Reproductive Toxicology 45 (2014) 1–7

Contents lists available at ScienceDirect

Reproductive Toxicology

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

Investigation of maternal environmental exposures in association with self-reported preterm birth☆

Chirag J. Patel[a], Ting Yang[b], Zhongkai Hu[b], Qiaojun Wen[b], Joyce Sung[c], Yasser Y. El-Sayed[c], Harvey Cohen[d], Jeffrey Gould[e], David K. Stevenson[e], Gary M. Shaw[e], Xuefeng Bruce Ling[b,1], Atul J. Butte[f,g,*,1], on behalf of the March of Dimes Prematurity Research Center at Stanford University School of Medicine

[a] Center for Biomedical Informatics, Harvard Medical School, Boston, MA, USA
[b] Department of Surgery, Stanford University School of Medicine, Stanford, CA, USA
[c] Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford, CA, USA
[d] Division of Hematology-Oncology, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA
[e] Division of Neonatal and Developmental Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA
[f] Division of Systems Medicine, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA
[g] Lucile Packard Children's Hospital, Palo Alto, CA, USA

ARTICLE INFO

Article history:
Received 16 July 2013
Received in revised form
11 December 2013
Accepted 18 December 2013
Available online 27 December 2013

Keywords:
Environmental exposure
Environment-wide Association Study
Preterm birth

ABSTRACT

Identification of maternal environmental factors influencing preterm birth risks is important to understand the reasons for the increase in prematurity since 1990. Here, we utilized a health survey, the US National Health and Nutrition Examination Survey (NHANES) to search for personal environmental factors associated with preterm birth. 201 urine and blood markers of environmental factors, such as allergens, pollutants, and nutrients were assayed in mothers (range of *N*: 49–724) who answered questions about any children born preterm (delivery <37 weeks). We screened each of the 201 factors for association with any child born preterm adjusting by age, race/ethnicity, education, and household income. We attempted to verify the top finding, urinary bisphenol A, in an independent study of pregnant women attending Lucile Packard Children's Hospital. We conclude that the association between maternal urinary levels of bisphenol A and preterm birth should be evaluated in a larger epidemiological investigation.

1. Introduction

Preterm birth has complex etiology and identification of environmental factors that influence its risk is a priority. Factors postulated to influence risk for preterm birth include those associated with adverse lifestyle and behavior, such as stress, smoking, drug use, and nutrition (as summarized in references in [1,2]). However, lifestyle and behavior represent a complex mixture of environmental exposures, such as particulates in air pollution or specific nutrients in food [3]. Pregnant women are exposed to a multitude of environmental factors [4]. Simultaneous investigation of a multitude of exposures is challenged by a lack of comprehensive data as well as analytic approaches to query data in a systematic fashion.

Recently, an analytical approach dubbed an "Environment-wide Association Study" (EWAS) has been proposed to search for multiple environmental factors connected to disease-related phenotypes, including blood pressure, type 2 diabetes, cholesterol, and mortality [5–8]. The objective of this investigation is exploratory and to apply the methodology to derive hypotheses between maternal levels of exposures of numerous factors with self-reported preterm birth. Specifically, we analyzed participants of four independent United States Nutrition and Examination Surveys (NHANES) between 1999 and 2006 whose serum, urine, or tap water had been assayed for levels of 201 environmental factors [9]. These factors included phenols, phthalates, industrial pollutants, and nutrition. We associated each of the 201 factors with history of self-reported preterm birth (delivery before 37 weeks). To lessen

Abbreviations: CDC, Centers for Disease Control and Prevention; EWAS, Environment-wide Association Study; FDR, false discovery rate; LPCH, Lucile Packard Children's Hospital; NCHS, National Centers for Health Statistics; NHANES, National Health and Nutrition Examination Survey; SES, socioeconomic status.

* Corresponding author at: Department of Pediatrics, Stanford School of Medicine, 1265 Welch Road, X-163 MS-5415, Stanford, CA 94305, USA. Tel.: +1 650 723 3465; fax: +1 650 723 7070.
E-mail address: abutte@stanford.edu (A.J. Butte).
[1] Co-last authors.

0890-6238/$ – see front matter © 2014 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.reprotox.2013.12.005

chances of reverse causality, we chose individuals who reported having their last birth at least 1 year prior to the survey. In doing so, the search for factors correlated with history of preterm birth is exploratory as measurement of exposures is subsequent to the delivery event.

The second objective of the study included exploration of the correlations between a top data-driven finding, bisphenol A, in an independent cohort study of 37 consenting pregnant women attending Lucile Packard Children's Hospital at Stanford University Medical Center. In this study we observed nominally higher levels of urinary bisphenol A in mothers who went on to have a preterm birth.

2. Material and methods

2.1. Data: NHANES 1999–2000, 2001–2002, 2003–2004, 2005–2006

We used the NHANES to conduct a systematic scan of directly assayed maternal environmental factors associated with self-reported preterm birth. We downloaded all available NHANES laboratory and questionnaire data for 1999–2000, 2001–2002, 2003–2004, and 2005–2006 surveys. Each survey is an independent and non-overlapping sampling of participants representative of the United States population administered by the Centers for Disease Control and Prevention (CDC) and the National Centers for Health Statistics (NCHS) [10–13].

NCHS asked eligible participants how many times they have been pregnant ("How many times have you been pregnant"), the age of their last pregnancy ("How old were you at the time of your last birth?") how many births resulted in "low birth weight" infants ("How many of your children weighed less than 5.5 pounds at birth?"). For individuals that answered "yes" to having a low birth weight infant, they were subsequently asked how many infants were born preterm ("How many children born preterm? A preterm delivery is one that occurs at 36 weeks or earlier in pregnancy."). Therefore, participants with a history of preterm birth were restricted to those who had low birth weight children. No information was provided about the mode of preterm delivery, such as iatrogenic or labor induced. Further, no information was provided regarding whether preterm births were singleton or non-singleton. There were 1446, 1550, 1415, and 1361 participant mothers who had at least 1 live birth in the 1999–2000, 2001–2002, 2003–2004, and 2005–2006 surveys respectively, a total of 5772 mothers in all surveys. We then restricted this sample of 5772 participants to those who reported at least one pregnancy just one year prior to the time of survey to lessen the impact of exposures being measured too distant pregnancies from the gestational period of interest. This restriction yielded a total sample of 780 participants. Those who responded yes to having any preterm children were classified as participants with history of preterm birth ($N = 62$) possible and those who responded to not having any preterm births were classified as a no history of preterm birth ($N = 718$).

Laboratory data included serum, urine, or water measures of environmental factors (Fig. 1B). We analyzed factors that were a direct measurement of environmental factors (e.g., amount of pesticide or heavy metal in urine or blood or amount of chemical compound in tap water sources of participants). There were 304 of these factors that were linkable to eligible participants with different sample sizes ranging from 2 to 62 participants with history of preterm and 44 to 700 without history of preterm birth. We eliminated from consideration 52 of these variables that had fewer than 10 participants with history of preterm. We further removed from consideration 51 variables because 99% of the observations were under the NCHS documented limits of detection. We also verified whether any variables that had NCHS documented limits of detection had a majority (>99%) of detected values belonging to participants who had a history of preterm birth. Supplementary Table 1 shows the number of participants with detected and non-detected factor values stratified by history of preterm. We found that none of the environmental variables were exclusively detected in participants with history of preterm birth (Table S1). No substitutions were made for variable values that were reported as less than the limit of detection. This left 201 variables (Fig. 1B) in diverse categories such as infectious agents (13 bacteria and 11 viruses), 23 polychlorinated biphenyls, 6 dioxins, 7 di-alkyl pesticide metabolites, 22 pesticides, 32 nutrients, 21 heavy metals, 4 furans, 9 hydrocarbons, 3 phenols, 11 phthalates, 6 phytoestrogens, 9 polyfluorochemicals, and 20 volatile organic compounds measured in participants' tap water or serum. Of these, 40 serum-measured variables representing lipophilic compounds, including furans, polychlorinated biphenyls, and organochlorinated pesticides were reported on both a whole weight in serum ("unadjusted") as well as relative to total serum lipids ("lipid-adjusted") basis. For these variables in our scan, we analyzed the whole-weight variables and compared the estimates and p-values of the whole weight variables to the lipid adjusted variables. In summary, the 201 environmental factors were measured in varying numbers of participants, ranging from $N = 106$ to $N = 762$. Individuals are selected randomly based on their demographic characteristics for the complex, stratified survey [14].

2.2. Systematic scan of environmental biomarkers of exposure associated with self-reported preterm birth

Our analysis consisted of performing 201 survey-weighted logistic regressions, where history of preterm birth was the dependent variable and modeled as a function of each environmental factor and age, race-ethnicity, education and socioeconomic status (SES), and number of births. For SES we used the tertile of poverty index (participant's household income divided by the time-adjusted poverty threshold), as previously described [5,6]. Race-Ethnicity was grouped as: Mexican American, Non-Hispanic Black, Non-Hispanic White, Other Hispanic, or Other. We chose these factors for adjustment as they are known to be associated with preterm birth and likely also associated with exposure [1]. We used *R survey* module for all survey-weighted analyses [15] with appropriate pseudo-strata, pseudo-sampling units, and weights to accommodate the complex sampling of the data. We chose weights corresponding to the smallest sub-sample for each environmental factor tested. Because of the comparison between individuals with and without history of preterm birth, exposure measurements follow the delivery events.

We transformed continuous measurements to "z-scores" (number of standard deviations from the mean) to compare effect sizes; specifically, effect sizes for these variables denote change in odds for preterm birth for a change in 1 standard deviation of exposure. For binary variables, such as presence/absence assays for infectious agents, effect sizes denote change in odds for preterm birth for those with a factor versus those without.

To account for multiple hypotheses, we calculated the false discovery rate (FDR), the estimated proportion of false discoveries made versus the number of total discoveries made for a given significance level [16]. Specifically, we used the Benjamini–Hochberg step-down procedure to compute the FDR [16]. We ranked findings from lowest to highest FDR (which corresponds to the lowest to highest p-values). We considered factors that achieved an FDR less than 40–50% to be the least susceptible to be a spurious finding and worth examining further in an independent cohort of patients attending Stanford Hospital and Clinics.