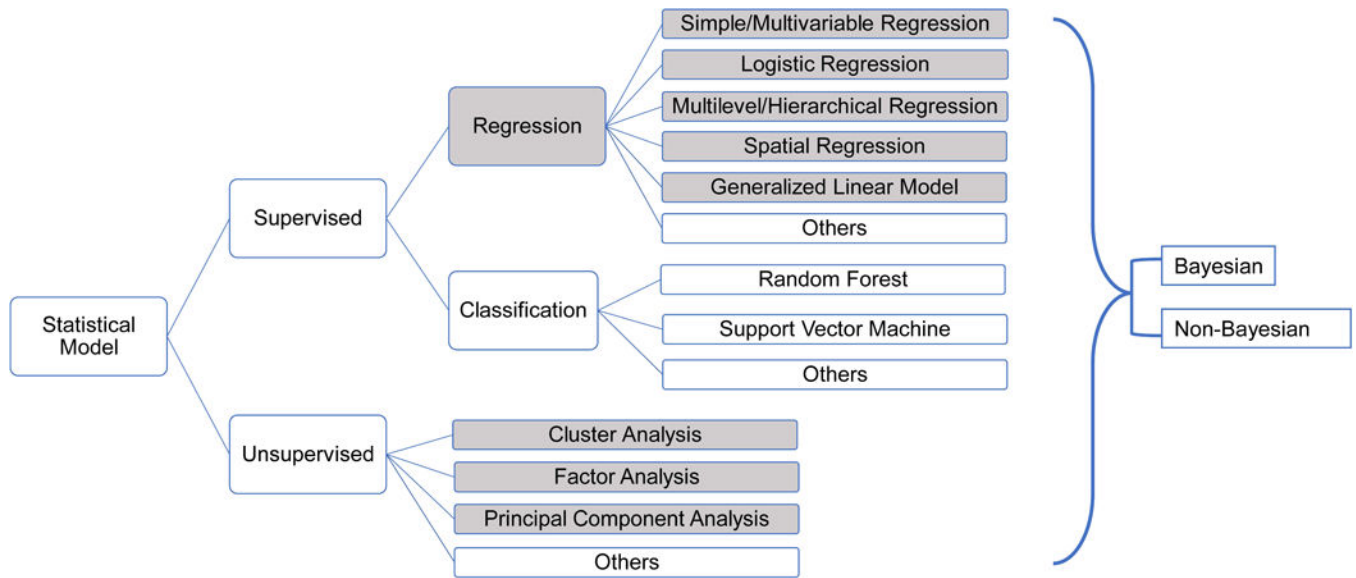


Figure 1. Different Types of Existing Statistical Models. Most of the current models used to capture the combined effects of multiple chemical/non-chemical stressors in the field of cumulative impact studies are regression models (highlighted in *gray shadows*).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1 Summary of Studies (n=31) Included from this Literature Review that Quantitatively Evaluate Combined Effects of both Chemical and Social Stressors on Health

Model Category	Model Subtype	Model	Number of Studies Used	Chemical Stressors Measured	Social Stressors Measured	Health Effects Studied	References
Supervised	Regression	Multivariable Linear/Non-linear Regression Model	10	Air pollutants, climate indicators, metals	SES, race, neighborhood features, psychological factors, material hardship	Mortality, children's BLL, blood pressure, cancer risk, IQ, respiratory disease	[37, 42, 43, 46, 52, 55, 61, 66, 68, 69]
		Logistic Regression Model	7	Air pollutants, silica, BPA	SES, education, access to health services, employment duration	Mortality, morbidity, nutritional status, IQ, cancer risk, child behavior, heart defects, lung function	[36, 38, 40, 45, 47, 59, 68]
		Multilevel/Hierarchical Model	5	Air pollutants, metals	SES, race, housing	BLL, pregnancy outcomes	[49, 56, 57, 63, 67]
		Generalized Linear Regression Model	3	Air pollutants, drinking water pollutants	SES, health prevention program	Mortality, morbidity	[54, 57, 62]
		Spatial Regression Model	2	Air pollutants	SES, race, physical disorder, violent crime	Respiratory disease, cancer risks	[58, 64]
Unsupervised	Cluster Analysis	Other Regression Models	5	Air pollutants	SES, physical activity, psychological factors, employment, urbanization, neighborhood features	Mortality, respiratory disease, autoimmune disease	[48, 50, 51, 53, 60]
		Factor Analysis	2	Air pollutants	Race, neighborhood deprivation, intervention program	Mortality, cancer risk	[50, 54, 60]
		Principal Component Analysis	2	Air pollutants	SES, race, physical disorder, crime, unemployment	Mortality, cancer risk, respiratory disease	[48, 64]
			2	Air pollutants	SES	Diabetes hospitalization	[61, 66]

Note: BLL – Blood Lead Level; BPA - Bisphenol A; IQ – Intelligence Quotient; SES – Socio-economic Status.

Table 2

Sample Studies Included that Quantitatively Evaluate both Chemical and Social Stressors on Health

Model Category	Model	Key Chemical Stressor	Key Social Stressor	Combined Effects	Location	Reference
Supervised/Regression/Multivariable Regression	Multivariable Regression	Organic solvent or pesticides	Maternal SES, occupation	Congenital heart defects	China	[59]
Supervised/Regression/Multivariable Regression, Unsupervised/PCA	Multivariable Regression/PCA	PM _{2.5}	SES	Diabetes hospitalization	Italy	[66]
Supervised/Regression/Logistic Regression	Conditional Logistic Regression	PM ₁₀ , NO ₂ , SO ₂ , CO, O ₃	SES	Mortality rate	Brazil	[40]
Supervised/Regression/Multilevel Modeling	Hierarchical Bayesian model	NO ₂ , NO _x	SES, exposure-related and land-use factors	Low term birth weight	U.S.	[56]
Supervised/Regression/GLM	Negative Binomial Regression	PM ₁₀ , NO ₂ , CO, THM in drinking water, nitrate in private well	Socioeconomic deprivation	Mortality	Europe	[62]
Supervised/Regression/Generalized Linear Model, Unsupervised/Cluster Analysis	Zero-inflated negative binomial (ZINB) regression, combined with cluster analysis	Arsenic in drinking water	Health intervention programs	Acute lower respiratory infection mortality	Bangladesh	[54]
Supervised/Regression/Spatial Model, Unsupervised/Factor Analysis	Simultaneous Autoregressive models	NO ₂	Violent crime and physical disorder	Childhood asthma exacerbation	U.S.	[64]
Supervised/Regression/Others	Cox proportional hazards model	PM _{2.5}	Physical activity, social engagement	Mortality rate	U.S.	[53]
Supervised/Regression/Others, Unsupervised/Cluster Analysis	Generalized estimating equations, combined with cluster analysis	Hazardous air pollutants	Race/ethnicity	Cancer risks	U.S.	[50]

Note: CO – Carbon Monoxide; NO₂ - Nitrogen dioxide; O₃ – Ozone; PM – Particulate Matter; SES – Socio-economic Status; SO₂ - Sulfur dioxide; THM – Trihalomethane