



Published in final edited form as:

Environ Res. 2023 January 15; 217: 114793. doi:10.1016/j.envres.2022.114793.

A Cross-Validation Based Approach for Estimating Specific Gravity in Elementary-School Aged Children Using a Nonlinear Model

Stefanie A. Busgang, MPH^a, Syam S. Andra, PhD^a, Paul Curtin, PhD^a, Elena Colicino, PhD^a, Matthew J. Mazzella, BA^a, Moira Bixby, MPH^a, Alison P. Sanders, PhD^b, John D. Meeker, ScD, CIH^c, Marissa Hauptman, MDMPH^{d,e,f}, Shirisha Yelamanchili, MS^a, Wanda Phipatanakul, MD, MS^{d,g}, Chris Gennings, PhD^a

^a. Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA

^b. Department of Environmental and Occupational Health, University of Pittsburgh, Pittsburgh, PA, USA.

^c. Department of Environmental Health Sciences, University of Michigan School of Public Health, Ann Arbor, MI, USA

^d. Harvard Medical School, Boston, MA, USA

^e. Division of General Pediatrics, Boston Children's Hospital, Boston, MA, USA

^f. Region 1: New England Pediatric Environmental Health Specialty Unit, Boston, MA, USA

^g. Division of Allergy and Immunology, Boston Children's Hospital, Boston, MA, USA

Abstract

Environmental research often relies on urinary biomarkers which require dilution correction to accurately measure exposures. Specific gravity (SG) and creatinine (UCr) are commonly measured

Corresponding Author – Stefanie A Busgang, Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, Box 1057, New York, NY, 10029, Stefanie.busgang@mssm.edu, 212-825-7115.

CRedit Statement

Stefanie A. Busgang: Conceptualization, Methodology, Formal analysis, Writing – Original draft, Visualization; **Syam S. Andra:** Investigation, Resources, Writing – Review & editing; **Paul Curtin:** Methodology, Supervision, Writing-Review & editing; **Elena Colicino:** Methodology, Writing – review & editing; **Matthew J. Mazzella:** Methodology, Writing – Review & editing; **Moira Bixby:** Methodology, Writing – Review & editing; **Alison P Sanders:** Conceptualization, Writing – Review & editing; **John D. Meeker:** Methodology, Writing – Review & editing; **Marissa Hauptman:** Investigation, Data Curation, Writing-Review & editing; **Shirisha Yelamanchili:** Investigation, Resources; **Wanda Phipatanakul:** Investigation, Data Curation, Funding Acquisition, Writing – Review & editing; **Chris Gennings:** Conceptualization, Methodology, Supervision, Acquisition, Writing – Review & editing.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Conflicts of Interest

The authors declare that they have no competing interests

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

urinary dilution factors. Epidemiologic studies may assess only one of these measures, making it difficult to pool studies that may otherwise be able to be combined.

Participants from the National Health and Nutrition Examination Survey 2007–2008 cycle were used to perform k-fold validation of a nonlinear model estimating SG from UCr. The final estimated model was applied to participants from the School Inner-City Asthma Intervention Study, who submitted urinary samples to the Children’s Health Exposure Analysis Resource. Model performance was evaluated using calibration metrics to determine how closely the average estimated SG was to the measured SG. Additional models, with interaction terms for age, sex, body mass index, race/ethnicity, relative time of day when sample was collected, log transformed 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), and asthma status were estimated and assessed for improvement. The association between monobenzyl phthalate (MBZP) and asthma symptom days, controlling for measured UCr, measured SG, and each estimated SG were compared to assess validity of the estimated SG.

The model estimating SG from UCr alone, resulted in a beta estimate of 1.10 (95% CI: 1.01, 1.19), indicating agreement between model-predicted SG and measured SG. Inclusion of age and sex in the model improved estimation ($\beta=1.06$, 95% CI: 0.98, 1.15). The full model accounting for all interaction terms with UCr resulted in the best agreement ($\beta= 1.02$, 95% CI: 0.94,1.10). Associations between MBZP and asthma symptoms days, controlling for each estimated SG, were within the range of effect estimates when controlling for measured SG and measured UCr (Rate ratios= 1.28–1.34). Our nonlinear modeling provides opportunities to estimate SG in studies that measure UCr or vice versa, enabling data pooling despite differences in urine dilution factors.

Keywords

dilution factors; data pooling; calibration metrics

Introduction

Urine is a useful biological matrix for collection and analysis given potential ease and quantity of collection, potential increased sensitivity given larger available volume, and the presence of metabolites that may have a longer biological half-life than parent compounds. When analyzing urinary biomarkers of environmental exposures it is crucial to adjust for dilution factors such as specific gravity (SG), urinary creatinine (UCr), and/or osmolality due to intra- and interpersonal variation of urinary concentrations and volume of urine per void [1, 2]. An accurate representation of exposure level may differ from what is measured due to a wide range of physiological changes in urinary flow rate which are dependent on the water and salt intake of the participant [3]. Without a correction, a participant’s biomarker measurement may appear to be an outlier; however, this urine sample may be highly concentrated, and a dilution correction may reveal that the biomarker concentration is within an acceptable range. Additionally, a dilution factor can account for some exposure misclassification that would result in biased effect estimates if concentrations were not corrected [4, 5].

There are several ways to correct for differential dilution across samples. Standard approaches include dividing the biomarker concentration by the dilution factor, controlling for the factor as a covariate in a regression model, or applying specific formulas such as the one proposed by Meeker et al (2005) [3, 6–8]. There is disagreement on which of these correction methods is most robust and the choice of the correction method may depend on the toxicant(s) or outcome of interest and dilution factor measured [5, 9, 10].

The Human Health Exposure Analysis Resource (HHEAR; a continuation of the Children's Health Exposure Analysis Resource) is a consortium intended to allow researchers to combine data from various epidemiological studies with environmental exposure analysis provided by a HHEAR Lab Hub [11]. One advantage of a consortium such as HHEAR is the ability to pool data across studies, resulting in an increased sample size, and the potential for increased power to detect subtle associations especially in subgroup analyses and rare disease groups, and improved generalizability [12]. However, studies may not be comparable if different dilution factors are measured. Because epidemiological study designs are diverse, each may have previously measured SG or UCr prior to HHEAR approval. Thus, there is a need to estimate a cross-study dilution factor to better integrate urine exposure data across studies.

Here, we used a HHEAR study, with HHEAR lab measures of both SG and UCr, to estimate SG based on measured UCr in school-age children using a nonlinear model. Estimated SG was compared to measured SG by plotting the linear relationship and assessing the beta coefficient to determine the model's accuracy in prediction. Further, we trained the nonlinear model using data from the National Health and Nutrition Examination Survey (NHANES), a U.S. based cohort of similarly aged children. Like our HHEAR study, NHANES cycle 2007–2008 measured both SG and UCr [13]. This allowed us to compare the estimated SG, derived from UCr, to the true measured SG in HHEAR. By using this nonlinear model, we can improve comparability of urinary biomarker data across studies and minimize error.

Methods

Participants.

The HHEAR approved project, with SG and UCr measures, included participants from the School Inner-City Asthma Intervention Study (SICAS2) [14, 15]. Details on recruitment and study design can be found elsewhere. In short, for the subset with samples funded by HHEAR, recruited during the first 3 years of the parent study, 168 children with asthma between the ages of 5–15 years were recruited if he/she attended one of 25 urban elementary schools in the Northeastern United States. Recruitment for this subset of children from the whole study included students with recurring enrollment from 2015–2017. Urine samples were collected during the baseline study visit, occurring from April through October, with 95% of samples collected between June–September prior to intervention deployment. For the purpose of our study, this population will be referred to as the validation cohort. Data for this project was obtained from the publicly available data in the Human Health Exposure Resource (HHEAR) Data Repository, which has been approved under Icahn School of Medicine at Mount Sinai IRB Protocol # 16–00947.

Children of the same age were selected from a nationally representative population from NHANES cycle 2007–2008 (N=2155) [13]. We did not require that children have asthma to keep the prediction model generalizable and to test if an asthma indicator could further improve model estimation. These NHANES participants will be referred to as the training cohort. The National center for Health Statistics Research Ethics Review Board approved documented consent for all NHANES participants.

SG and UCr.

The Lautenberg Laboratory at Icahn School of Medicine at Mount Sinai, a HHEAR lab hub, measured urinary UCr of participants from samples collected at baseline in SICAS2 using a well-established colorimetric method (detection limit was 0.3125 mg dL⁻¹) [16]. Urine SG was measured with a small volume method (10 μ L) (Rudolph Research Analytical, Hackettstown, NJ, USA), with a detection limit 1.0000 kg m⁻³ [17]. Quality controls (QC) included in each batch were experimental blanks, lab urine pools, NIST SRMs and proficiency testing material where available. Batch-wise coefficient of variation of QCs during analysis of the study specimens were <20% of target concentration.

Urinary specimens from NHANES participants are stored and shipped to University of Minnesota in Minneapolis, MN for analyses of UCr and to Division of Laboratory Sciences, National Center for Environmental Health, and Centers for Disease Control and Prevention for the analyses of SG. UCr was measured using a Beckman Synchron CX3 Clinical Analyzer in 2007 and a Roche/Hitachi Modular P Chemistry Analyzer in 2008. Additional information on laboratory methods is available online (CDC, NHANES 2007–2008; ; <https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/labmethods.aspx?BeginYear=2007>). Urinary SG was measured using a digital hand-held refractometer ATAGO PAL-10S from ATAGO (Bellevue, WA, USA), with automatic temperature compensation. One blank and one QC were measured per batch for SG analyses. Reported results met the accuracy and precision requirements of the quality control program of the Division of Laboratory Sciences, National Center for Environmental Health, CDC.

Statistical Analysis.

SG was plotted against UCr using data from the training cohort. We modeled the relationship between UCr and SG using a Gompertz model to define the asymmetric sigmoidal shape visually determined by Figure 1 [18]. Since SG cannot be measured below 1, a constraint was added to the model, such that estimated values always exceed 1. Using the estimated equation structure, k-fold cross validation was performed by splitting the NHANES dataset in 5 approximately equal partitions [19]. The parameters of the model were estimated using partitions 1–4 and the resulting model was tested on partition 5. This was repeated 5 times, setting one partition of the data aside for testing and estimating the parameters using the rest of the data. The final model parameters were the average across the 5 validations. A second and third model were estimated to include several additional terms that have been previously shown to be related to UCr and/or SG [20]. These include age, sex, body mass index (BMI), race/ethnicity, and relative time of day when sample was collected (i.e. morning, afternoon, or evening) as well as their interaction with UCr; i.e.,

$$SG = 1 + \alpha \exp\left(-\exp\left(-\left[\beta(UCr - \delta) + \sum_j \{\theta_j z_j + \gamma_{cj} z_j (UCr - \delta)\}\right]\right)\right), \quad (1)$$

where the z 's are covariates for age (centered), BMI (centered), sex, race/ethnicity, and relative time of day when sample was collected; the θ_{js} are linear parameters and the γ_{cjs} are interaction parameters for the j th covariate with UCr, where j indexes the covariates in sequence up to including all five covariates. The second model included age and sex only and the third model include all five covariates. Lastly, a fourth model was estimated to include two additional terms that overlap in both studies and have some evidence of a relationship with kidney function but have not been directly linked to UCr or SG [21–23]. This full model includes all previously described terms in addition to natural log transformed urinary 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) concentration (metabolite of tobacco smoke), and current asthma status. For each study, values of NNAL below the limit of detection (LOD) or the limit of quantification (LOQ) were imputed as the LOD/LOQ/ 2 prior to log transformation. For NHANES, the LOD was 0.000566 ng/ml and for the CHEAR study, the LOQ was 0.0105 ng/ml. Since continuous variables are centered using a nationally representative sample, if this equation is applied to another study and is missing one variable, the term can be removed if the national average is a reasonable assumption.

To confirm that these variables have a relationship with UCr and/or SG, as previously described in the literature, we ran individual associations between each variable and each dilution factor, per study, using Pearson correlations for continuous variables (age, BMI, and log transformed NNAL), t-tests for binary variables (sex and asthma), and Analysis of Variance (ANOVA) for variables with three or more categories (race and session time). The associations of UCr and SG with asthma status were not tested in the validation cohort since all children had asthma.

To assess the performance of our estimation models, we report calibration metrics [24] and visualize a calibration plot to determine the linear relationship between estimated/predicted SG and measured/observed SG for both the full training sample and the HHEAR validation cohorts separately. A linear regression between the two terms produced an adjusted R^2 , beta estimate, 95% confidence intervals, and p-value for each model tested. Near perfect agreement between measured SG and estimated SG would be represented by a beta coefficient equal to 1. The adjusted R^2 is used to guide our final model selection to ensure we are not overfitting the model.

To further assess validity of the estimated SG, we compared the association between a phthalate metabolite and asthma severity, since previous studies have found a positive association between phthalates and asthma. Specifically, zero-inflated Poisson regression models were constructed, using asthmatic children from the HHEAR cohort, to analyze the effect of log transformed monobenzyl phthalate (MBZP) on the number of asthma symptoms days assessing both the count model and logit model with the metabolite without a dilution correction, the metabolite controlling for measured creatinine, controlling for measured specific gravity, and the metabolite controlling for each estimated specific gravity. As a comparison to methods previously proposed, we ran sensitivity analyses to assess

associations adjusting for covariates used in SG estimation (age, sex, BMI, relative time of sample collection, and log transformed NNAL) instead of adjusting for estimated SG and to assess associations using the z-score of UCr for a randomly selected half of the validation cohort and using z-score of SG for the other half of the cohort.

There is evidence that UCr is sensitive to the number of freeze-thaw cycles [25]. To address this concern, a sensitivity analysis was examined using our first estimation model with only UCr as a predictor of SG, without considering other variables, stratified by time since sample collection (between 0 and 3 years) in the validation cohort. All analyses were conducted using SAS 9.4 (Cary, NC).

Results

One hundred sixty-eight children ages 5–15 were recruited during the first three years of the SICAS-2 parent study. Of those, 141 children (ages 6 – 14) provided urine specimens that were able to be included in this analysis. Of the 2155 children, ages 6–14, who participated in NHANES, 1530 had measures of SG, UCr, and BMI and were included in the analysis. Table 1 provides descriptive statistics of UCr and SG of each cohort. Ranges of UCr and SG were included to show the high variability in measurements and potential outliers. Since age, sex, BMI, race/ethnicity, and time of day of sample collection have been associated with UCr and SG measures in previous literature [20], we examined the individual associations of these selected demographic variables with UCr and SG in our training and validation cohorts separately (Table 2). Additionally, associations with NNAL and asthma were tested since these were potential variables of interest for our estimation model. All variables except asthma status were significantly associated with UCr and all variables except race/ethnicity and asthma status were significantly associated with SG in the training cohort. In the validation cohort, only age, race/ethnicity, and NNAL were significantly associated with both UCr and SG.

The graphical representation of the relationship between UCr and SG in the NHANES training dataset is shown in Figure 1. The Pearson correlation estimate between SG and natural log transformed UCr was 0.82 (Figure 2). Table 3 provides parameter estimates and their standard errors (SE) for each nonlinear model, estimated using the NHANES training cohort. The first estimated nonlinear model, referred to as Model 1, included a term for UCr only. We observed a strong linear association between the estimated SG and measured SG (Figure 3A and Table 4; $\beta=1.00$, CI: 0.97, 1.04) within the training sample with a moderate adjusted R^2 of 0.691. However, when our equation was applied in the validation sample, SG was overestimated compared to measured SG (Figure 3B and Table 4; $\beta=1.10$, CI: 1.01, 1.19). The confidence interval does not include 1 for this model so we explored other factors to improve this estimation. The adjusted R^2 was 0.812 which will be compared to subsequent models.

A secondary nonlinear model, referred to as Model 2, was created to estimate SG, accounting for potential interactions between UCr and age and sex. When we applied this equation in the validation cohort, SG estimation improved as indicated by a linear beta coefficient closer to 1 (Figure 3C and Table 4; $\beta=1.06$, CI: 0.98, 1.15). There was also some

improvement in model fit as indicated by an adjusted R^2 of 0.826. Next, we added additional terms that are often, but not always, collected across studies. In addition to the covariates in Model 2, Model 3 includes terms for BMI and relative time of sample collection (i.e. morning, afternoon, evening). Lastly, we explored a fourth nonlinear model, referred to as Model 4, using the same demographics included in Model 3, with additional terms for current asthma status and log transformed NNAL concentrations. In the training dataset, asthma was defined by children who self-reported currently having asthma in comparison to children who indicated never having asthma or no longer having it. When estimating SG in the validation sample using this equation, asthma was coded as 1 (does have asthma) for all participants. Estimation of SG with this model was the best (Figure 3D and Table 4; $\beta=1.02$, CI: 0.94, 1.10); however, the adjusted R^2 decreased slightly from the previous model (0.825).

Next, the association of MBZP, controlling for no dilution or different renditions of dilutions, on asthma symptoms days was compared. For the model without a dilution, there was a significant increase in asthma symptom days (RR= 1.26, CI: 1.05,1.52). The association controlling for log transformed UCr was higher (RR=1.34, CI: 1.09,1.64). Finally, associations with measured SG or estimated SG from models 1–3 fell somewhere in between without the dilution and the model with log transformed UCr (RR=1.28, CI: 1.06,1.55; RR=1.34, CI: 1.06, 1.55; RR=1.33, CI: 1.08, 1.63; RR=1.30, CI=1.06, 1.60 respectively).

For the sensitivity analysis determining if freeze thaw cycles could affect our modeling, we found that estimation was similar across each time periods and thus time since sample collection is not likely to be problematic (Figure S1; <1 year, 1–2 years, and 2–3 years).

Discussion

In this study, using training and validation data from two large cohorts of US children, we demonstrate a method to estimate SG from measured UCr using a k-fold validated nonlinear model. The observed nonlinear relationship between UCr and SG is consistent with trends identified in prior literature, such that there is a linear relationship between UCr and SG at lower concentrations, whereas the relationship with SG remains constant at higher values of UCr.[26, 27]. The correlation between log transformed UCr and SG ($\rho=0.82$; Figure 2) is similar to the correlation that others have found [28–30]. Figure 2 also provides evidence for the use of a nonlinear model, rather than simply using the log transformation of UCr, since the smoothed curve asymptotes at 1 and appears to be sigmoidal in shape in comparison to the linear curve.

The HHEAR validation cohort revealed that our estimated Model 1 overestimates SG for participants when using UCr alone but gets closer to the true measured SG when the model includes interaction terms for age, sex, BMI, race/ethnicity, relative time of day of sample collection, log-transformed NNAL, and asthma status. Similarly, the association of MBZP with asthma symptoms days was highest, in comparison to other SG models, when using UCr alone to estimate SG. The models with additional interactions terms result in an association closer to that of the association controlling for measured SG. Though not

a perfect estimator, our nonlinear method enables researchers to pool urinary data across datasets that have measured either creatinine or specific gravity, allowing for increased power to test for subgroup specific associations.

By using NHANES to train models and by centering continuous variables, our model provides the flexibility to assume population-level averages when one of these variables is not available or the term can be omitted if it is a reasonable assumption; for example, in a cohort that is unrelated to asthma, in which the assumption that the population does not have asthma is reasonable, then values for this term would be 0, thus resulting in the omission of the beta for asthma. Likewise, if a population level average is appropriate for a continuous variable, the value would also be 0 since continuous variables are centered, again resulting in the omission of the beta related to that variable. Another option is to re-fit a model with appropriate overlapping variables or default to Model 2, which includes variables that are more likely to be measured in any study.

We note potential limitations of our model based on the way we chose to operationalize certain variables and the assumptions made. For example, race and ethnicity can be challenging to harmonize and is defined differently among studies. Specifically for datasets used herein, NHANES has established a 5-level categorization on self-reported race/ethnicity whereas HHEAR used a 6-level categorization of race and separately asked about ethnicity. Both cohorts required collapsing of categories to establish a commonality between studies, resulting in a 4-level race/ethnicity categorization. The reduction in granularity may limit the preciseness of our equation. For example, the NHANES cycle used in the current study did not report Asians as a racial group, thus reducing the ability to discern potentially important differences in creatinine clearance for this group. Using children ages 6–14 from NHANES cycles 2013–2018, which did discern Non-Hispanic Asian as a racial identity, we find that Asians have a lower median creatinine level compared to other races, and thus this distinction may improve our modeling in the future (Figure S2). Similarly, time of sample collection is provided differently between cohorts. NHANES defines session times as ‘morning,’ ‘afternoon,’ and ‘evening’ but does not indicate what exact time frame is represented by each session. Since HHEAR provides exact time of sample collection, we categorized time frames based on the assumption that morning occurs before noon, afternoon is from noon to 4 pm, and evening is any time after 4 pm. Van der Steen (2008) notes that operationalization of variable definitions and variable options often requires careful consideration when pooling data from different study populations; thus, this will always require thoughtful consideration when combining studies. For this reason, we recommend using the age and sex adjusted nonlinear model in most scenarios since these covariates should require little need for operationalizing or harmonizing. Additional models can be trained to improve upon these assumptions currently being made as well as include other terms that may further improve estimation.

Ogburn et al (2021) performed a simulation study to assess potential biases when using covariates estimated by auxiliary models, reporting on cases where the auxiliary variable was either the outcome, exposure, or covariate. The authors noted potential biases and concluded that variables used in the auxiliary model should not be adjusted for when applying the auxiliary covariate to the regression as this can further bias the association.

As such, we followed this recommendation and did not adjust for the variables used to estimate SG when we assessed the association between MBZP and asthma symptom days. Ogburn et al. also noted that it may be better to adjust for the variables used in the auxiliary model, rather than estimate the new covariate. As a sensitivity analysis we assessed the same association controlling for age, sex, BMI, and log transformed NNAL and found an estimate outside of the range of estimates found with true dilution measures, whereas associations controlling for estimated SG were within the range (Table 5). We recognize that there is potential for biased estimates using our method and future studies should assess similar models using additional datasets that also have both measured UCr and SG to further evaluate scenarios where there is bias. Despite this concern, our method has value in being able to combine studies in which a study may or may not have all the necessary auxiliary variables.

We observed associations between measured SG or UCr that varied by demographics among children including age, sex, BMI, and race/ethnicity. Previously Barr et al. showed that UCr increases until ages 12–19 and then decreases in older age groups [1]. Our results are consistent with this finding such that children ages 9–14 had higher average UCr levels compared to children ages 6–8 (Table 1). In addition, boys tend to have higher UCr levels than girls [28, 31]. In our study, data from the NHANES subset were consistent with this finding; however, we observed a higher mean level of UCr among girls compared to boys in the HHEAR cohort. SG remained consistent between sexes and between age groups, which supports the notion that SG is less impacted by these factors. Non-Hispanic blacks tend to have higher UCr concentrations compared to Non-Hispanic whites and Mexican Americans counterparts, across all age groups [1]. We observed this in this study where creatinine concentrations were at least 20 mg/dL higher among non-Hispanic Black participants than non-Hispanic White and Hispanic participants from both HHEAR and NHANES having UCr concentrations at least 20 mg/dL higher than non-Hispanic whites and Hispanics on average (Table 1). Additionally, our comparison of median creatinine among racial groups from NHANES cycles 2013–2018 further confirms higher levels among Non-Hispanic blacks (Figure S2).

We examined interaction terms because UCr excretion is influenced by age, sex, and muscle mass [31] and thus may be influential on the relationship between UCr and SG. Sauvé's findings support our inclusion of an interaction between UCr and sex due to smaller differences in SG levels between males and females compared to the differences in UCr [29]. However, Barr et al. also showed that sex differences in UCr are not apparent until after age 19 [1].

Our model that included a term for asthma and log transformed NNAL and their interaction with UCr (Model 3) resulted in the best estimation of SG in our validation cohort. This provides evidence that the relationship between UCr and SG may be affected by our disease of interest. Kartha et al. found co-occurrence of asthma and nephrolithiasis in a pediatric population, which suggests that asthma could impact the kidneys, resulting in a change in UCr clearance [23]. It is possible that other diseases and disorders can contribute to individual differences in dilution factors. For example, the use of UCr in assessing exposures related to kidney disease can be problematic since UCr is both filtered and secreted by

the kidneys [32]. In addition, some evidence suggests that exposure to second-hand smoke could have implications on kidney function, and thus may have an effect on UCr production [21, 22]. Both studies measured NNAL using liquid chromatography and mass spectrometry (LC/MS/MS). Though many studies may not have measures on NNAL, other indicators of tobacco exposure can be used to refit a model to include potential changes in UCr clearance due to this exposure.

Some debate exists in the literature where SG and UCr can be used interchangeably. Some evidence in adult studies suggests this may be appropriate when the measures are highly correlated [7, 33]; whereas when low correlation exists between urinary UCr and SG, these measures may not be interchangeable [4]. Our data support that in studies of children, a nonlinear model can be applied to appropriately estimate SG from UCr, or vice versa. Meanwhile, Mendiola et al. combined two studies of adult populations that either measured UCr or SG by taking the rank order of the dilution adjusted concentration [33]. SG was measured using a handheld refractometer, similar to the analytic method used in the NHANES cohort; however, the analytic method for UCR was not reported in this paper. More recently, Kuiper et al (2022) applied data collected in a cohort study of pregnant women to compare associations using the z-score of UCr for two sites, and the z-score of SG for two other sites in comparison to associations using all UCr across sites or using all SG across sites. They reported similar associations but note somewhat larger differences with these assignments. For comparison with our method, we included a comparison by randomly selecting half of our validation cohort using z-scored log transformed UCr and the other half using z-scored SG. The estimate was similar but somewhat higher and outside of the range of associations using measured log transformed UCr or SG for the entire population. We suggest our approach for child populations herein as an additional method to establish a commonality between varying dilution measures that is further refined by the inclusion of factors that may vary in one dilution measure and not the other. Future studies can assess nonlinear modeling in adult populations and pregnant women with available data. Likewise, future studies should assess models with other phenotypes besides asthma, which we were limited due to the nature of the validation dataset and lack of other phenotypes relevant to phthalate exposures in the training dataset among the study years selected.

How we chose to apply SG or UCr analytically in an epidemiological study testing for association (as a covariate or normalizing directly) varies based on the exposure of interest. For example, metals vs organics have different toxicological properties. Suwazano et al. (2005) suggested using a formula to correct for SG when measuring cadmium because it is less affected by age and sex [34]. Whereas organic compounds, such as phthalates, are actively excreted by the renal tubules and thus correcting for SG is preferred over UCr [31]. Meanwhile, Xia et al. (2014) demonstrated a better model fit when UCr was included as a separate covariate in models rather than SG when studying NNAL [35]. Muscat et al. (2011) found that the effect of UCr on urinary cotinine levels differed significantly by race and suggested using SG in a multiracial population, indicating the need to model these more complex relationships that are dependent on the exposure and population demographics [7]. Regardless of the reason for measuring UCr or SG, there continues to be a lack of consensus on which should be the standard measure. Thus, dilution estimation models provide the ability to make studies comparable, regardless of the correction factor measured.

Conclusion

We examined nonlinear models that were estimated using a nationally representative sample which can be applied to external cohorts to estimate SG from UCr. This can be a useful tool when combining across studies with different dilution measures so that biomarker concentrations can be more easily harmonized. We present four models to provide flexibility for the end user to choose which model is most appropriate based on the available data but show that estimation performs best when additional factors that potentially influence dilution are accounted for. We recommend using Model 2 in most scenarios as it performed as well as Model 3 and does not require additional operationalization of covariates that may or may not be present in studies being combined.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank the Centers for Disease Control and Prevention (CDC) for conducting NHANES, as well as the participants of the 2015–2016 NHANES cycle for making this research possible. We would like to thank the participants of the School Inner-City Asthma Intervention Study. Finally, we would like to express gratitude to the National Institute of Environmental Health Sciences (NIEHS) for providing support for HHEAR. This work was supported in part by funding from NIH/NIEHS: U2CES026561, U2CES026553, U2CES026555, R00ES027508, R01AI073964, R01AI073964-02S1, K24AI106822, K23ES031663, U01AI110397, and P30ES000002. Dr. Hauptman was also supported by the American Academy of Pediatrics (AAP) and funded in part by cooperative agreement award with the Centers for Disease Control and Prevention/Agency for Toxic Substances and Disease Registry (CDC/ATSDR) FAIN:NU61TS000296. The U.S. Environmental Protection Agency (U.S.EPA) supports the Pediatric Environmental Health Specialty Units (PEHSUs) by providing partial funding to the ATSDR under Inter-Agency Agreement DW-75-95877701. The findings and conclusions presented have not been formally disseminated by CDC/ATSDR or EPA and should not be construed to represent any agency determination or policy. Use of trade names that may be mentioned is for identification only and does not imply endorsement by the CDC/ATSDR or EPA.

Data Availability

Data related to the NHANES cohort is publicly available and can be accessed at <https://wwwn.cdc.gov/nchs/nhanes/>. Data related to the HHEAR cohort is publicly available and can be accessed at [HHEARdatacenter.mssm.edu](https://hheardatacenter.mssm.edu) under the following DOIs: [10.36043/2016-1407_EPI_68](https://doi.org/10.36043/2016-1407_EPI_68) and [10.36043/1407_118](https://doi.org/10.36043/1407_118).

References:Uncategorized References

1. Barr DB, et al. , Urinary creatinine concentrations in the US population: implications for urinary biologic monitoring measurements. *Environmental health perspectives*, 2004. 113(2): p. 192–200.
2. Gagnebin Y, et al. , Metabolomic analysis of urine samples by UHPLC-QTOF-MS: Impact of normalization strategies. *Analytica chimica acta*, 2017. 955: p. 27–35. [PubMed: 28088278]
3. Hertel J, et al. , Dilution correction for dynamically influenced urinary analyte data. *Analytica chimica acta*, 2018. 1032: p. 18–31. [PubMed: 30143216]
4. Alessio L, et al. , Reliability of urinary creatinine as a parameter used to adjust values of urinary biological indicators. *International archives of occupational and environmental health*, 1985. 55(2): p. 99–106. [PubMed: 3988361]

5. LaKind JS, et al. , Factors affecting interpretation of national biomonitoring data from multiple countries: BPA as a case study. *Environmental research*, 2019. 173: p. 318–329. [PubMed: 30951958]
6. Meeker JD, et al. , Temporal variability of urinary levels of nonpersistent insecticides in adult men. *Journal of Exposure Science and Environmental Epidemiology*, 2005. 15(3): p. 271.
7. Muscat JE, Liu A, and Richie JP Jr, A comparison of creatinine vs. specific gravity to correct for urinary dilution of cotinine. *Biomarkers*, 2011. 16(3): p. 206–211. [PubMed: 21288164]
8. O'Brien KM, et al. , Environmental chemicals in urine and blood: improving methods for creatinine and lipid adjustment. *Environmental health perspectives*, 2015. 124(2): p. 220–227. [PubMed: 26219104]
9. Middleton DR, Watts MJ, and Polya DA, A comparative assessment of dilution correction methods for spot urinary analyte concentrations in a UK population exposed to arsenic in drinking water. *Environment international*, 2019. 130: p. 104721. [PubMed: 31207477]
10. O'Brien KM., Upson K, and Buckley JP, Lipid and creatinine adjustment to evaluate health effects of environmental exposures. *Current environmental health reports*, 2017. 4(1): p. 44–50. [PubMed: 28097619]
11. Balshaw DM, et al. , The Children's Health Exposure Analysis Resource (CHEAR): Enabling Research into the Environmental Influences on Children's Health Outcomes. *Current opinion in pediatrics*, 2017. 29(3): p. 385. [PubMed: 28383342]
12. Van der Steen J, et al. , Benefits and pitfalls of pooling datasets from comparable observational studies: combining US and Dutch nursing home studies. *Palliative Medicine*, 2008. 22(6): p. 750–759. [PubMed: 18715975]
13. Statistics N.C.f.H., National Health and Nutrition Examination Survey: questionnaires, datasets, and related documentation. 2018.
14. Phipatanakul W, et al. , The school inner-city asthma intervention study: design, rationale, methods, and lessons learned. *Contemporary clinical trials*, 2017. 60: p. 14–23. [PubMed: 28619649]
15. Phipatanakul W, et al. , Effect of School Integrated Pest Management or Classroom Air Filter Purifiers on Asthma Symptoms in Students With Active Asthma: A Randomized Clinical Trial. *JAMA*, 2021. 326(9): p. 839–850. [PubMed: 34547084]
16. Taussky HH and Kurzmann G, A microcolorimetric determination of creatine in urine by the Jaffe reaction. *Journal of Biological Chemistry*, 1954. 208: p. 853–861. [PubMed: 13174594]
17. Cone EJ, et al. , Normalization of urinary drug concentrations with specific gravity and creatinine. *Journal of analytical toxicology*, 2009. 33(1): p. 1–7. [PubMed: 19161663]
18. Gompertz B, XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c. *Philosophical transactions of the Royal Society of London*, 1825(115): p. 513–583.
19. Geisser S, The predictive sample reuse method with applications. *Journal of the American statistical Association*, 1975. 70(350): p. 320–328.
20. Kuiper JR, et al. , Urinary specific gravity measures in the US population: Implications for the adjustment of non-persistent chemical urinary biomarker data. *Environment international*, 2021. 156: p. 106656. [PubMed: 34062395]
21. García-Esquinas E, et al. , Kidney function and tobacco smoke exposure in US adolescents. *Pediatrics*, 2013. 131(5): p. e1415–e1423. [PubMed: 23569089]
22. Jhee JH, et al. , Secondhand smoke and CKD. *Clinical Journal of the American Society of Nephrology*, 2019. 14(4): p. 515–522. [PubMed: 30846462]
23. Kartha GK, et al. , Co-occurrence of asthma and nephrolithiasis in children. *PloS one*, 2017. 12(1): p. e0168813. [PubMed: 28081237]
24. Steyerberg EW and Vergouwe Y, Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European heart journal*, 2014. 35(29): p. 1925–1931. [PubMed: 24898551]
25. Garde A, Hansen ÅM, and Kristiansen J, Evaluation, including effects of storage and repeated freezing and thawing, of a method for measurement of urinary creatinine. *Scandinavian journal of clinical and laboratory investigation*, 2003. 63(7–8): p. 521–524. [PubMed: 14743962]

26. Pearso MA., et al. ., Evaluation of physiological measures for correcting variation in urinary output: Implications for assessing environmental chemical exposure in children. *Journal of exposure science & environmental epidemiology*, 2009. 19(3): p. 336–342. [PubMed: 18841168]
27. Wang B, et al. ., Influence of body mass index status on urinary creatinine and specific gravity for epidemiological study of children. *European journal of pediatrics*, 2015. 174(11): p. 1481–1489. [PubMed: 25956275]
28. Carrieri M, Trevisan A, and Bartolucci GB, Adjustment to concentration-dilution of spot urine samples: correlation between specific gravity and creatinine. *International archives of occupational and environmental health*, 2000. 74(1): p. 63–67.
29. Sauvé J-F, et al. ., Creatinine and specific gravity normalization in biological monitoring of occupational exposures. *Journal of occupational and environmental hygiene*, 2015. 12(2): p. 123–129. [PubMed: 25192246]
30. Trevisan A.J.A.j.o.i.m., Concentration adjustment of spot samples in analysis of urinary xenobiotic metabolites. 1990. 17(5): p. 637–642.
31. Boeniger MF, Lowry LK, and Rosenberg J, Interpretation of urine results used to assess chemical exposure with emphasis on creatinine adjustments: a review. *American Industrial Hygiene Association Journal*, 1993. 54(10): p. 615–627. [PubMed: 8237794]
32. Weaver VM, et al. ., Challenges for environmental epidemiology research: are biomarker concentrations altered by kidney function or urine concentration adjustment? *Journal of exposure science & environmental epidemiology*, 2016. 26(1): p. 1–8. [PubMed: 25736163]
33. Mendiola J, et al. ., Urinary concentrations of di (2-ethylhexyl) phthalate metabolites and serum reproductive hormones: pooled analysis of fertile and infertile men. *Journal of andrology*, 2012. 33(3): p. 488–498. [PubMed: 21597090]
34. Suwazono Y, et al. ., Creatinine versus specific gravity-adjusted urinary cadmium concentrations. *Biomarkers*, 2005. 10(2–3): p. 117–126. [PubMed: 16076727]
35. Xia Y, et al. ., Comparison of Creatinine and Specific Gravity for Hydration Corrections on Measurement of the Tobacco-Specific Nitrosamine 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanol (NNAL) in Urine. *Journal of clinical laboratory analysis*, 2014. 28(5): p. 353–363. [PubMed: 24648246]

Highlights

- K-fold validation can be used to estimate specific gravity from urinary creatinine
- Individual characteristics related to dilution improve model estimation
- Calibration plots indicate that specific gravity estimation performs well on average
- Effect estimates using measured or estimated specific gravity are similar

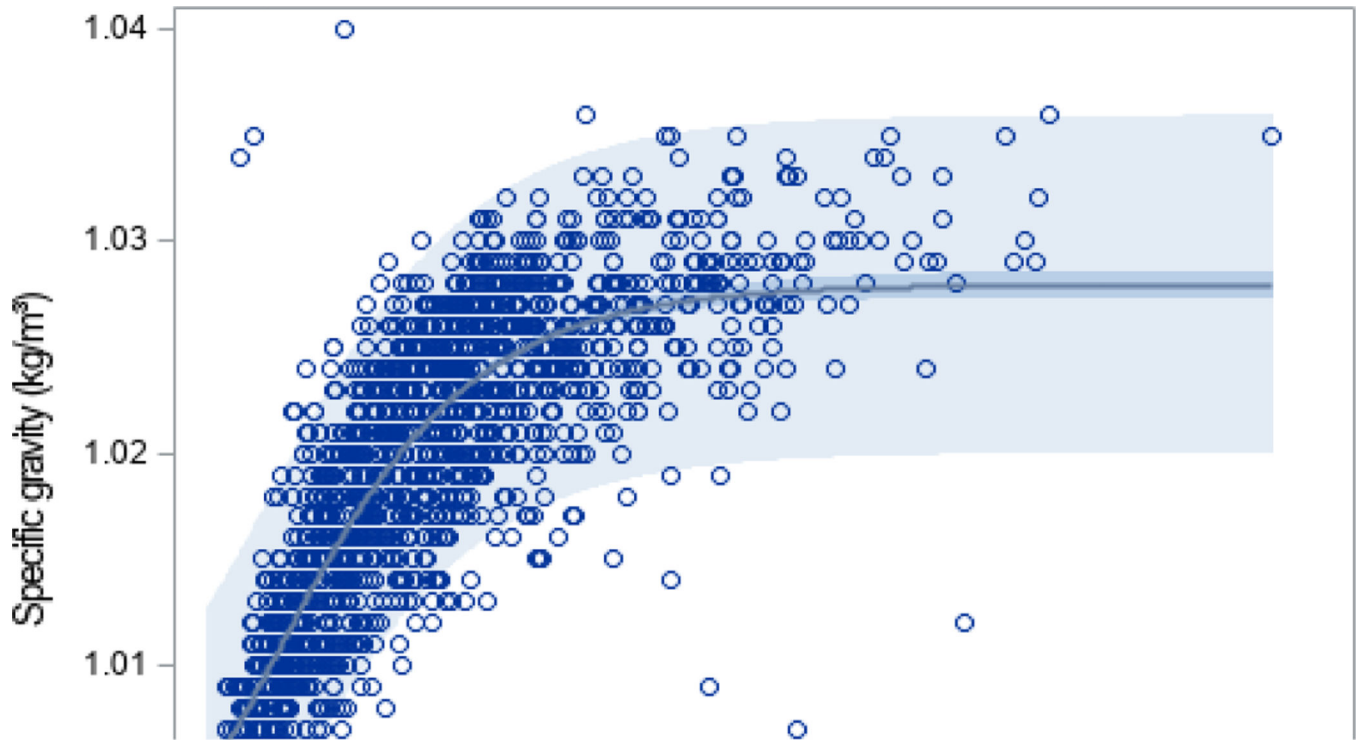


Figure 1:
SG and UCr were measured in 1530 elementary school aged children from NHANES cycle 2007–2008.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

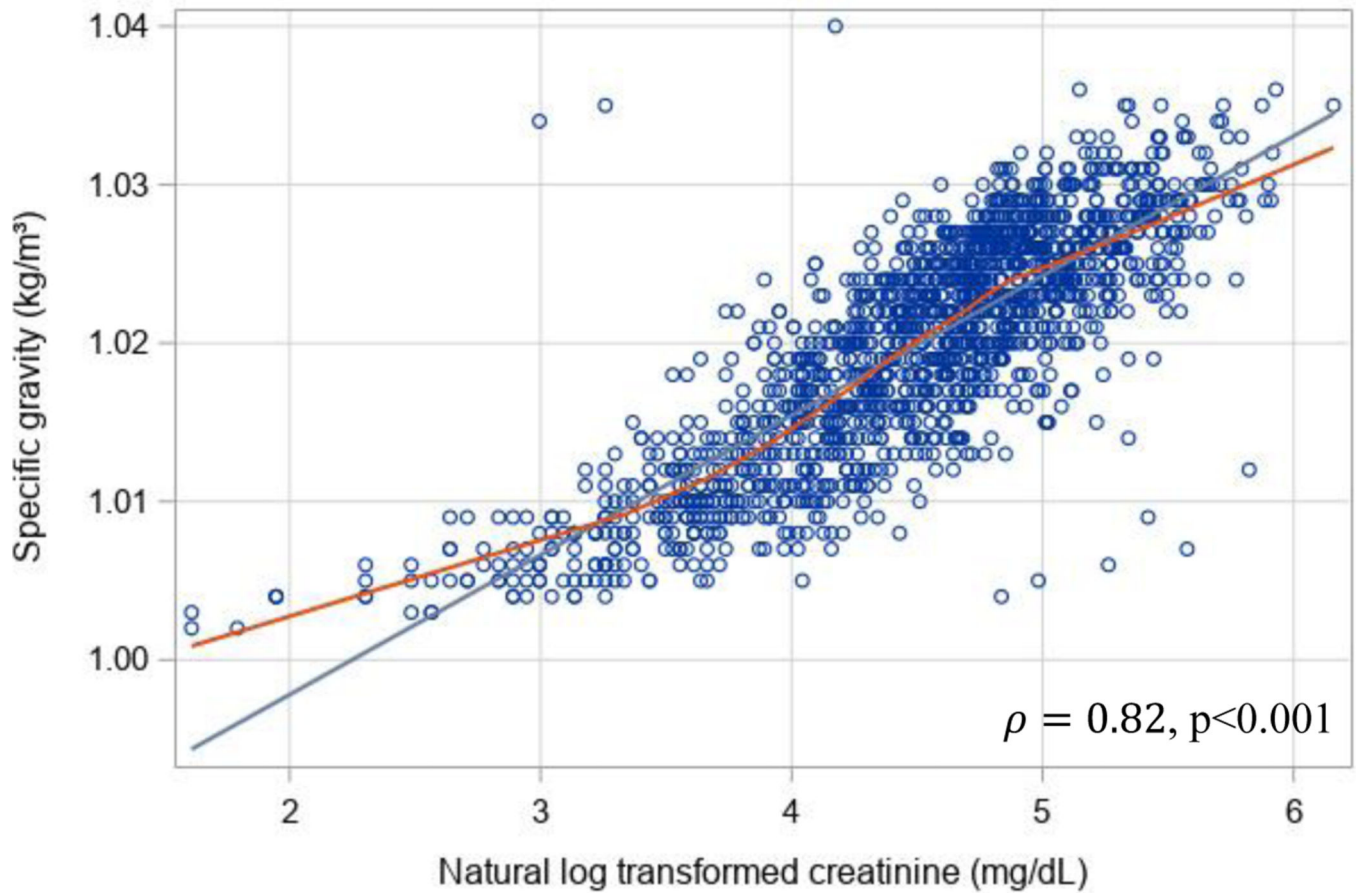


Figure 2: Linear relationship between natural log-transformed UCr and SG in the training sample, showing the correlation between measures. A smoothing line (in red) indicates a possible sigmoidal relationship between log transformed UCr and SG.

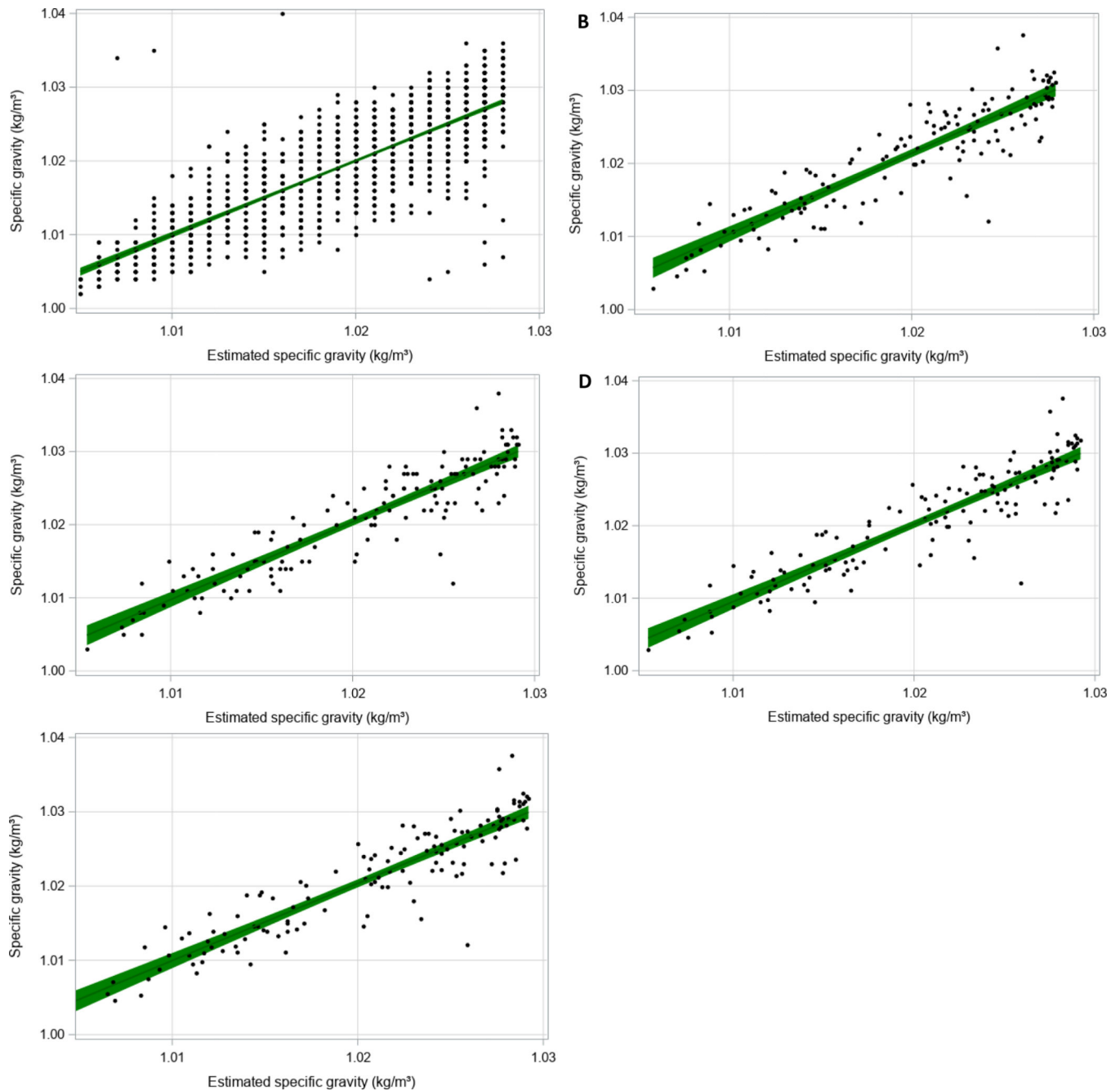


Figure 3: Calibration plots showing the linear relationship between estimated SG and measured SG using a nonlinear model accounting for UCr, using NHANES as a training dataset (N=1530). Parameter estimates can be found in Table 3. Data estimated for model 1 is shown in (A) for the training cohort and (B) for the validation cohort. Data estimated from model 2 is shown in (C) for validation cohort and model 3 in (D) for validation cohort and model 4 in (E) for validation cohort.

Table 1:

Descriptive statistics stratified by HHEAR (validation) and NHANES (training) cohorts

	HHEAR (N=139)	NHANES (N=1530)
	Freq (%)	Freq (%)
<i>Sex</i>		
Boys	74 (54.2)	779 (50.9)
Girls	65 (46.8)	751 (49.1)
<i>Race</i>		
White, non-Hispanic	14 (10.1)	437 (28.6)
Black, non-Hispanic	33 (23.7)	419 (27.4)
Hispanic	89 (64.0)	605 (39.5)
Other/multiracial	3 (2.2)	69 (4.5)
<i>Session</i>		
Morning	36 (25.9)	629 (41.1)
Afternoon	80 (57.5)	570 (37.3)
Evening	23 (16.6)	331 (21.6)
	Mean (SD) [Range]	Mean (SD) [Range]
<i>Age (years)</i>	8.19 (1.77) [6.00–14.00]	9.79 (2.51) [6.00–14.00]
<i>BMI (kg/m²)</i>	19.6 (4.68) [13.7–37.2]	20.0 (5.17) [12.5–48.5]
<i>BMI percentile</i>	72.3 (26.6) [4.00–99.7]	66.0 (30.1) [0.0138–99.9]
<i>NNAL (ng/mL)^a</i>	0.0133 (0.0173) [0.0074–0.117]	0.0116 (0.0392) [0.0004–0.649]
<i>Creatinine (mg/dL)</i>	109 (66.0) [10.9–304]	108.9 (65.2) [5.00–473]
Boys	105 (62.3)	113 (63.4)
Girls	114 (70.2)	104.6 (66.8)
Young children (6–8) ^b	95.8 (57.3)	83.5 (46.3)
Older children (9–14) ^c	132 (74.2)	123 (69.8)
White, non-Hispanic	57.1 (35.2)	105 (62.3)
Black, non-Hispanic	130 (65.4)	123 (73.8)
Hispanic	106 (63.6)	103 (60.3)
Other/multiracial	215 (56.6)	100 (55.9)
<i>SG (kg/m³)</i>	1.021 (0.008) [1.003–1.038]	1.020 (0.010) [1.002–1.040]
Boys	1.021 (0.007)	1.021 (0.007)
Girls	1.021 (0.008)	1.019 (0.007)
Young children (6–8) ^b	1.020 (0.008)	1.019 (0.007)
Older children (9–14) ^c	1.022 (0.007)	1.020 (0.007)
White, non-Hispanic	1.015 (0.007)	1.020 (0.007)
Black, non-Hispanic	1.023 (0.006)	1.020 (0.008)
Hispanic	1.021 (0.008)	1.020 (0.007)
Other/multiracial	1.028 (0.006)	1.019 (0.007)

^aFor HHEAR cohort N=132; for NHANES cohort N=1413

^bFor the HHEAR cohort N=90; for NHANES cohort N=552

^cFor the HHEAR cohort N=51; for NHANES cohort N=978

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2:

P-values from spearman correlations between demographics and log transformed UC_r and SG in the HHEAR (validation) and NHANES (training) cohorts.

	HHEAR (N=139)		NHANES (N=1530)	
	UC_r^a	SG	UC_r^a	SG
Sex ^b	0.629	0.641	<0.001	<0.001
Age ^c	<0.001	0.067	<0.001	<0.001
BMI ^c	0.203	0.473	<0.001	<0.001
Race ^d	<0.001	0.005	<0.001	0.799
Session time ^d	0.439	0.190	<0.001	<0.001
NNAL ^{a,c}	0.002	0.004	<0.001	<0.001
Asthma ^b	---	---	0.290	0.672

^a Natural log transformed

^b p-values are calculated from t-test of UC_r or SG between groups.

^c p-values are calculated from pearson correlation

^d p-values are calculated from ANOVA of UC_r or SG between groups.

Table 3:

Average parameter estimates from k-fold validation of each nonlinear model taking the form of equation 1, generated using the training cohort (N=1530).

	Model 1	Model 2	Model 3	Model 4 ^a
	<u>Avg Estimate (SE)</u>	<u>Avg Estimate (SE)</u>	<u>Avg Estimate (SE)</u>	<u>Avg Estimate (SE)</u>
α	0.0281 (0.0002)	0.0293 (0.0002)	0.0293 (0.0004)	0.0293 (0.0003)
β	0.0194 (0.0007)	0.0195 (0.0002)	0.0192 (0.0009)	0.0196 (0.0006)
δ	34.5 (0.710)	32.7 (1.65)	33.0 (1.32)	32.5 (1.27)
θ_{age}	---	0.0283 (0.005)	0.0309 (0.008)	0.0347 (0.012)
θ_{sex}	---	0.068 (0.119)	0.0723 (0.010)	0.0787 (0.006)
θ_{BMI}	---	---	-0.00156 (0.002)	-0.0005 (0.002)
θ_{black}	---	---	0.0615 (0.034)	0.0678 (0.046)
$\theta_{hispanic}$	---	---	-0.0409 (0.023)	-0.0230 (0.028)
$\theta_{Other/multiracial}$	---	---	-0.0358 (0.023)	-0.0820 (0.024)
$\theta_{afternoon}$	---	---	0.00726 (0.015)	0.0320 (0.008)
$\theta_{evening}$	---	---	-0.0575 (0.008)	-0.0455 (0.005)
θ_{NNAL}	---	---	---	0.539 (0.142)
θ_{asthma}	---	---	---	0.0261 (0.025)
γ_{age}	---	0.00137 (0.0001)	0.00132 (0.0001)	0.00132 (0.0001)
γ_{sex}	---	0.000768 (0.0002)	0.000915 (0.0003)	0.000777 (0.0001)
γ_{BMI}	---	---	-5.09E-6 (0.00004)	0.0000306 (0.000007)
γ_{black}	---	---	0.00222 (0.0006)	0.00231 (0.0007)
$\gamma_{hispanic}$	---	---	-0.000148 (0.0003)	-0.000323 (0.0003)
$\gamma_{Other/multiracial}$	---	---	0.00127 (0.009)	0.00245 (0.0004)
$\gamma_{afternoon}$	---	---	-0.00162 (0.0002)	-0.00160 (0.00008)
$\gamma_{evening}$	---	---	-0.00284 (0.0004)	-0.00271 (0.0006)
γ_{NNAL}	---	---	---	0.00494 (0.001)
γ_{asthma}	---	---	---	0.000284 (0.0005)

^aN=1413 due to missing NNAL

SE=Standard error

Table 4:

Summary of beta coefficients, assessing the linear association between measured SG and estimated SG from each nonlinear model described in Table 3.

Validation Dataset	Model #	Adjusted R ²	Beta	95% CI	p-value
NHANES (N=1530)	1	0.691	1.00	0.97, 1.04	<0.001
HHEAR (N=139)	1	0.803	1.10	1.01, 1.19	<0.001
	2	0.826	1.06	0.98, 1.15	<0.001
	3	0.829	1.06	0.98, 1.14	<0.001
	4 ^a	0.826	1.01	0.93, 1.09	<0.001

^aN=132 due to missing NNAL

Table 5:

Association of MBZP with asthma symptoms days in HHEAR cohort using zero inflated Poisson regression, controlling for the dilution factor as a covariate where the dilution is either log transformed UCr, measured SG, or SG estimated from the 3 models described in Table 3 (N=139).

Method	Rate Ratio	95% CI	p-value
No dilution	1.26	(1.05, 1.52)	0.013
Log transformed UCr	1.34	(1.09, 1.64)	0.005
Measured SG	1.28	(1.06, 1.55)	0.012
SG estimated from Model 1	1.34	(1.09, 1.64)	0.006
SG from estimated Model 2	1.33	(1.08, 1.63)	0.007
SG estimated from Model 3	1.33	(1.08, 1.63)	0.007
SG estimated from Model 4	1.30	(1.06, 1.60)	0.012
z-scored dilution	1.37	(1.07, 1.75)	0.013
covariates	1.25	(1.05, 1.50)	0.013