

Original Article

A sociodemographic index identifies sex-related effects on insomnia in the Hispanic Community Health Study/Study of Latinos

Natali Sorajja^{1,2,†}, Joon Chung^{2,†}, Carmela Alcántara³, Sylvia Wassertheil-Smoller⁴, Frank J Penedo⁵, Alberto R Ramos⁶, Krista M Perreira⁷, Martha L Daviglius⁸, Shakira F Suglia⁹, Linda C Gallo¹⁰, Peter Y Liu¹¹, Susan Redline², Carmen R Isasi⁴ and Tamar Sofer^{1,2,12,*}

¹Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA,

²Division of Sleep and Circadian Disorders, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA,

³School of Social Work, Columbia University, New York, NY, USA,

⁴Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, NY, USA,

⁵Department of Psychology, University of Miami, Miami, FL, USA,

⁶Department of Neurology, University of Miami Miller School of Medicine, Miami, FL, USA,

⁷Department of Social Medicine, University of North Carolina School of Medicine, Chapel Hill, NC, USA,

⁸Institute for Minority Health Research, University of Illinois at Chicago, Chicago, IL, USA,

⁹Department of Epidemiology, Rollins School of Public Health, Atlanta, GA, USA,

¹⁰Department of Psychology, San Diego State University, San Diego, CA, USA,

¹¹Division of Genetics, Lundquist Institute at Harbor-UCLA Medical Center, Torrance, CA, USA and

¹²CardioVascular Institute (CVI), Beth Israel Deaconess Medical Center, Boston, MA, USA

[†]These authors equally contributing authors.

*Corresponding author. Tamar Sofer, Center for Life Sciences CLS-934, 3 Blackfan Street, Boston, MA 02115, USA. Email: tsofer@bidmc.harvard.edu.

Abstract

Study Objectives: Sex differences are related to both biological factors and the gendered environment. We constructed measures to model sex-related differences beyond binary sex.

Methods: Data came from the baseline visit of the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). We applied the least absolute shrinkage and selection operator penalized logistic regression of male versus female sex over sociodemographic, acculturation, and psychological factors jointly. Two “gendered indices,” the gendered index of sociodemographic environment (GISE) and gendered index of psychological and sociodemographic environment, summarizing the sociodemographic environment (GISE) and psychosocial and sociodemographic environment (GIPSE) associated with sex, were calculated by summing these variables, weighted by their regression coefficients. We examined the association of these indices with insomnia, a phenotype with strong sex differences, in sex-adjusted and sex-stratified analyses.

Results: The distribution of GISE and GIPSE differed by sex with higher values in male individuals. In an association model with insomnia, male sex was associated with a lower likelihood of insomnia (odds ratio [OR] = 0.60, 95% CI [0.53, 0.67]). Including GISE in the model, the association was slightly weaker (OR = 0.63, 95% CI [0.56, 0.70]), and weaker when including instead GIPSE in the association model (OR = 0.78, 95% CI [0.69, 0.88]). Higher values of GISE and of GIPSE, more common in the male sex, were associated with a lower likelihood of insomnia, in analyses adjusted for sex (per 1 standard deviation of the index, GISE OR = 0.92, 95% CI [0.87, 0.99], GIPSE OR = 0.65, 95% CI [0.61, 0.70]).

Conclusions: New measures such as GISE and GIPSE capture sex-related differences beyond binary sex and have the potential to better model and inform research studies of sleep health.

Key words: sex differences; gendered environment; gendered indices; sociodemographic sex-specific associations; insomnia; sleep health

Submitted for publication: April 17, 2024; Revised: July 19, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Sleep Research Society.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Statement of Significance

Sex differences are complex, as they are determined by sex chromosomes and reproductive organs, while lifestyle and environmental factors also interact with these determinants to influence the expression of biological sex. Here, we used data from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) to develop “gendered indices”: variables that have different distributions in male and female individuals, but that are made up completely of non-biological variables. We show that the gendered indices explain some of the (binary) sex variable association with insomnia. Further, within male and female strata, the gendered indices are associated with insomnia with the same effect size estimate as in the combined dataset. This approach has the potential to improve modeling and understanding of determinants of sex differences in insomnia and other health measures, and ultimately improve health disparities.

Sex differences are observed in many aspects of health [1–4]. Modern genomics, epigenomics, and other omics technologies have advanced the study of sex differences [5–7]. Yet, the relationship between sex and health is more complicated than determined by binary definitions of sex as male or female. Table 1 provides an overview of terms and definitions related to sex and gender that are relevant for public health research and for this manuscript. First, sex differences are driven by a range of potential biological sex effects. Here, “biological sex effects” refer to the collection of biologically measurable quantities related to sex, such as sex chromosome combinations, X-linked genes, reproductive organs and history, and sex hormone levels. The latter are related to, but are not completely determined by, the genetic categories of male and female sex as XY and XX sex chromosome combinations, which characterize more than 99% of individuals [9–11]. Sex hormone levels change with age, for all individuals regardless of chromosomal combinations. Thus, biological sex effects encompass a range of nonbinary, time-varying, effects. Second, additional health effects are due to sociocultural factors reflecting “gendered” structural and social environment, i.e. differences in occupation, income, educational attainment, family roles, etc., experienced by men, women, and nonbinary individuals [12]. The gendered environment of an individual is impacted by their sex assigned at birth (Figure 1), as well as by their gender identity and other gender dimensions [8]. Finally, sociocultural “gender”-related factors may also intersect, affect, and interact with biological sex effects [13, 14]. Indeed, both biological effects and dimensions of gender have been shown to contribute

to differences in cardiovascular outcomes between women and men [15]. With this complexity in mind, we note that studies of sex differences in health, compared to studies examining sex- or gender-specific health determinants, still primarily focus on strata of men and women (sometimes implicitly assuming a one-to-one correspondence between binary sex and gender identity of man or woman). Thus, even when other biological and sociocultural factors are utilized, they are ultimately used to explain differences between two dichotomous groups. This manuscript addresses the issue of assessing sex differences within the context of two sex strata while improving upon the dichotomization of sex.

A prime example of the multi-faceted sex-related effects on health is sleep, because it demonstrates strong sex differences [16, 17], with some differences clearly driven by biology [18] and others likely driven by the gendered environment [19, 20], including occupation, health status, and social habits ([21, 22]. The association is also bidirectional because sleep habits/exposures such as sleep restrictions and sleep patterns related to shift work, may affect hormonal levels, including sex hormones which impact biological sex differences [23–25]. Furthermore, of all behaviors, the impact of sleep is of major interest because it is intimately related to many if not all aspects of health in general [26, 27]. For these reasons, it is important to try to disentangle sex-related effects on sleep, and more generally, on health and disease, to better understand modifiable components of sleep health, in men, women, and all individuals.

Here, we develop gendered indices, which are data-reducing summaries of non-biological, sociodemographic measures. They

Table 1. Sex- and Gender-Related Terms Used in this Manuscript

Term/expression	Meaning and context in this manuscript
Sex differences	Differences between two groups: females and males. These groups may not be defined biologically or by gender identity, often because appropriate data are lacking, i.e. there may not be information about sex chromosomes.
Biological sex effects	Objective measures of biological factors related to sex, such as sex chromosomes, sex hormones, reproductive organs and history, menopause, and puberty.
Sex assigned at birth	Sex is assigned at birth, typically by a physician according to observed physiological organs.
Binary sex	The categorization of individuals to male and female sex, typically with the intention to reflect underlying binary biological sex. Historically, binary sex definitions may not be accurate as individuals may be categorized into binary sex groups without an explicit check of the biological definition of sex chromosomes, and are categorized instead based on gender identity or perceived gender. In research that does not focus on sex and/or gender minorities, Individuals may be excluded from analyses focusing on binary sex.
Gender dimensions	Barr et al. [8] described four gender dimensions that relate to health: identity and expression; roles and norms; relations; and power.
Gender identity	A person’s sense of their own gender. In this study, we do not have data about self-reported gender identity.
Gendered environment	The idea is that individuals of different gender expressions experience the world differently through differences in the environment available to them. This is reflected in differences in the distribution of demographic and sociocultural measures between individuals of different sexes assigned at birth (here focus on males and females).

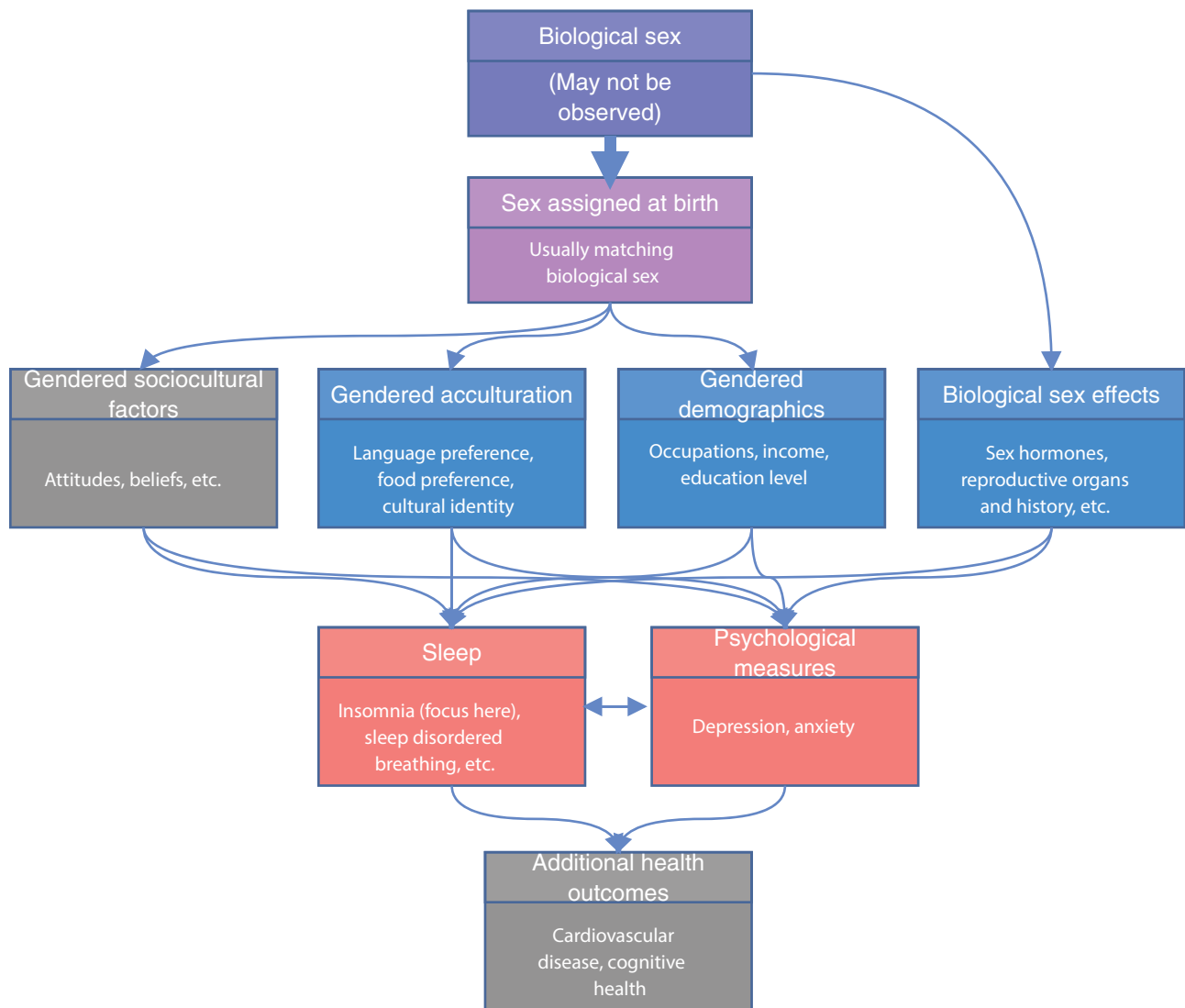


Figure 1. Model of sex and gender effects on sleep and health. Conceptual model of sex and gender effects on sleep. Biological sex (male, female, or intersex) is determined pre-birth though it may not be observed, sex is assigned at birth, usually matching the biological sex as more than 99% of individuals are either male or female. Biological sex, where in this work we focus on binary sex, leads to biological sex-related factors, and sex assigned at birth leads to demographic, acculturation, and sociocultural (not well measured in our data, hence grayed) factors via their gendered characteristics. These may affect psychological measures and sleep, which interact. Downstream, these may affect other health outcomes (not investigated here, hence grayed).

are not developed to reflect nor to quantify gender identity, but rather, to summarize non-biological measures experienced by men and women independent of the direct effects of sex. We use the term “gendered” rather than, say, “sexed,” because differences in the indices distribution between male and female individuals are due to non-biological measures that are available in this dataset. These indices are constructed using measures available in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) that are associated with binary sex. Our goal is to concisely summarize non-biological measures related to the gendered environment into a “gendered index” to be used when studying sex differences and sex-related determinants of health. We demonstrate the utility of these indices in studying sex differences in insomnia, a sleep disorder with strong sex differences—insomnia is approximately 1.5 times more common in female compared to male adults [28, 29]. HCHS/SOL participants are Hispanic and Latino adults, typically first- (i.e. foreign-born with foreign-born

parents) or second-generation (i.e. foreign-born with at least one US-born parent) immigrants to the United States, and often of low socioeconomic status. As such, we included measures of acculturation (i.e. of the level of assimilation, or acceptance, of the dominant US culture compared to the country-of-origin culture [30]) and, in secondary analysis, psychological measures of depression and anxiety, as these are also determinants and/or correlates of both sleep health and health in general, which may show gendered patterns.

Methods

The Hispanic Community Health Study/Study of Latinos

Data were collected from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) epidemiologic study. The purpose of the HCHS/SOL has been to evaluate the prevalence

of cardiovascular disease (CVD) risk factors, and to study the roles of CVD risk factors, socioeconomic, and genetic factors in the development of cardiovascular disease in the Hispanic/Latino population [31–33]. HCHS/SOL participants were selected for the study via a two-stage sampling design as previously described [34]. In brief, the target population was defined as all non-institutionalized, self-identified, Hispanic/Latino adults aged 18–74 years, from four US metropolitan areas: Miami, San Diego, Chicago, and the Bronx area of New York. Older adults aged 45–74 were oversampled, but there were no other inclusion or exclusion criteria. Our project used data from the baseline HCHS/SOL visit which took place in 2008–2011 and included 16 415 participants. At baseline, sex was included in the questionnaire form and labeled “Gender” with potential responses being “Male” or “Female.” We refer to this variable as “reported sex,” because it was ascribed by the interviewer. This variable likely represents assigned sex at birth and may not describe gender identity. No options for categorization as intersex/difference in sexual development or identification as nonbinary or transgender were provided at baseline. However, these options have been added as part of HCHS/SOL Visit 3 which was completed in early 2024 [35] and will become available in the future.

Insomnia symptoms measure

Insomnia symptoms were assessed in the HCHS/SOL baseline exam using the Women Health Initiative Insomnia Rating Scale (WHIIRS [36]). The WHIIRS is a 5 items instrument querying participants’ symptoms of insomnia in the past 4 weeks. The first four items are the questions: “Did you have trouble falling asleep?,” “Did you wake up several times at night?,” “Did you wake up earlier than you planned to?,” and “Did you have trouble getting back to sleep after you woke up too early?.” The responses to these questions are “No, no in the past 4 weeks” (0 points), “Year, less than once a week” (1 point), “Yes, 1 or 2 times a week” (2 points), “Year, 3 or 4 times a week” (3 points), and “Yes, 5 or more times a week” (4 points). The fifth item is the statement “Overall, was your typical night’s sleep during the past 4 week,” with responses “Very sound or restful,” “Sound or restful,” “Average quality,” “Restless,” “Very restless” (points ranging from 0 to 4). Responses to all items are summed to indicate the severity of insomnia symptoms (with higher scores indicating more symptoms). The WHIIRS was developed by the Women’s Health Initiative (WHI) study researchers, where WHI is a large study of postmenopausal women [37]. The measure was assessed in populations of postmenopausal women with validity ascertained based on (1) test–retest reliability over the same day (correlation 0.96) and up to more than 1 year later (correlation 0.66) [36], (2) convergent validity, based on correlations of WHIIRS scores with other outcomes related to insomnia being in the expected direction [36, 38].

Insomnia was defined by dichotomizing the responses to the WHIIRS such that WHIIRS ≥ 10 was defined as having insomnia, and no insomnia if WHIIRS < 10 . This cutoff is used based on the original publication reporting that it resulted in a probability of 0.65 to correctly predict those with insomnia and those without [36]. Because a dichotomized insomnia measure has a more intuitive interpretation, our analyses focused on the dichotomized insomnia measure. However, we repeated some analyses with WHIIRS quantified continuously as secondary results, as described in the relevant sections.

Demographic, acculturation, and psychosocial measures, and assessment of associations with sex and insomnia

We selected sociodemographic variables that are typically associated with sex from the HCHS/SOL study. These variables included marital status, household income level, employment status, longest-held occupation type, language preference, language acculturation subscale, social acculturation subscale, both extracted from the short acculturation scale for Hispanics [39], ethnic identity score (average of two 7-point Likert scale items: “I have a strong sense of belonging to my own ethnic group” and “I have a lot of pride in my ethnic group”), current health insurance status, number of years lived in the United States, number of years of education, and psychological scores: the Spielberger Trait Anxiety Inventory scale (STAI10) measuring anxiety [40], and the Center for Epidemiological Studies Depression scale (CESD10) measuring depression symptoms [41]. Because one of the CESD10 questions refers to quality of sleep (“my sleep was restless”), and we are interested in insomnia which is related, we recomputed the responses to this questionnaire, to generate “CESD9,” excluding the sleep-related question. For interpretation purposes, we re-weighted the resulting score, dividing by 9 and multiplying by 10, so it is on the same scale as the standard CESD10. In all association analyses, we controlled for the covariates age, Hispanic/Latino background, and study center. We refer to these as covariates, while we refer to other variables explicitly as sex, socioeconomic variables, or psychological variables. We used the R *survey* package [42] to characterize the study population by sex, accounting for survey design and weights so that estimates are generalizable for the HCHS/SOL target population.

We evaluated the associations of these variables with sex using multinomial regression; each variable was treated as an outcome and individually regressed over sex as an exposure (adjusted to covariates) using the *svy_VGAM* package (version 1.2) to allow for the inclusion of multiple levels of a variable within the sample survey analysis framework. We next used multivariable survey logistic regression to measure the association of each sociodemographic variable with insomnia while adjusting for covariates, and with and without adjusting for sex. In these analyses, insomnia was the outcome variable. In secondary analysis, we also evaluated the sex-specific associations of these variables with insomnia, in association analysis limited to each sex stratum, and tested for differences in associations between sex strata by fitting a model, using the sex-combined data, with a sex interaction term with the variable of interest. The *p*-value for the interaction test is the association *p*-value of the relevant multiplicative sex interaction term. Because interaction analyses have low statistical power when binary traits are used (such as sex), compared to main effect analyses, we also included sex-stratified and interaction testing with WHIIRS scores, rather than insomnia, as the outcome, as continuous outcomes have increased power compared to their dichotomized versions.

Missing data imputation

We visualized missing patterns in data using the R *nanian* package (version 1.0.0). We performed both a complete-case analysis (primary) and a sensitivity analysis based on imputed data. The sensitivity analysis was performed by imputing the data five times with the *mi* R package (version 1.1) and follow-up analysis in each imputed dataset as described below. The *mi* function performs

regression-based predictive imputation, and we used default values (other than the use of 5 imputations, whereas the default is 3).

Development of gendered indices: gendered index of sociodemographic environment and gendered index of psychological and sociodemographic environment

Using the sociodemographic variables in Table 2, we created two gendered indices: the gendered index of the sociodemographic environment (GISE; primary), which included all of the considered sociodemographic variables, and the gendered index of the psychological and sociodemographic environment (GIPSE; secondary), which included these variables as well as participants' CESD9 scores and STAI10 scores. These indices were created using a least absolute shrinkage and selection operator (LASSO) regression with the *glmnet* package (version 4.1-7). This model regressed sex on these variables with penalization and potential variable selection. A tuning parameter λ was used to control the level of penalization. Cross-validation was performed to select the tuning parameter value based on the minimization of the mean-squared-error. Next, the complete dataset was used with the selected tuning parameter value to train the LASSO model. The result of the LASSO regression is a set of coefficients corresponding to each sociocultural (GISE) or sociocultural and psychological (GIPSE) measure, with some coefficients being potentially zero. Both GISE and GIPSE were formed by extracting the fitted values from each of the two LASSO models. In the secondary analysis, we also computed indices in each of the imputed datasets and averaged the indices across the imputed dataset for visualization.

Secondary analysis: verifying the consistency of relationships between variables used and with reported sex

It is possible that the gendered indices are associated with sex due to overfitting the data, as they are constructed by optimizing their association with sex. To verify that GISE and GIPSE are indeed gendered (meaning, have different associations by sex), and that observed sex differences in the indices distributions are not the result of overfitting to the data, we ran a secondary analysis utilizing a training and testing set approach, by splitting the HCHS/SOL dataset into separate training and testing sets. Some of HCHS/SOL individuals were sampled from the same household. Such individuals may share similar sociodemographic and cultural environments. To ensure that the training and testing sets are independent, we verified that if an individual was in the training set, then another HCHS/SOL individual from the same household was not in the testing set. Thus, the training and testing sets were created by grouping all individuals with the same primary sampling unit (PSU) identification number into only the training or only the testing set. Finally, the training/testing sets were created using a 70%–30% split of the PSUs.

In another secondary analysis, we assessed whether the relationship among sociodemographic measures differed between males and females. We performed principal component analysis within each sex stratum. We compared via visualization the loadings of the first male-specific and female-specific PCs to study whether the distribution of sociodemographic measures appears different between the sex strata.

Association analysis of sex and gendered indices with insomnia

We scaled the indices by dividing them by their standard deviation as computed in the complete-case analysis to increase the interpretability of the otherwise dimensionless variables. To measure the associations of GISE and GIPSE with insomnia, insomnia was regressed on each gendered index separately in a survey logistic regression, adjusted for the covariates, including sex. The coefficient of reported sex and of the gendered indices were used to measure, correspondingly, the effects of binary sex and of the gendered environment independently of binary sex effects. The proposed interpretation of the estimated associations of GISE and GIPSE with insomnia is that they represent sex-related effects that are explained by demographic, sociodemographic, and psychological (GIPSE) variables, on insomnia.

Because our working assumption is that the gendered indices mediate sex effects on insomnia via the sociodemographic gendered environment, we also fit a detailed association model in which insomnia was regressed on reported sex while adjusting for all the components of each of the indices, instead of GISE and GIPSE themselves. The goal was to assess whether the estimated effect of reported sex is similar in the case where only the gendered index is used compared with the case where each component variable is used individually. Finally, to analyze the relationship between insomnia and each of the indices within each sex group, we also performed sex-stratified analyses.

In addition to the estimation of sex and gendered indices associations, as well as of the sociodemographic variables when used separately, we also computed the pseudo- R^2 , a measure of model prediction that is applicable for complex survey analyses [43]. This was done in the sex-combined and sex-stratified analyses.

Because association analysis with a binary, dichotomized, measure usually has lower statistical power compared with association analysis with the underlying continuous measures, we repeated the same analyses with continuously assessed WHIIRS scores as the outcome. The analyses followed the same approach other than the use of survey linear, rather than logistic regression.

Sensitivity analysis: gendered indices in imputed datasets

A new index was computed for each imputed dataset. Then, for each dataset, we repeated the association analysis with insomnia and applied Rubin's rule to combine the estimated effects. We also computed averaged indices across the five imputed datasets and studied their sex-specific distributions. Imputed indices may be useful for optimizing sample sizes of other analyses.

Interaction analyses of gendered indices association with insomnia by sex

In another analysis, we tested whether the associations of GISE and GIPSE with insomnia further differed by sex. We performed combined-sex association analyses with insomnia and WHIIRS as the outcomes, using the same regression models as before, but now with an additional multiplicative interaction term of GISE and sex and (separately) of GIPSE and sex.

Results

Table 2 characterizes the HCHS/SOL dataset and its target population. There were 9835 female and 6580 male participants. On average, female individuals in the HCHS/SOL target population were 46 years old, while male individuals were 45. In the target

Table 2. Characteristics of the HCHS/SOL Study and Target Population, Stratified by Sex

		Female	Male
n		9835	6580
Age (mean [SD])		46.54 (15.12)	44.81 (14.84)
Study site (%)	Bronx	2518 (30.3)	1600 (27.6)
	Chicago	2313 (14.6)	1821 (17.0)
	Miami	2353 (28.5)	1724 (30.1)
	San Diego	2651 (26.6)	1435 (25.3)
Background (%)	Dominican	963 (11.5)	510 (8.2)
	Central American	1049 (7.5)	683 (7.3)
	Cuban	1250 (18.3)	1098 (21.8)
	Mexican	4022 (38.1)	2450 (36.6)
	Puerto Rican	1589 (15.3)	1139 (17.1)
	South American	635 (5.2)	437 (4.7)
Education (%)	More than one/other heritage	276 (4.0)	227 (4.3)
	No high school diploma or GED	3768 (32.9)	2439 (31.8)
	At most high school diploma/GED	2353 (26.3)	1827 (30.2)
Income level (%)	Greater than high school/GED	3660 (40.8)	2277 (38.0)
	Less than \$10 000	1587 (17.6)	751 (11.5)
	\$10 001–\$20 000	3035 (33.7)	1834 (29.5)
	\$20 001–\$40 000	2873 (32.2)	2185 (34.4)
	\$40 001–\$75 000	1024 (12.5)	992 (16.7)
Marital status (%)	More than \$75 000	283 (4.0)	363 (7.8)
	Single	2552 (31.6)	1970 (37.9)
	Married or living with a partner	4726 (47.0)	3710 (50.8)
	Separated, divorced, or widow(er)	2505 (21.4)	864 (11.3)
Employment status (%)	Retired/not currently employed	5264 (56.9)	2689 (41.1)
	Employed part-time(≤35 hours/week)	1802 (18.8)	926 (15.0)
	Employed full-time(>35 hours/week)	2596 (24.3)	2832 (43.9)
Occupation (%)	Non-skilled worker	2725 (24.0)	2025 (27.5)
	Service worker	1702 (18.7)	669 (12.2)
	Skilled worker	1774 (17.7)	1680 (25.8)
	Professional/technical, administrative/executive, or office staff	1666 (19.6)	699 (11.7)
	Other occupation	1802 (20.0)	1383 (22.8)
	Current health insurance (%)	Yes	5065 (53.2)
Years in United States (%)	Less than 10 years	2328 (28.3)	1477 (27.0)
	10 years or more	5829 (50.8)	3798 (48.0)
	US born	1604 (20.9)	1259 (25.1)
Language preference (%)	Spanish	7997 (76.7)	5122 (72.9)
	English	1838 (23.3)	1458 (27.1)
SASH	Social acculturation subscale (mean [SD])	2.16 (0.60)	2.24 (0.59)
	Language acculturation subscale (mean [SD])	1.88 (1.12)	2.10 (1.17)
Ethnic identity score	Mean (SD)	3.20 (0.56)	3.22 (0.57)
Insomnia measures	Insomnia = Yes (%)	3424 (33.6)	1578 (23.6)
	WHIIRS (mean [SD])	7.68 (5.49)	6.22 (5.03)
Psychological measure	STAI10 (mean [SD])	17.98 (6.10)	16.18 (5.07)
	CESD9 (mean [SD])	10.57 (4.97)	9.15 (4.24)

Proportions, means, and standard deviations (SDs) were weighted using sampling weights to represent the HCHS/SOL target population. HCHS/SOL, Hispanic Community Health Study/Study of Latinos; GED, General Education Development; SASH, Social Acculturation Scale for Hispanics; WHIIRS, Women Health Initiative Insomnia Rating Scale. STAI10, Spielberger Trait Anxiety Inventory scale based on 10 items. CESD9, Center for Epidemiological Studies Depression scale based on 9 items (excluding a sleep quality-related question).

population, 33.6% of female individuals but only 23.6% of male individuals are estimated to have insomnia. Household income had the most missing values ($n = 1488$), with social acculturation subscale ($n = 729$) and insomnia ($n = 560$) following next. There were 13 666 complete cases out of 16 415 individuals, of which 8083 were female and 5593 were male participants.

Demographic, acculturation, and psychosocial measures driving sex differences in insomnia

Table 3 provides results from the multivariable regressions of each of the sociodemographic variables over sex, and the associations of these variables with insomnia. These association analyses were performed on each variable separately. These associations are further visualized in Figure 2. The results show that sociodemographic measures demonstrate sex-related patterns: all are associated with binary, reported sex. Further, most of the variables were also associated with insomnia, including when adjusting for reported sex. This supports the derivation of a combined gendered index to capture the multidimensionality of the gendered sociodemographic environment, expected to play a role in observed sex differences in insomnia.

In the secondary analysis, we also studied whether the associations of the sociodemographic variables with insomnia and insomnia symptoms potentially differ by sex stratum. The results are reported in Supplementary Table 1 (insomnia) and Supplementary Table 2 (WHIIRS scores). The results were qualitatively similar in analyses examining insomnia and WHIIRS scores, and only two associations had statistically significant sex interaction: education and anxiety score (STAI10), but the evidence of interaction was still weak. Having an education level greater than high school/GED was associated with a reduced likelihood of insomnia in female individuals (OR = 0.8, 95% CI [0.67, 0.96]), but with a higher likelihood of insomnia in male individuals (OR = 1.07, 95% CI [0.86, 1.35]), although the association among males was not statistically significant. The interaction p -value was .05. With STAI10 anxiety score, there was stronger evidence of interaction (interaction p -value = .007; STAI10 was associated with insomnia in both sex strata), with a stronger association of STAI10 scores with insomnia observed in males than in females. However, the estimated association of a unit increase of STAI10 with insomnia likelihood was still quite similar in the female and male strata: female OR = 1.12, 95% CI [1.10, 1.13], and male OR = 1.15, 95% CI [1.13, 1.17].

Sociodemographic gendered indices

Table 4 provides the coefficients of the variables used in GISE and GIPSE created using penalized regression. The penalized regression model was fit using the entire data set (complete cases). Variables labeled “Not Selected” were removed by the LASSO algorithm in the variable selection process. Figure 3 visualizes the distributions of GISE and GIPSE by sex, demonstrating that male participants tend to have higher values compared to female participants.

In secondary analysis, we split the data into independent training and testing datasets, where we developed gendered indices on the training dataset and then constructed them in the testing dataset, and assessed whether their distribution differed by sex group. Supplementary Figure 1 visualizes the distribution of the indices in the test dataset, demonstrating that their distribution differs by reported sex, and therefore, we conclude that the gendered indices are not associated with sex only due to overfitting. Further, the figure is stratified by age groups, demonstrating that

sex difference exists in the indices across adulthood. Interestingly, the median values of the gendered indices monotonically decline between the 30 and 40 age stratum to the 60 and 74 age stratum; the pattern is observed in both male and female participants.

In another sensitivity analysis, we used imputed datasets to develop gendered indices. Supplementary Figure 2 visualizes the missingness patterns in the data. Household income had a large number of observations with missing values ($n = 1488$), with social acculturation subscale ($n = 729$) and insomnia ($n = 560$) following next. There were 13 705 complete cases out of the 16 415 participants. Supplementary Figure 3 visualizes the distribution of the indices based on the imputed data, averaged over the five imputed values for each person. In the imputed dataset, males still tend to have higher values compared to female participants, but the difference was less pronounced than in the primary analysis and in the secondary analysis of the independent test dataset within the complete-case dataset, suggesting that individuals with missing data differ from complete cases in a non-ignorable manner.

Association of gendered indices with insomnia

Table 5 provides results from the association analysis of reported sex and the developed gendered indices in multiple regression models. In a model adjusting for covariates, self-reported male sex was associated with a lower likelihood of having insomnia symptoms (OR = 0.61, 95% CI [0.53, 0.67]). The association was attenuated after adjustment to GISE (OR = 0.64, 95% CI [0.57, 0.71]), and even more so after adjustment to GIPSE (OR = 0.82, 95% CI [0.72, 0.92]). However, adjustment to the components of the gendered indices rather than to the gendered indices themselves (models 4 and 5) did not attenuate the association between male sex and insomnia to the same level, perhaps because the gendered index captures sex-related effects better. Higher gendered indices, typical of male participants, were associated with lower insomnia likelihood (GISE OR = 0.90, 95% CI [0.68, 0.85], GIPSE OR = 0.62, 95% CI [0.58, 0.66]) in models adjusted for sex. In separate analyses within male and female strata, GISE and GIPSE were still associated with insomnia, with a similar effect size as in the combined analysis. Pseudo R^2 from these analyses is provided in Supplementary Table 3. In general, pseudo R^2 were higher in the female stratum compared to the male stratum, in all models. On the percentage scale, pseudo R^2 was 5.3% in the female stratum in a model with only covariates, and increased to 5.42% with GISE and 8.53% with GIPSE. It was substantially higher, with 15.27% when including all components of GIPSE. In the male stratum, pseudo R^2 ranged from 3.55% in the model with only covariates, to 12.64% in the model with all GIPSE components modeled individually.

Supplementary Table 4 reports results from secondary analysis using individuals with imputed data as well ($N = 16 415$ vs $N = 13 705$ using complete cases). The association of GISE with insomnia was attenuated, and the association was not statistically significant in male individuals. The associations of GIPSE remained about the same. Supplementary Tables 5 and 6 report results from similar analyses with the WHIIRS assessed on a continuous scale, rather than dichotomized to indicate insomnia. Supplementary Table 5 provides results from analyses that used the complete dataset, and Supplementary Table 6 from analyses that used the imputed data. The results are very similar to those from the main analysis using insomnia (dichotomized WHIIRS scores) as the outcome. One exception is that the results appear to be less attenuated between the imputed and complete

Table 3. Estimated Association of Demographic, Acculturation, and Psychological Measures With Sex and With Insomnia

Characteristic	Beta	95% CI	P-value	Beta	95% CI	P-value	Beta	95% CI	P-value
	Association with male sex			Association with insomnia (sex-unadjusted)			Association with insomnia (sex-adjusted)		
Demographic variables									
Marital status (ref = Single)									
Married/living with partner	0.91	(0.8,1.02)	.104	0.82	(0.72,0.94)	.004**	0.81	(0.71,0.93)	.003**
Divorced	0.45	(0.38,0.54)	<.001***	1.09	(0.92,1.28)	.326	0.99	(0.84,1.17)	.935
Income (ref = less than \$10 000)									
\$10 001–20 000	1.27	(1.08,1.5)	.003**	0.84	(0.72,0.99)	.038*	0.87	(0.74,1.02)	.087
\$20 001–40 000	1.59	(1.36,1.86)	<.001***	0.71	(0.59,0.84)	<.001***	0.74	(0.62,0.89)	.001**
\$40 001–75 000	1.99	(1.67,2.38)	<.001***	0.66	(0.53,0.82)	<.001***	0.71	(0.56,0.89)	.003**
More than \$75 000	3.06	(2.36,3.96)	<.001***	0.47	(0.32,0.69)	<.001***	0.53	(0.36,0.78)	.001**
Employment (ref = Retired/unemployed)									
Employed Part-Time	1.11	(0.96,1.28)	.156	0.78	(0.67,0.92)	.003**	0.79	(0.67,0.93)	.005**
Employed Full-Time	2.54	(2.24,2.87)	<.001***	0.63	(0.56,0.72)	<.001***	0.69	(0.61,0.79)	<.001***
Occupation (ref = Non-skilled worker)									
Service Worker	0.53	(0.45,0.63)	<.001***	1.17	(0.96,1.43)	.110	1.09	(0.89,1.33)	.417
Skilled Worker	1.33	(1.15,1.53)	<.001***	1.02	(0.87,1.2)	.774	1.06	(0.9,1.24)	.499
Professional/technical	0.49	(0.4,0.6)	<.001***	1.08	(0.9,1.31)	.412	0.99	(0.82,1.2)	.947
Other occupation	0.96	(0.82,1.12)	.594	1.10	(0.92,1.32)	.312	1.1	(0.91,1.31)	.324
Education (ref = No high school diploma or GED)									
At most High School Diploma/GED	1.09	(0.95,1.24)	.218	1.00	(0.85,1.18)	.987	1.01	(0.86,1.19)	.910
Greater than high school/GED	0.89	(0.78,1)	.053	0.92	(0.8,1.06)	.251	0.91	(0.78,1.05)	.192
Has current health insurance coverage	0.78	(0.71,0.86)	<.001***	0.89	(0.78,1.02)	.102	0.86	(0.75,0.99)	.033*
Acculturation measures									
Years in the US (Ref = less than 10 years)									
10 years or more	1.05	(0.93,1.19)	.458	1.11	(0.97,1.28)	.125	1.13	(0.98,1.3)	.094
US born	1.19	(1,1.42)	.056	1.45	(1.19,1.77)	<.001***	1.5	(1.23,1.82)	<.001***
Language preference: English	1.15	(0.99,1.33)	.070	1.28	(1.08,1.51)	.004**	1.3	(1.1,1.54)	.002**
SASH language acculturation subscale	0.07	(0.05,0.09)	<.001***	1.1	(1.03,1.18)	.004**	1.13	(1.06,1.21)	<.001***
SASH social acculturation subscale	0.03	(0.02,0.04)	<.001***	0.99	(0.88,1.11)	.885	1.02	(0.91,1.14)	.764
Ethnic identity score	0.00	(−0.01,0.01)	.793	0.97	(0.88,1.08)	.581	0.97	(0.88,1.08)	.610
Psychological measures									
CESD9	−0.10	(−0.11,−0.08)	<.001***	1.13	(1.12,1.14)	<.001***	1.13	(1.12,1.14)	<.001***
STAI10	−0.13	(−0.15,−0.11)	<.001***	1.13	(1.12,1.15)	<.001***	1.13	(1.12,1.14)	<.001***

The table provides the estimated associations of each considered variable with sex and insomnia from analyses adjusted for covariates: age, study center, and Hispanic/Latino background. The associations were estimated separately for each variable (multiple levels of the same variables were included in the same analysis). Associations with reported sex were estimated for a multinomial regression using the `svyVGAM` R package, to allow for inclusion of multiple outcome levels in the same analysis. Associations with insomnia were estimated for a logistic regression with insomnia as the outcome. Insomnia associations were further estimated in both reported sex-unadjusted and sex-adjusted analyses.

HCHS/SOL: Hispanic Community Health Study/Study of Latinos; SASH: Social Acculturation Scale for Hispanics; WHIIRS: Women Health Initiative Insomnia Rating Scale. STAI10: Spielberger Trait Anxiety Inventory scale based on 10 items. CESD9: Center for Epidemiological Studies Depression scale based on 9 items (excluding sleep quality-related question).

p-value significance: *** corresponds to $p < .001$, ** corresponds to $p < .01$, * corresponds to $p < .05$.

dataset when using the continuous WHIIRS scores (for example, in [Supplementary Table 6](#) the association of GISE with WHIIRS in male participants is statistically significant, while it was not statistically significant in the corresponding imputed data analysis that used insomnia as the outcome).

[Supplementary Tables 7](#) (complete data) and [8](#) (imputed data) provide results from interaction analysis of sex and the gendered indices in association with insomnia and WHIIRS scores. The

estimated interaction effects were mostly not statistically significant, with the exception of the interaction between the estimated association of GIPSE and male sex on WHIIRS (estimated effect of 0.32 in both complete and imputed data analysis). Note that the estimated effect is in the opposite direction to that of the male variables: male sex, and higher GIPSE (which, on average, tend to be higher in male, compared to female, participants) are associated with lower WHIIRS, while the interaction term, modeling the

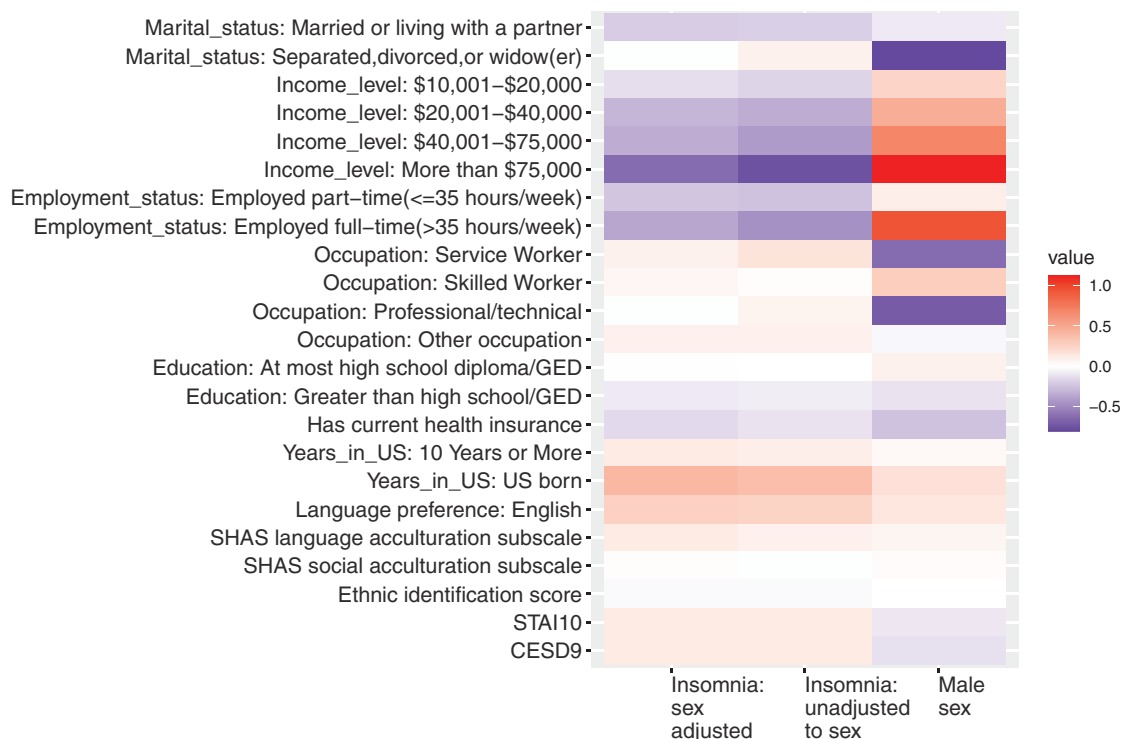


Figure 2. Visualization of the beta coefficients from the associations of sociodemographic and psychological variables with reported sex and insomnia. The figure visualizes the estimated associations (beta coefficients) from regression of each of the sociodemographic and psychological variables considered with insomnia as the outcome (sex-adjusted; left column), male sex as the exposure (middle column), and insomnia as the outcome in an analysis adjusted for sex (right column). Multi-level variables were used in the same regression model. All models were adjusted for covariates: age, study center, and Hispanic/Latino background. Opposite colors between sex and insomnia associations indicate that a variable is associated with higher (lower) likelihood of being male and lower (higher) likelihood of having insomnia. HCHS/SOL, Hispanic Community Health Study/Study of Latinos; SASH, Social Acculturation Scale for Hispanics; WHIIRS, Women Health Initiative Insomnia Rating Scale. STAI10, Spielberger Trait Anxiety Inventory scale based on 10 items. CESD9, Center for Epidemiological Studies Depression scale based on 9 items (excluding sleep quality-related question).

effect of increased GIPSE specifically in males, is associated with higher WHIIRS. This is probably due to some collinearity of the gendered indices and sex.

Discussion

We developed indices, GISE and GIPSE, summarizing non-biological sociodemographic, and psychological measures, such that the indices' distributions differ by sex. Our goal was to develop a measure that will aide in distinguishing sociodemographic contributions to sex differences in health outcomes from factors that could be attributed to biological sex effects in Hispanic/Latino population. The gendered indices are data-driven, and they do not measure gender, as there is no specific gender-related construct that they quantify. They are "gendered" in that they have different distributions across reported male and female individuals, with typically higher values in male participants (by construction). As an exemplar, we studied the association of the gendered indices with insomnia, a phenotype that has a higher prevalence in females compared to male individuals [29] with evidence relating these sex differences to both biological and sociocultural factors. Higher values of the indices were associated with a lower likelihood of insomnia, even when adjusting for reported sex, and in sex-stratified analysis. Using the GISE when modeling insomnia risk weakened somewhat the estimated association of reported sex with insomnia, suggesting that gendered sociodemographic factors explain some of the effects of reported sex on insomnia. Using GIPSE in the insomnia association model resulted in

a higher change in the estimated effect of reported sex, further suggesting that psychological measures, depression, and anxiety, have a substantial role in sex differences in insomnia.

The GISE aggregates marital status, household income, employment status, occupation category, education level, health insurance status, and cultural-acculturation measures that are unique to immigrant populations, including years in the United States, language preference, etc. The variables that had the highest weight in GISE, suggesting that they demonstrate strong sex differences, by order, are having an office-based occupation (more likely in the female stratum) being employed full-time (more likely in the male stratum), and being a service worker (more likely in the female stratum). Having full-time employment was strongly, and negatively associated with odds of insomnia (adjusted OR = 0.69, 95% CI [0.61,0.79]), while occupation variables did not have statistically significant associations with insomnia individually. Still, the association of GISE with insomnia and WHIIRS scores was due to the associations of its aggregated components with both sex and insomnia: a variable, for example, being employed full time, is both more common in males (say) and is associated with a lower likelihood of insomnia, or the converse. GISE variables are strictly sociodemographic, suggesting that the sociocultural and demographic trends affecting choices (or "outcomes," as sometimes choices are limited) of male and female individuals have health implications, such as insomnia and, perhaps, its comorbid or incident health outcomes. In other words, these results add to existing evidence that the gendered environment contributes to sex disparities in health outcomes.

Table 4. Gendered Indices Variables and Coefficients

	Coefficient: GISE	Coefficient: GIPSE (includes psychological variables)
Demographic variables		
Marital status		
Married or living with a partner	0.03	0
Separated, divorced, or widow(er)	-0.64	-0.62
Income		
\$10 001–20 000	0.05	0
\$20 001–40 000	0.13	0.05
\$40 001–75 000	0.32	0.19
More than \$75 000	0.6	0.44
Employment		
Employed part-time	-0.09	-0.15
Employed full-time	0.66	0.59
Occupation		
Service worker	-0.62	-0.62
Skilled worker	0.16	0.17
Professional/technical	-0.72	-0.73
Other occupation	0.06	0.06
Education		
At most high school diploma/GED	0.02	0
Greater than high school/GED	-0.21	-0.28
Has current health insurance coverage		
-0.33		-0.32
Acculturation measures		
Language preference (English)	0.3	0.31
SASH language acculturation subscale	0.07	0.06
SASH social acculturation subscale	0.15	0.14
Ethnic identity score	0.3	0.31
Years live in the US (ref = Less than 10 Years)		
Years in US: 10 Years or More	-0.08	-0.07
Years in US: US born	-0.29	-0.28
Psychological measures		
STAI10	NA	-0.04
CESD9	NA	-0.03

Coefficients of the various variables were estimated in a LASSO penalized logistic regression of male sex over all variables jointly. Adjusted covariates were not used in this analysis. Thus, positive values for a binary variable suggest that it is more likely to be seen in male individuals, and for a continuous variable, it suggests that male individuals tend to have higher values.

GISE: gendered index of the sociodemographic environment; GIPSE: gendered index of the psychological and sociodemographic environment; HCHS/SOL: Hispanic Community Health Study/Study of Latinos; SASH: Social Acculturation Scale for Hispanics; WHIIRS: Women Health Initiative Insomnia Rating Scale. STAI10: Spielberger Trait Anxiety Inventory scale based on 10 items. CESD9: Center for Epidemiological Studies Depression scale based on 9 items (excluding sleep quality-related question).

Clinically, these findings reinforce the importance of accounting for social determinants of health in the primary care setting, and identifying individuals that may be at higher risk based on marital and employment status, etc. Notably, GIPSE has a different interpretation, as it includes anxiety and depression measures, which may have bidirectional relationships with insomnia (where insomnia and anxiety may increase the risk of each other, for instance).

The gendered indices are continuous rather than categorically male or female, representing a weighted aggregate of sociodemographic factors resulting in a range of values. This agrees with reality, as both biological sex and gender effects are continuous.

While binary sex is usually used in research, there are individuals who fall outside the standard definition based on sex chromosomes [44]. Furthermore, sex hormone levels are continuous and change throughout the reproductive cycle of women [45], throughout the life course in both men and women, and precipitously at menopause in women [46]. The concept of gender is also suggested to be continuous, with measures of gender identity [47, 48] and other gendered expressions and norms [49] being potentially measured by scales. Thus, binary representations of both sex and gender are limited and may not capture nuances in how sex and gender, as encompassing both biological and cultural determinants, relate to health.

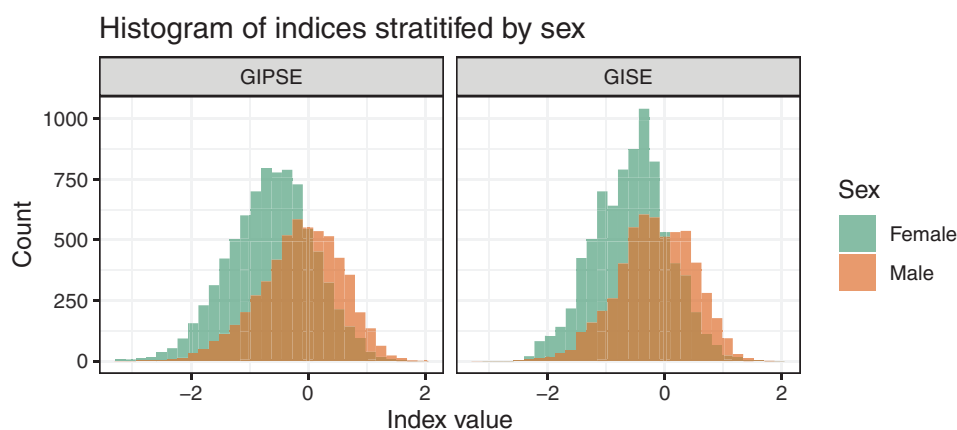


Figure 3. Distribution of the gendered indices in female and male participants. Distribution of the primary and secondary gendered indices by strata of reported sex. GISE, gendered index of sociodemographic environment; GIPSE, gendered index of psychological and sociodemographic environment.

Table 5. Association of Reported Sex and Standardized Gendered Indices With Insomnia

Model	Estimated male sex association			Estimated gendered index association		
	Estimated OR	95% CI	P-value	Estimated OR	95% CI	P-value
Model 1: adjusting for covariates.	0.60	(0.53,0.67)	<.001***	—	—	—
Model 2: adjusting for covariates, GISE	0.63	(0.56,0.7)	<.001***	0.92	(0.87,0.99)	.017*
Model 3: adjusting for covariates, GIPSE	0.78	(0.69,0.88)	<.001***	0.65	(0.61,0.7)	<.001***
Model 4: adjusting for covariates, components of GISE	0.62	(0.55,0.7)	<.001***	—	—	—
Model 5: adjusting for covariates, components of GIPSE	0.73	(0.64,0.84)	<.001***	—	—	—
Analysis in male stratum						
Model adjusting for covariates, GISE	—	—	—	.93	(0.85,1.03)	.158
Model adjusting for covariates, GIPSE	—	—	—	.68	(0.62,0.75)	<.001***
Analysis in female stratum						
Model adjusting for covariates, GISE	—	—	—	.92	(0.84,0.99)	.034*
Model adjusting for covariates, GIPSE	—	—	—	.63	(0.58,0.69)	<.001***

The estimated association of male sex and of the gendered indices with insomnia in various regression models. Association effects are adjusted odds ratios (OR). Thus, values lower than 1 indicate a protective association. For gendered indices, ORs are per 1 standard deviation (SD) increase in the index. Male individuals tend to have higher values of the gendered indices. Adjusting covariates were age, Hispanic/Latino background, and study center. GISE: gendered index of sociodemographic environment; GIPSE: gendered index of psychological and sociodemographic environment.

Consideration of socially constructed gendered roles and factors in relation to health and sleep is especially warranted in under-studied, diverse populations such as US-residing Hispanic/Latino individuals whose social experiences may differ in meaningful ways from populations already well-described by current literature, e.g. due to a multitude of structural and cultural components unique to this population ([50–52], and whose distribution of risk for sleep disorders may not be equivalent to that of other population groups [53]. For example, Kaufmann et al. ([54] reported that the severity of insomnia increases with age in Hispanic adults but decreases with age in non-Hispanic white adults, strongly suggesting that assumptions regarding insomnia in Hispanic/Latino populations should not be inferred from non-Hispanic White populations. While the constructed gendered indices do not quantify any previously specified measure of gender, we plan to use them to estimate the potential contribution of gendered sociodemographic-related costs with sleep health and other health outcomes in HCHS/SOL. As ongoing and future data collection efforts incorporate better measures of sex and gender, it will allow for more nuanced inference of the role of sex-related effects on health.

While the gendered indices replicated well in an independent subset of HCHS/SOL adults, these indices are not likely to be generalizable to other race or ethnic populations due to the cultural differences between population segments, as well as due to differences in surveys applied in different studies. The specificity of the gendered indices to HCHS/SOL (a strength) suggests that new indices will need to be tailored for other studies. Notably, we did not study potential differences in the distribution of sociodemographic measures used or in the potential resulting gendered indices between Hispanic background groups (Mexican, Cuban, Puerto Rican, etc.), which are both socioculturally diverse and exhibit different distributions of health outcomes in the HCHS/SOL [55–57]. While this is certainly feasible, the indices will likely be less accurate as they will suffer from overfitting due to small sample sizes.

It is interesting to note that higher acculturation to the United States was associated with a higher likelihood of insomnia and worse insomnia symptoms (WHIIRS), compared to less acculturation. For instance, being born in the United States had the strongest association with higher insomnia likelihood compared to other variables, while also living 10 or more years in the United States

was associated with higher insomnia likelihood compared to living less than 10 years in the United States. Similarly, preferring the English language and having higher language acculturation scores were associated with a higher likelihood of insomnia. These findings are in line with those previously reported in an analysis of sleep duration patterns in US Latinos from the National Health Interview Survey [58]. The authors found that US-born Latinos, disaggregated by Hispanic/Latino background, had worse sleep patterns (too short or too long sleep) compared to foreign-born Latinos from the same background. García et al. hypothesized that this trend is the result of the adoption of unhealthy lifestyles (diet, smoking, etc., as supported by other studies [59]), which in turn results in worse sleep quality.

Strengths of this study include the use of a large cohort of diverse Hispanic/Latino adults, representing multiple Hispanic/Latino backgrounds, in the United States, a rigorous statistical analysis that included secondary analysis validating the scores by independent training-testing data split, and imputed data analysis. There are also some limitations. The baseline HCHS/SOL dataset only included binary sex definition and did not assess gender identity. Such data will become available in the future, based on HCHS/SOL Visit 3. The WHIIRS measure used to assess insomnia was developed and validated in a research study of postmenopausal women. While it has been used in multiple studies of other populations (such as HCHS/SOL), it is a limitation that the WHIIRS was not validated in more age, sex, and gender-diverse populations. Our gendered indices used household income, not the income of the individuals themselves. As household income was associated with reported sex, it is likely that either household income reporting depends on reported sex, or those household compositions in HCHS/SOL differ by reported sex. Notably, the idea of gendered indices was inspired by Smith et al. [60], who developed a gendered index in the context of labor force participation. They used a few areas of demographic measures that differ between men and women (occupation segregation, work hours, level of education, responsibility of caring for children), and included differences in income between household male and female participants in their index. By that, they interpret their index as measuring gender roles in the labor force. The indices developed here are not interpreted in such a specific context, as they have both sociodemographic (GISE) and psychological (GIPSE) components. Generally, data collection in HCHS/SOL was not specific to study sex effects, and we do not have data about additional factors that are known to have strong gendered patterns such as caregiving roles. Such measures could potentially improve the indices but do not negate the value of the indices that have been constructed. In the context of household income, it is possible that it contributes to the gendered indices via a relationship with occupation status and type, which also differ by sex, but it is difficult to pinpoint anything with certainty given the limitations of the available data. Regardless, it is unlikely that the gendered indices association results are driven by household income, as it is just one of multiple variables associated with both sex and insomnia.

To summarize, we constructed two gendered indices, each summarizing sociodemographic factors, one also including psychological measures, into a single continuous variable that on average tends to have higher values in reported male individuals. The indices explain some of the association of sex with insomnia. In future work, we will continue using these gendered indices to identify potentially behaviorally modifiable aspects of sleep health and other health outcomes.

Supplementary Material

Supplementary material is available at *SLEEP Advances* online.

Funding

The Hispanic Community Health Study/Study of Latinos is a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (HHSN268201300001I / N01-HC-65233), University of Miami (HHSN268201300004I / N01-HC-65234), Albert Einstein College of Medicine (HHSN268201300002I / N01-HC-65235), University of Illinois at Chicago (HHSN268201300003I / N01-HC-65236 Northwestern Univ), and San Diego State University (HHSN268201300005I / N01-HC-65237). The following Institutes/Centers/Offices have contributed to the HCHS/SOL through a transfer of funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements. This research was partially supported by NHLBI grants R01HL161012 (to TS), R35HL135818 (to SR), and by a microgrant from the Brigham Research Institute (to TS). Additional support was provided by the Life Course Methodology Core (LCMC) at the New York Regional Center for Diabetes Translation Research (P30 DK111022).

Disclosure Statement

Financial disclosure: Redline discloses consulting relationships with Eli Lilly Inc.

Nonfinancial disclosure: Redline serves as an unpaid member of the Apnimed Scientific Advisory Board, as an unpaid board member for the Alliance for Sleep Apnea Partners, and has received loaned equipment for a multi-site study: oxygen concentrators from Philips Respironics and polysomnography equipment from Nox Medical. A previous version of this manuscript has been deposited as a preprint on medRxiv.

Author Contributions

Natali Sorajja (Formal analysis [lead], Writing—original draft [equal]), Joon Chung (Conceptualization [equal], Writing—original draft [equal]), Carmela Alcántara (Writing—review & editing [supporting]), Sylvia Wassertheil-Smoller (Writing—review & editing [supporting]), Frank Penedo (Writing—review & editing [supporting]), Alberto R Ramos (Writing—review & editing [supporting]), Krista Perreira (Writing—review & editing [equal]), Martha Daviglius (Writing—review & editing [supporting]), Shakira Suglia (Writing—review & editing [supporting]), Linda Gallo (Writing—review & editing [equal]), Peter Liu (Writing—review & editing [equal]), Susan Redline (Funding acquisition [equal], Writing—review & editing [equal]), Carmen Isasi (Writing—review & editing [equal]), and Tamar Sofer (Conceptualization [lead], Formal analysis [equal], Funding acquisition [equal], Methodology [lead], Supervision [lead], Writing—original draft [equal])

Data Availability

HCHS/SOL data are available via a data use agreement with the HCHS/SOL Data Coordinating Center. See <https://sites.csc.unc.edu/hchs/> for study procedures. HCHS/SOL data are

also available on the National Heart Lung and Blood Institute's BioLINCC (Biologic Specimen and Data Repository Information Coordinating Center) repository under accession number HLB01141422a.

Code Availability

Code used in this analysis will become publicly available upon paper acceptance on the GitHub repository https://github.com/tamartsi/Gendered_indices_and_insomnia.

Ethics Statement

The HCHS/SOL was approved by the institutional review boards (IRBs) at each field center, where all participants gave written informed consent, and by the Non-Biomedical IRB at the University of North Carolina at Chapel Hill, to the HCHS/SOL Data Coordinating Center. All IRBs approving the HCHS/SOL study are: Non-Biomedical IRB at the University of North Carolina at Chapel Hill, Chapel Hill, NC; Einstein IRB at the Albert Einstein College of Medicine of Yeshiva University, Bronx, NY; IRB at Office for the Protection of Research Subjects (OPRS), University of Illinois at Chicago, Chicago, IL; Human Subject Research Office, University of Miami, Miami, FL; Institutional Review Board of San Diego State University, San Diego, CA. All methods and analyses of HCHS/ SOL participants' materials and data were carried out in accordance with human subject research guidelines and regulations. This work was approved by the Mass General Brigham IRB and by the Beth Israel Deaconess Medical Center Committee on Clinical Investigations.

References

- Oksuzyan A, Juel K, Vaupel JW, Christensen K. Men: good health and high mortality. Sex differences in health and aging. *Aging Clin Exp Res*. 2008;**20**(2):91–102. doi: [10.1007/BF03324754](https://doi.org/10.1007/BF03324754)
- Tronieri JS, Wurst CM, Pearl RL, Allison KC. Sex differences in obesity and mental health. *Curr Psychiatry Rep*. 2017;**19**(6):29. doi: [10.1007/s11920-017-0784-8](https://doi.org/10.1007/s11920-017-0784-8)
- Merz AA, Cheng S. Sex differences in cardiovascular ageing. *Heart*. 2016;**102**(11):825–831. doi: [10.1136/heartjnl-2015-308769](https://doi.org/10.1136/heartjnl-2015-308769)
- McCarrey AC, An Y, Kitner-Triolo MH, Ferrucci L, Resnick SM. Sex differences in cognitive trajectories in clinically normal older adults. *Psychol Aging*. 2016;**31**(2):166–175. doi: [10.1037/pag0000070](https://doi.org/10.1037/pag0000070)
- Gilks WP, Abbott JK, Morrow EH. Sex differences in disease genetics: evidence, evolution, and detection. *Trends Genet*. 2014;**30**(10):453–463. doi: [10.1016/j.tig.2014.08.006](https://doi.org/10.1016/j.tig.2014.08.006)
- Kassam I, Wu Y, Yang J, Visscher PM, McRae AF. Tissue-specific sex differences in human gene expression. *Hum Mol Genet*. 2019;**28**(17):2976–2986. doi: [10.1093/hmg/ddz090](https://doi.org/10.1093/hmg/ddz090)
- Guo L, Zhong MB, Zhang L, Zhang B, Cai D. Sex differences in alzheimer's disease: insights from the multiomics landscape. *Biol Psychiatry*. 2022;**91**(11):61–71. doi: [10.1016/j.biopsych.2021.02.968](https://doi.org/10.1016/j.biopsych.2021.02.968)
- Barr E, Popkin R, Roodzant E, Jaworski B, Temkin SM. Gender as a social and structural variable: research perspectives from the National Institutes of Health (NIH). *Transl. Behav. Med.*. 2024;**14**(1):13–22. doi: [10.1093/tbm/ibad014](https://doi.org/10.1093/tbm/ibad014)
- Samango-Sprouse C, Kirkizlar E, Hall MP, et al. Incidence of X and Y chromosomal aneuploidy in a large child bearing population. *PLoS One*. 2016;**11**(8):e0161045. doi: [10.1371/journal.pone.0161045](https://doi.org/10.1371/journal.pone.0161045)
- Konishi A, Samura O, Muromoto J, et al. Prevalence of common aneuploidy in twin pregnancies. *J Hum Genet*. 2022;**67**:261–265. doi: [10.1038/s10038-021-01001-0](https://doi.org/10.1038/s10038-021-01001-0)
- Committee on Practice Bulletins—Obstetrics, Committee on Genetics, and Society for Maternal-Fetal Medicine. Practice bulletin no. 226: screening for fetal chromosomal abnormalities. *Obstet Gynecol*. 2020;**136**(4):e48–e69.
- Doyal L. Sex, gender, and health: the need for a new approach. *BMJ*. 2001;**323**(7320):1061–1063. doi: [10.1136/bmj.323.7320.1061](https://doi.org/10.1136/bmj.323.7320.1061)
- Connelly PJ, Azizi Z, Alipour P, Delles C, Pilote L, Raparelli V. The importance of gender to understand sex differences in cardiovascular disease. *Can J Cardiol*. 2021;**37**(5):699–710. doi: [10.1016/j.cjca.2021.02.005](https://doi.org/10.1016/j.cjca.2021.02.005)
- Torgirimson BN, Minson CT. Sex and gender: what is the difference? *J Appl Physiol*. 2005;**99**(3):787785–787787. doi: [10.1152/jappphysiol.00376.2005](https://doi.org/10.1152/jappphysiol.00376.2005)
- Gauci S, Cartledge S, Redfern J, et al. Biology, bias, or both? the contribution of sex and gender to the disparity in cardiovascular outcomes between women and men. *Curr Atheroscler Rep*. 2022;**24**(9):701–708. doi: [10.1007/s11883-022-01046-2](https://doi.org/10.1007/s11883-022-01046-2)
- Baker FC, Yüksel D, de Zambotti M. Sex differences in sleep. In: Attarian H, Viola-Saltzman M, editors. *Sleep disorders in women: A guide to practical management*. Cham: Springer International Publishing; 2020. p. 55–64.
- Krishnan V, Collop NA. Gender differences in sleep disorders. *Curr Opin Pulm Med*. 2006;**12**(6):383–389. doi: [10.1097/01.mcp.0000245705.69440.6a](https://doi.org/10.1097/01.mcp.0000245705.69440.6a)
- Mong JA, Cusmano DM. Sex differences in sleep: impact of biological sex and sex steroids. *Philos Trans R Soc Lond B Biol Sci*. 2016;**371**(1688):20150110. doi: [10.1098/rstb.2015.0110](https://doi.org/10.1098/rstb.2015.0110)
- Decker AN, Fischer AR, Gunn HE. Socio-Ecological Context of Sleep: Gender Differences and Couples' Relationships as Exemplars. *Curr Psychiatry Rep*. 2022;**24**(12):831–840. doi: [10.1007/s11920-022-01393-6](https://doi.org/10.1007/s11920-022-01393-6)
- Arber S, Hislop J, Bote M, Meadows R. Gender roles and women's sleep in mid and later life: a quantitative approach. *Sociol Res Online*. 2007;**12**(5):182–199. doi: [10.5153/sro.1609](https://doi.org/10.5153/sro.1609)
- Jackson CL, Hu FB, Redline S, Williams DR, Mattei J, Kawachi I. Racial/ethnic disparities in short sleep duration by occupation: the contribution of immigrant status. *Soc Sci Med*. 2014;**118**:71–79. doi: [10.1016/j.socscimed.2014.07.059](https://doi.org/10.1016/j.socscimed.2014.07.059)
- Jarrin DC, McGrath JJ, Silverstein JE, Drake C. Objective and subjective socioeconomic gradients exist for sleep quality, sleep latency, sleep duration, weekend oversleep, and daytime sleepiness in adults. *Behav Sleep Med*. 2013;**11**(2):144–158. doi: [10.1080/15402002.2011.636112](https://doi.org/10.1080/15402002.2011.636112)
- Czeisler CA, Klerman EB. Circadian and sleep-dependent regulation of hormone release in humans. *Recent Prog Horm Res*. 1999;**54**:97–130; discussion 130.
- St-Onge M-P, Pizinger T, Kovtun K, RoyChoudhury A. Sleep and meal timing influence food intake and its hormonal regulation in healthy adults with overweight/obesity. *Eur J Clin Nutr*. 2019;**72**(suppl 1):76–82. doi: [10.1038/s41430-018-0312-x](https://doi.org/10.1038/s41430-018-0312-x)
- Kelly MR, Yuen F, Satterfield BC, et al. Endogenous diurnal patterns of adrenal and gonadal hormones during a 24-hour constant routine after simulated shift work. *J Endocr Soc*. 2022;**6**(12):bvac153. doi: [10.1210/jendso/bvac153](https://doi.org/10.1210/jendso/bvac153)
- Stewart EM, Landry S, Edwards BA, Drummond SPA. The bidirectional relationship between sleep and health. In: Paul RH, Salminen LE, Heaps J, Cohen LM, eds. *The wiley encyclopedia of health psychology*. Hoboken, NJ: Wiley; 2020. p. 165–188.

27. Grandner MA. Sleep, health, and society. *Sleep Med Clin*. 2017;**12**(1):1–22. doi: [10.1016/j.jsmc.2016.10.012](https://doi.org/10.1016/j.jsmc.2016.10.012)
28. Suh S, Cho N, Zhang J. Sex Differences in Insomnia: from Epidemiology and Etiology to Intervention. *Curr Psychiatry Rep*. 2018;**20**(99):69. doi: [10.1007/s11920-018-0940-9](https://doi.org/10.1007/s11920-018-0940-9)
29. Zhang B, Wing Y-K. Sex differences in insomnia: a meta-analysis. *Sleep*. 2006;**29**(1):85–93. doi: [10.1093/sleep/29.1.85](https://doi.org/10.1093/sleep/29.1.85)
30. Thomson MD, Hoffman-Goetz L. Defining and measuring acculturation: a systematic review of public health studies with Hispanic populations in the United States. *Soc Sci Med*. 2009;**69**(7):983–991. doi: [10.1016/j.socscimed.2009.05.011](https://doi.org/10.1016/j.socscimed.2009.05.011)
31. Lavange LM, Kalsbeek WD, Sorlie PD, et al. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol*. 2010;**20**(8):642–649. doi: [10.1016/j.annepidem.2010.05.006](https://doi.org/10.1016/j.annepidem.2010.05.006)
32. Sorlie PD, Avilés-Santa LM, Wassertheil-Smoller S, et al. Design and implementation of the Hispanic Community Health Study/Study of Latinos. *Ann Epidemiol*. 2010;**20**(8):629–641. doi: [10.1016/j.annepidem.2010.03.015](https://doi.org/10.1016/j.annepidem.2010.03.015)
33. Pirzada A, Cai J, Cordero C, et al. Risk factors for cardiovascular disease: knowledge gained from the Hispanic Community Health Study/Study of Latinos. *Curr Atheroscler Rep*. 2023;**25**(11):785–793. doi: [10.1007/s11883-023-01152-9](https://doi.org/10.1007/s11883-023-01152-9)
34. Pirzada A, Cai J, Heiss G, et al. Evolving science on cardiovascular disease among hispanic/latino adults: JACC international. *J Am Coll Cardiol*. 2023;**81**(15):1505–1520. doi: [10.1016/j.jacc.2023.02.023](https://doi.org/10.1016/j.jacc.2023.02.023)
35. Poteat T, Gallo LC, Harkness A, et al. Influence of stress, gender, and minority status on cardiovascular disease risk in the hispanic/latino community: protocol for a longitudinal observational cohort study. *JMIR Res Protoc*. 2021;**10**(5):e28997. doi: [10.2196/28997](https://doi.org/10.2196/28997)
36. Levine DW, Kripke DF, Kaplan RM, et al. Reliability and validity of the Women's Health Initiative Insomnia Rating Scale. *Psychol Assess*. 2003;**15**(2):137–148. doi: [10.1037/1040-3590.15.2.137](https://doi.org/10.1037/1040-3590.15.2.137)
37. The Women's Health Initiative Study Group. Design of the women's health initiative clinical trial and observational study. *Control Clin Trials*. 1998;**19**(1):61–109.
38. Levine DW, Dailey ME, Rockhill B, Tipping D, Naughton MJ, Shumaker SA. Validation of the Women's Health Initiative Insomnia Rating Scale in a multicenter controlled clinical trial. *Psychosom Med*. 2005;**67**(1):98–104. doi: [10.1097/01.psy.0000151743.58067.f0](https://doi.org/10.1097/01.psy.0000151743.58067.f0)
39. Marin G, Sabogal F, Marin BV, Otero-Sabogal R, Perez-Stable EJ. Development of a Short Acculturation Scale for Hispanics. *Hisp J Behav Sci*. 1987;**9**(2):183–205.
40. Spielberger CD. *State-Trait Anxiety Inventory for Adults*. PsycTests Database Record: American Psychological Association (APA); 2012.
41. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psychol Measurement*. 1977;**1**(3):385–401. doi: [10.1177/014662167700100306](https://doi.org/10.1177/014662167700100306)
42. Lumley T. *Complex surveys: a guide to analysis using R*. Hoboken, NJ: Wiley & Sons, Inc., 1st edition; 2010.
43. Lumley T. Pseudo-R² statistics under complex sampling. *Aust N Z J Stat*. 2017;**59**(2):187–194.
44. Štrkalj G, Pather N. Beyond the sex binary: toward the inclusive anatomical sciences education. *Anat Sci Educ*. 2021;**14**(4):513–518. doi: [10.1002/ase.2002](https://doi.org/10.1002/ase.2002)
45. Lovejoy JC. The influence of sex hormones on obesity across the female life span. *J Womens Health*. 1998;**7**(10):1247–1256. doi: [10.1089/jwh.1998.7.1247](https://doi.org/10.1089/jwh.1998.7.1247)
46. Gurvich C, Hoy K, Thomas N, Kulkarni J. Sex differences and the influence of sex hormones on cognition through adulthood and the aging process. *Brain Sci*. 2018;**8**(9):163. doi: [10.3390/brainsci8090163](https://doi.org/10.3390/brainsci8090163)
47. Brenøe AA, Heursen L, Ranehill E, Weber RA. Continuous gender identity and economics. *AEA Papers Proc*. 2022;**112**:573–577. doi: [10.1257/pandp.20221083](https://doi.org/10.1257/pandp.20221083)
48. Lindqvist A, Sendén MG, Renström EA. What is gender, anyway: a review of the options for operationalising gender. *Psychol Sexual*. 2020;**12**:332–344. doi: [10.1080/19419899.2020.1729844](https://doi.org/10.1080/19419899.2020.1729844)
49. Smiler AP, Epstein M. Measuring gender: options and issues. In: Chrisler JC, McCreary DR, eds. *Handbook of gender research in psychology*. NY, NY: Springer New York; 2010. p. 133–157.
50. Molina KM, Estrella ML, Durazo-Arvizu R, et al. Perceived discrimination and physical health-related quality of life: The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) Sociocultural Ancillary Study. *Soc Sci Med*. 2019;**222**:91–100. doi: [10.1016/j.socscimed.2018.12.038](https://doi.org/10.1016/j.socscimed.2018.12.038)
51. Nuñez A, González P, Talavera GA, et al. Machismo, Marianismo, and Negative Cognitive-Emotional Factors: Findings From the Hispanic Community Health Study/Study of Latinos Sociocultural Ancillary Study. *J Lat Psychol*. 2016;**4**(4):202–217. doi: [10.1037/lat0000050](https://doi.org/10.1037/lat0000050)
52. Arellano-Morales L, Roesch SC, Gallo LC, et al. Prevalence and correlates of perceived ethnic discrimination in the hispanic community health study/study of latinos sociocultural ancillary study. *J Lat Psychol*. 2015;**3**(3):160–176. doi: [10.1037/lat0000040](https://doi.org/10.1037/lat0000040)
53. Redline S, Sotres-Alvarez D, Loredó J, et al. Sleep-disordered breathing in Hispanic/Latino individuals of diverse backgrounds. The Hispanic Community Health Study/Study of Latinos. *Am J Respir Crit Care Med*. 2014;**189**(3):335–344. doi: [10.1164/rccm.201309-1735OC](https://doi.org/10.1164/rccm.201309-1735OC)
54. Kaufmann CN, Mojtabei R, Hock RS, et al. Racial/ethnic differences in insomnia trajectories among U.S. older adults. *Am J Geriatr Psychiatry*. 2016;**24**(7):575–584. doi: [10.1016/j.jagp.2016.02.049](https://doi.org/10.1016/j.jagp.2016.02.049)
55. Daviglus ML, Talavera GA, Avilés-Santa ML, et al. Prevalence of major cardiovascular risk factors and cardiovascular diseases among Hispanic/Latino individuals of diverse backgrounds in the United States. *JAMA*. 2012;**308**(17):1775–1784. doi: [10.1001/jama.2012.14517](https://doi.org/10.1001/jama.2012.14517)
56. Allison MA, Gonzalez F, Raji L, et al. Cuban Americans have the highest rates of peripheral arterial disease in diverse Hispanic/Latino communities. *J Vasc Surg*. 2015;**62**(3):665–672. doi: [10.1016/j.jvs.2015.03.065](https://doi.org/10.1016/j.jvs.2015.03.065)
57. Heiss G, Snyder ML, Teng Y, et al. Prevalence of metabolic syndrome among Hispanics/Latinos of diverse background: the Hispanic Community Health Study/Study of Latinos. *Diabetes Care*. 2014;**37**(8):231–299.
58. García C, Sheehan CM, Flores-Gonzalez N, Ailshire JA. Sleep Patterns among US Latinos by nativity and country of origin: results from the National Health Interview Survey. *Ethn Dis*. 2020;**30**(1):119–128. doi: [10.18865/ed.30.1.119](https://doi.org/10.18865/ed.30.1.119)
59. Abraído-Lanza AF, Chao MT, Flórez KR. Do healthy behaviors decline with greater acculturation? Implications for the Latino mortality paradox. *Soc Sci Med*. 2005;**61**(6):1243–1255. doi: [10.1016/j.socscimed.2005.01.016](https://doi.org/10.1016/j.socscimed.2005.01.016)
60. Smith PM, Koehoorn M. Measuring gender when you don't have a gender measure: constructing a gender index using survey data. *Int J Equity Health*. 2016;**15**:82. doi: [10.1186/s12939-016-0370-4](https://doi.org/10.1186/s12939-016-0370-4)