



# Environment-wide association study to comprehensively test and validate associations between nutrition and lifestyle factors and testosterone deficiency: NHANES 1988–1994 and 1999–2004

D. S. Lopez<sup>1</sup> · W. Wulaningsih<sup>2,3</sup> · K. K. Tsilidis<sup>4,5</sup> · J. Baillargeon<sup>1</sup> · S. B. Williams<sup>6</sup> · R. Urban<sup>7</sup> · S. Rohrmann<sup>8</sup>

Received: 24 July 2019 / Accepted: 3 February 2020  
© Hellenic Endocrine Society 2020

## Abstract

**Purpose** Testosterone (T) plays an important role in men's health and its deficiency is linked with poorer health. However, the role of nutritional and lifestyle factors in T regulation and production remains unclear. The objectives are to comprehensively test the cross-sectional associations of nutritional and lifestyle factors with T deficiency and to validate the associations in the NHANES survey.

**Methods** We performed weighted multivariable logistic regression analysis to examine the association of 173 nutritional and lifestyle factors with T deficiency (total testosterone  $\leq 3.5$  ng/mL) in NHANES III as the discovery set (mean age 41). We controlled for multiple comparisons with a false discovery rate (FDR)  $< 5\%$  and replicated in NHANES 1999–2004 (mean age 44).

**Results** We identified seven nutritional factors as being inversely associated with T deficiency in NHANES 1999–2004, namely dietary intake of vitamin A, protein, saturated fatty acids, monounsaturated fatty acids, total fats, saturated fatty acid 16:0, and phosphorus. In a multivariable model, only vitamin A intake remained significantly associated with T deficiency (OR 0.97, 95% CI 0.94–0.99). Principal component analysis suggested that the two principal components, (1) dietary fats, protein, and phosphorus and (2) total vitamin A, may be associated with T deficiency.

**Conclusion** Our systematic evaluation provided new insight into the modifiable factors that could play a role in the regulation of T production. This study has the potential to contribute to the current body of literature which seeks to formulate a clinical definition of T deficiency after taking into account nutritional and lifestyle factors.

**Keywords** Testosterone: nutrition · EWAS · NHANES

## Introduction

While testosterone (T) is linked with many positive physiological effects, including improved sexual function, physical

performance, muscle strength, lean body mass, and cognitive function, low levels are associated with an increased risk of cardiovascular mortality and comorbidities [1–3]. Specifically, low levels of T are associated with cardiovascular

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s42000-020-00179-w>) contains supplementary material, which is available to authorized users.

✉ D. S. Lopez  
davlopez@utmb.edu

<sup>1</sup> Department of Preventive Medicine and Community Health, University of Texas Medical Branch, 301 University Blvd, Galveston, TX, USA

<sup>2</sup> Cancer Epidemiology Group, Division of Cancer Studies, King's College London, London, UK

<sup>3</sup> MRC Unit for Lifelong Health and Ageing, University College London, London, UK

<sup>4</sup> Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece

<sup>5</sup> Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK

<sup>6</sup> Department of Surgery, Division of Urology, University of Texas Medical Branch, Galveston, TX, USA

<sup>7</sup> Department of Internal Medicine, University of Texas Medical Branch, Galveston, TX, USA

<sup>8</sup> Division of Chronic Disease Epidemiology; Epidemiology, Biostatistics and Preventive Institute (EBPI), University of Zurich, Zurich, Switzerland

diseases [4], cancer [5], and all-cause mortality [6]. In the USA alone, studies have reported that as many as 38.7% of men 45 years or older demonstrate low levels of T [7, 8] or T deficiency (total testosterone  $\leq 350$  ng/dL) [9], it being projected that by 2025 approximately 6.5 million men will have T deficiency [8, 10]. The obesity epidemic and the aging of the US male population have contributed, in part, to the increasing prevalence of T deficiency [11]. However, the role of other nutritional and environmental factors in the regulation of T production has not been widely investigated.

Recently, Patel et al. proposed a study design analogous to genome-wide association studies (GWAS), namely “the environment-wide association study (EWAS)” to comprehensively test and analytically validate nutritional and environmental factors associated with chronic diseases and other clinical phenotypes [12]. Similarly to GWAS, EWAS does not test only one or a few associations at a time. Rather, it simultaneously evaluates multiple environmental factors for associations, with proper adjustment for multiple comparisons. Subsequently, the identified significant associations are further validated across independent datasets [12]. Therefore, in this study, we propose to utilize an EWAS approach to comprehensively test nutritional and lifestyle factors in relation to T deficiency in the NHANES III (1988–1994) discovery set and further replicate significant associations in the continuous NHANES 1999–2004.

## Methods

### Study population

The National Center for Health Statistics (NCHS) conducted the Third National Health and Nutrition Examination Survey (NHANES III), a cross-sectional study of the US civilian non-institutionalized population aged 2 months and older, between 1988 and 1994. NHANES III used a multistage, stratified, and clustered probability sampling in which Mexican-Americans, non-Hispanic blacks, the elderly, and young children were oversampled to ensure adequate sample sizes [13]. NHANES III was conducted in two phases (1988–1991 and 1991–1994) from which independent, unbiased national estimates of health and nutrition can be calculated. Participants were interviewed at home and asked detailed demographic and health-related questions. Participants also underwent extensive physical and laboratory examinations in a mobile examination center or during a home visit. Trained personnel collected blood samples from participants after an overnight fast under standardized conditions [13]. For studies on sex steroid hormones, we selected men who participated in the morning session to reduce diurnal variation in serum levels. Men with a self-reported history of cardiovascular disease or prostate cancer were excluded because certain treatments may

affect hormone levels. From this population, we selected 1316 male participants aged 20 and older in the first phase of NHANES III as the discovery set, and findings from this data set were replicated in the continuous NHANES 1999–2004 ( $n = 890$ ).

The continuous NHANES 1999–2004 is a program of studies undertaken by the NCHS of the US Centers for Disease Control and Prevention (CDC) to assess the health and nutritional status of adults and children in the USA. This study includes data from 890 male participants in the 1999–2000, 2001–2002, and 2003–2004 NHANES cycles with complete sex hormone data. Similarly to NHANES III, continuous NHANES uses a multi-stage, stratified, and clustered probability sampling strategy in which minority populations and the elderly are oversampled to ensure adequate representativeness of the total US civilian, non-institutionalized population. Continuous NHANES consists of an interview and an examination component that includes morning blood collection to reduce diurnal variation of hormones, and men with a history prostate cancer and younger than 20 years of age were excluded for this study. Investigators are allowed to access surplus sera for approved studies. Details of the survey design, methods, and data collections are available on the NHANES website (<https://www.cdc.gov/nchs/nhanes/index.htm>). Accessed May 2018).

The NHANES program is approved by the Institutional Review Board of the NCHS at CDC. Informed consent was obtained from all participants. Addressing questions about hormones and men’s health in NHANES was approved by the National Institutes of Health Office of Human Subjects Research and the NCHS Ethics Review Board at CDC.

### Assessment of testosterone, estradiol, and SHBG

In NHANES III, serum concentrations of total T, total estradiol, and SHBG were previously measured in Dr. Nader Rifai’s laboratory at Boston Children’s Hospital, MA, by electrochemiluminescence immunoassays on a 2010 Elecsys autoanalyzer (Roche Diagnostics, Indianapolis, IN, USA). The lowest limits (LOD) of detection were 2 ng/mL for T, 5 pg/mL for estradiol, and 3 nmol/L for SHBG; no samples had concentrations below LOD for these analytes. Values of total T below or equal to 350 ng/dL [ $12.15 \text{ nmol L}^{-1}$ ] are considered T deficiency [9].

Concerning continuous NHANES 1999–2004, details on the blood draw, process, storage, and shipping methods are provided elsewhere [14]. As in NHANES III, total T, estradiol, and SHBG were measured using the Elecsys 2010 system (Roche Diagnostics, Laval, QC, Canada). The laboratory technicians were blinded to the participant characteristics of the samples. The LOD of the assays were 2 ng/dL for T, 5 pg/mL for estradiol, and 3 nmol/L for SHBG. One sample had a concentration below the LOD for T and 10 samples for estradiol, which were assigned to half the limit of detection. T

deficiency was also defined as total T below or equal to 350 ng/dL [12.15 nmol L<sup>-1</sup>] [9]. In this study, the term T deficiency does not imply that a deficit needs to be replaced; therefore, its use is analogous to low T, but we did not use the term to avoid confusion with the business centers using the same words. Twenty-one samples were assayed in duplicate for quality control purposes, and the coefficients of variation were 4.8% for testosterone, 21.4% for estradiol, and 5.6% for SHBG. Testosterone deficiency with calculated free testosterone (< 225 pM) was also calculated [9].

### Assessment of exposures and covariates

There was a total of 173 factors, including 133 nutritional and 40 lifestyle factors, in our EWAS in NHANES III (Supplementary Table S1, Supplementary Data). The data collected ranged from information obtained through the interview, such as smoking history to physical and laboratory examination, for example serum vitamin C concentrations. Examples of markers and categories are presented in

Table 1. These factors were assessed either as continuous or as categorical variables. The majority of continuous variables had a right-skewed distribution. We transformed these variables into standardized z-scores by subtracting the mean and dividing by the standard deviation (SD) of the population. For categorical variables, we consistently defined one value as the referent category or the “negative” result, for example “never smoker” as the reference for “current smoker.” Vigorous physical activity (yes, no) was defined as participating three or more times per week in leisure-time physical activities with metabolic equivalent  $\geq 6$  for those aged 60 and older, and metabolic equivalent  $\geq 7$  for those younger than 60. Secondary exposure to smoking among never smokers (never smoked  $\geq 100$  cigarettes) was defined as exposure to smoke at home ( $\geq 1$  person smoke at home) or at work ( $\geq 1$  h smoke exposure at work). All exposure variables were assessed with standard procedures as detailed in the NHANES III documentation [13].

The following variables have been suggested as strongly affecting nutritional and environmental factors and T levels

**Table 1** Number and examples of nutrition and lifestyle factors in NHANES III

Factor category	Number of variables <sup>†</sup>	Examples
<b>Nutrition</b>		
Food nutrient recall	104	Dietary fiber (continuous) Aspartame (continuous) Energy from protein (continuous)
Healthy Eating Index (HEI)	10	Total HEI score (continuous)
Nutrients and minerals (serum and urine)	17	Serum vitamin A (continuous) Serum selenium (continuous)
Alcohol use	1	Drink alcohol twice or more a day (yes/no)
Caffeinated beverages	1	Drink caffeinated beverages twice or more a day (yes/no)
<b>Lifestyle factors</b>		
Personal smoking	14	Current smoker (reference: never smoker) Ever smoked 20 cigars in life (yes/no) Serum cotinine (continuous)
Environmental smoking	4	Does anyone smoke in the home (yes/no) At work, hours per day can smell smoking (ordinal)
Physical activity	1	Vigorous physical activity (yes/no)
Social support	2	How many times talking on the phone with family, friends, or neighbors per week (ordinal)
Environmental pollutants	2	Serum lead (continuous) Urine cadmium (continuous)
Bacterial infection	2	<i>Helicobacter pylori</i> antibody (continuous)
Viral infection	8	Herpes simplex virus I antibody (positive/negative) Hepatitis A antibody (positive/negative)
Parasite infection	1	Toxoplasma antibody (continuous)
Painkiller use	6	Taken aspirin in the past month (yes/no) # NSAIDs taken in the past month (ordinal)

<sup>†</sup> This table summarizes the number of variables available for each factor category from the Supplementary Table S1 (e.g., there were 104 available variables under the Nutrition category, and 10 variables to define the Healthy Eating Index)

and were therefore considered as confounders in our study: age, race/ethnicity, education, body mass index (BMI), estradiol, SHBG, and socioeconomic status (SES). Race/ethnicity was categorized into non-Hispanic white, non-Hispanic black, Mexican-American, and other. We classified educational attainment as less than high school, high school equivalent, and higher than high school. BMI was calculated from measured weight and height (weight in kilograms divided by height in meters squared). SES was estimated with poverty-to-income ratio (PIR), a ratio of total family income to the official poverty threshold at the family level. A PIR < 1 indicated that income was less than the level of poverty. We categorized PIR into < 1, 1–< 2, 2–< 3, and ≥ 3, indicating the lowest to highest SES as previously described [15].

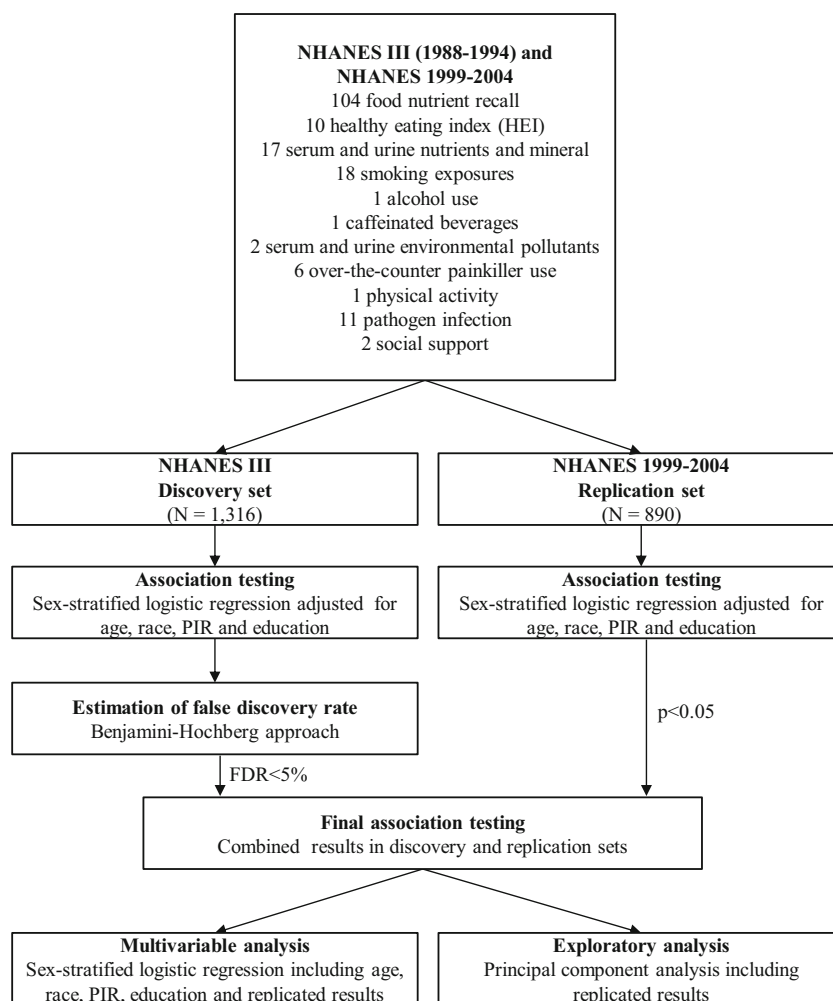
Detailed information on the laboratory methods used and quality measures for metabolic factors/disorders in NHANES III have been reported previously [13, 16]. Briefly, hypertension was defined by blood pressure of ≥ 130/≥ 85 mmHg or any use of antihypertensive drugs. Diabetes was defined as fasting glucose ≥ 110 mg/dL or any use of insulin or

glucose-lowering drugs. Any HDL cholesterol levels < 40 mg/dL for men were considered as low. Triglyceridemia was defined as any levels of triglycerides ≥ 150 mg/dL. Blood pressure was measured with mercury manometer and the average of the second and third blood pressure measurements was taken. Fasting plasma glucose levels were measured by using a modified hexokinase enzymatic method (Roche Diagnostic Systems, Inc., Montclair, NJ, USA). Blood lipids were enzymatically measured using the Hitachi 704 Analyzer (Boehringer Mannheim Diagnostics, Indianapolis, IN, USA).

## Statistical analysis

Sampling weights specific for each phase were applied to take into account selection probabilities, over-sampling, non-response, and differences between the sample and the total US population. Figure 1 depicts the analytical steps in this study which were similar to previously published nutrient- and environment-wide studies [12, 17]. T deficiency was used in the dichotomized format as the main outcome [9]

**Fig. 1** Diagram representing environment-wide analysis of nutrition and lifestyle factors in the NHANES III and NHANES 1999–2004



given that this definition is becoming widely known and accepted due to its clinical relevance with health outcomes, including chronic diseases [9, 18]. We selected factors corresponding to categories of environmental exposures used in previous EWAS studies as summarized in Table 1 [12, 17]. First, each of the 173 nutrition and lifestyle factors was assessed in relation to T deficiency in the discovery set, phase I of NHANES III. Survey-weighted logistic regression was used to examine the association of continuous and dichotomous nutrition and lifestyle factors in men. All models were linearly adjusted for age, race/ethnicity, education, BMI, and PIR by adding each term into the regression model. Pearson correlation coefficients were calculated for environmental and nutritional factors visualized with a heatmap, as shown in Supplementary Fig. S2.

We estimated the false discovery rate (FDR) among findings in the discovery set. FDR is the expected proportion of false discoveries, among all significant findings at a given significance level [19]. We estimated FDR using the Benjamini–Hochberg stepdown method to select factors with statistically significant association with T deficiency and  $FDR < 5\%$  in the discovery set [20].

Replicated analysis of assessed nutrition and lifestyle factors was subsequently performed by re-running similar logistic regression models in the second dataset, namely NHANES 1999–2004 for 16 out of 28 factors identified in the discovery dataset (because not all of them were available). Only nutrition or lifestyle factors with both  $FDR < 5\%$  in the first dataset and  $P$  value  $< 0.05$  in the replication set were considered valid. Furthermore, analysis for replicated factors in the overall study population was repeated in additional multivariable models adjusted for age, race, PIR, education, SHBG, and estradiol and models incorporating all replicated factors.

Finally, we conducted principal component analysis to identify any underlying factors associated with T deficiency based on intercorrelation among tentatively replicated nutrition and lifestyle factors in the overall study population. Principal component analysis was performed with an orthogonal varimax rotation procedure. An eigenvalue of 41 was used to define the number of principal components to be extracted from our data [21]. Proportion of variance in abdominal obesity explained by each principal component was estimated, and 95% confidence intervals of this estimation were obtained using 1000 bootstrap resampling [22]. We further estimated the value of the principal components identified and assessed them in relation to T deficiency using a similar multivariable approach.

The NHANES III datasets were prepared with Statistical Analysis Software release 9.3 (Statistical Analysis Software Institute, Cary, NC, USA). All analyses were performed with R version 3.1.2 (R Foundation for Statistical Computing) [23]. The *survey* package was used to account for sampling weights and the *psych* package was used to perform principal

component analysis. The *Circlize* package was used to generate the circle plot.

## Results

Descriptive characteristics of the study participants are shown in Table 2 (NHANES III and NHANES 1999–2004), and the means (s.d.) and frequencies of investigated nutrition ( $n = 133$ ) and lifestyle factors ( $n = 40$ ) are presented in Supplementary Table S1. Approximately 19% of men with T deficiency were older, non-Hispanic white, had higher BMI, hypertension, diabetes, low HDL cholesterol, and triglyceridemia, but lower poverty-to-income ratio, SHBG, and estradiol levels compared to men without deficiency (Table 2). Similar characteristics were observed for men in NHANES 1999–2004 with the exception of SHBG, which was not statistically different between the groups and which showed that poverty-to-income ratio was higher among men with T deficiency (Table 2). The distribution of these characteristics was also observed by calculated free testosterone deficiency, and the differences were similar (Supplementary Table S3).

We performed a systematic screening of the relationships of the 173 nutrition and lifestyle factors with T deficiency in men. A total of 28 factors with  $FDR < 5\%$  in the discovery set were examined for significance ( $P < 0.05$ ) in the replication set but only 16 were available to replicate (Supplementary Fig. S1 and Supplementary Table S2). Supplementary Figure 1 shows the distribution of  $P$  values for each investigated factors and effect sizes in a circle Manhattan plot and the detailed results on these associations are presented in Supplementary Table S2. This resulted in seven tentatively replicated factors showing significant inverse associations for vitamin A intake ( $OR = 0.64$ , 95% CI 0.46–0.89,  $P$  value = 0.01), protein intake ( $OR = 0.60$ , 95% CI 0.45–0.81,  $P$  value = 0.003), saturated fatty acid intake ( $OR = 0.67$ , 95% CI 0.51–0.89,  $P$  value = 0.01), monounsaturated fatty acid intake ( $OR = 0.68$ , 95% CI 0.52–0.88,  $P$  value = 0.008), total fat intake ( $OR = 0.70$ , 95% CI 0.54–0.91,  $P$  value = 0.01), SFA 16:0 ( $OR = 0.68$ , 95% CI 0.52–0.90,  $P$  value = 0.01), and phosphate intake ( $OR = 0.61$ , 95% CI 0.42–0.87,  $P$  value = 0.01) with T deficiency (Table 3).

Correlation between replicated factors is displayed as a heatmap in Supplementary Fig. S2. The larger the correlation between a pair of variables, the closer in proximity they appear in the heatmap. We performed a principal component factor analysis to assess any structure underlying replicated factors and to identify common underlying factors. Two principal components (PC) were identified. The first PC1 consisted of total fat intake, SFA 16:0, saturated fat, monounsaturated fat, protein, and phosphate intake. The second PC2 mainly comprised vitamin A intake. The total variance of replicated variables explained by PC1 was 73% (95% CI 72–75%) and 14%



**Table 2** Descriptive characteristics of study participants in NHANES III (discovery set) and NHANES 1999–2004

	NHANES III			NHANES 1999–2004		
	Testosterone deficiency ( <i>N</i> = 243)	No testosterone deficiency ( <i>N</i> = 1073)	<i>P</i> value	Testosterone deficiency ( <i>N</i> = 213)	No testosterone deficiency ( <i>N</i> = 677)	<i>P</i> value
Age, mean (SD)	51.40 (1.08)	41.36 (0.73)	< 0.001	50.50 (1.62)	41.35 (0.79)	< 0.001
Race (%) <sup>†</sup>			0.572			0.060
Non-Hispanic white	76.72	78.89		74.95	71.05	
Non-Hispanic black	9.54	9.52		6.73	11.37	
Mexican-American	4.03	4.65		12.82	13.16	
Others	9.71	6.94		5.50	4.42	
Poverty-to-income ratio (PIR) (%) <sup>†</sup>			0.110			0.984
< 1	3.94	10.76		10.77	10.76	
1–2	21.16	18.30		17.30	22.38	
2–3	32.09	19.74		10.78	14.62	
≥ 3	42.81	51.20		61.16	52.24	
Body mass index (kg/m <sup>2</sup> ), mean (SD)	28.70 (0.70)	25.93 (0.70)	< 0.001	31.82 (0.80)	26.67 (0.24)	< 0.001
Metabolic disorder (%) <sup>†</sup>						
Hypertension	61.69	37.37	< 0.001	59.25	40.82	< 0.001
Diabetes	19.17	9.94	0.023	26.84	11.59	0.001
Low HDL cholesterol	40.73	30.89	0.040	34.96	28.71	0.073
Triglyceridemia	49.02	28.84	< 0.001	44.62	30.89	< 0.001
Hormone levels, mean (SD)						
Testosterone (ng/mL)	2.62 (0.10)	5.92 (0.10)	< 0.001	2.66 (0.07)	5.82 (0.10)	< 0.001
SHBG (nmol/L)	32.30 (1.95)	39.24 (0.78)	0.008	23.24 (1.99)	32.33 (1.17)	0.133
Estradiol (ng/mL)	31.33 (1.64)	38.22 (0.63)	< 0.001	28.89 (1.75)	39.24 (1.00)	< 0.001

<sup>†</sup> These categorical variables are presented as relative frequencies (%)

(95% CI 13–15%) by PC2 (Supplementary Fig. S3). For the final analysis, we obtained estimates for replicated nutritional and lifestyle factors in relation to T deficiency in multivariable analysis in the NHANES 1999–2004 dataset (Table 4). The only factor that remained statistically significant associated to T deficiency was vitamin A (OR = 0.67, 95% CI 0.49–0.92).

## Discussion

In a systematic screening of 173 nutrition and lifestyle factors, seven factors were identified and validated as being inversely associated with T deficiency after applying the EWAS methodology in a representative sample of the US population. These seven factors included intake of total vitamin A,

**Table 3** Associations between replicated nutrition and lifestyle factors in relation to T deficiency in the discovery and replication datasets

Factors	Discovery dataset (NHANES III)			Replication dataset (NHANES 1999–2004)			
	<i>N</i> = T deficiency	<i>N</i> = Total	OR (95% CI)	<i>N</i> = T deficiency	<i>N</i> = Total	OR (95% CI)	<i>P</i> value
Vitamin A intake (g)	234	1273	0.90 (0.71–0.91)	140	565	0.64 (0.46–0.89)	0.010
Protein intake (g)	234	1273	0.67 (0.53–0.85)	140	565	0.60 (0.45–0.81)	0.003
Saturated fatty acid intake (g)	234	1273	0.71 (0.58–0.87)	140	565	0.67 (0.51–0.89)	0.011
Monounsaturated fatty acid intake (g)	234	1273	0.76 (0.64–0.89)	140	565	0.68 (0.52–0.88)	0.008
Total fat intake (g)	234	1273	0.75 (0.63–0.89)	140	565	0.70 (0.54–0.91)	0.012
SFA 16:0 (g)	234	1273	0.71 (0.58–0.88)	140	565	0.68 (0.52–0.90)	0.014
Phosphate intake (mg)	234	1273	0.70 (0.56–0.87)	140	565	0.61 (0.42–0.87)	0.010

All models were adjusted for age, BMI, race/ethnicity, education, PIR, SHBG, and estradiol. Benjamini–Hochberg adjusted *P* values for FDR < 5% are presented for the discovery dataset, and *P* values from significance testing are presented for replication dataset

**Table 4** Multivariable analysis of replicated factors and T deficiency in NHANES 1999–2004

	Weighted mean (SD)	OR (95% CI) per SD change	P value
Vitamin A intake (g)	0.18 (0.05)	0.67 (0.49–0.92)	0.012
Protein intake (g)	0.67 (0.05)	0.71 (0.39–1.29)	0.127
Saturated fatty acid intake (g)	0.49 (0.06)	0.55 (0.18–1.73)	0.307
Monounsaturated fatty acid intake (g)	0.61 (0.05)	0.38 (0.11–1.37)	0.133
Total fat intake (g)	0.62 (0.05)	2.48 (0.54–11.42)	0.245
SFA 16:0 (g)	0.55 (0.06)	1.49 (0.38–5.83)	0.580
Phosphate intake (mg)	0.60 (0.05)	1.08 (0.55–2.13)	0.828

† Adjusted for age, race, PIR, BMI, SHBG, and estradiol

protein, total saturated fatty acids, total monounsaturated fatty acids, total fats, saturated fatty acid 16:0, and phosphorus. In a multivariable model in NHANES 1999–2004, the seven replicated nutritional and lifestyle factors were investigated in relation to T deficiency and only vitamin A was significantly associated. On the basis of intercorrelation between these factors, two PCs were found: (1) total fats, saturated fatty acid 16:0, total saturated fatty acids, total monounsaturated fatty acids, protein, and phosphorous; and (2) total vitamin A. The two PCs also displayed significant associations with T deficiency.

To our knowledge, this is the first study to investigate the association of a large number of nutritional and lifestyle factors with T deficiency using an EWAS approach—a study design analogous to GWAS. Although EWAS has previously been used with other health outcomes, there was a research gap in studying the role of nutritional and lifestyle factors in the production and secretion of T levels using this method. The seven nutritional factors that were tested and replicated in this study have previously been suggested as being linked with serum T levels. However, the emerging body of literature related to these associations has provided conflicting results.

Therefore, several aspects of the findings of our investigation merit further discussion. First, in our study, all associations between the nutritional factors and T deficiency were inverse, meaning that every unit increase of these nutritional factors had the potential to reduce the likelihood of T deficiency, possibly due to a direct increase in serum T levels. In our study, nutritional factors are dietary intakes that were compared with serum levels from previous papers, as described in the paragraphs below, and although in some instances our sample size was larger compared to previous studies, our dietary intake data are also more prone to recall and selection bias. Second, vitamin A was investigated in an experimental study conducted in guinea pigs and it was reported that low levels of vitamin A are accompanied by corresponding low levels of testosterone [24]. However, a population-based nutrition study, which included 155 male twins, identified an inverse association between vitamin A and T levels [25]. Years later, a review study, mainly conducted in animals, reported that low levels of vitamin A lead to decrease in T

production [26]. Further, a more recent experimental study by Yang et al. in 2018 seemed to agree with the latter review, as has our study, by reporting that low levels of vitamin A adversely affect T secretion through regulation of the Leydig cell differentiation in mouse and rat models [27]. However, this association should be interpreted with caution because dietary recalls and cross-sectional studies are more prone to recall and selection biases, which could have influenced these findings.

Third, a number of studies have reported that dietary fats (e.g., total fat, and saturated and unsaturated fats) play a role in the production and secretion of T, although these findings remain inconsistent as dietary fat has been found to have both positive and negative effects on circulating total T [28–32]. Our study observed an inverse association between total fat, saturated fats and unsaturated fats, and T deficiency. This observation has plausibility as it has been previously shown that saturated fat intake increased total cholesterol concentration [33], and total cholesterol is a precursor for T [34, 35]. Another plausibility derived from a high-fat diet study, rich in saturated fatty acids, found that this diet caused a decrease in SHBG and consequently an increase in the levels of free T [36]. Yet, an aforementioned male twin study of 155 pairs also investigated the relationship between T levels and total fat and reported an inverse relationship [25]. Similarly, Nagata et al. 2000 found that intake of saturated, monounsaturated, and polyunsaturated fats was inversely correlated with serum total testosterone among Japanese men, but the correlation was statistically significant only for the polyunsaturated fat [37]. However, Key et al. 1990 found a positive correlation only between T and polyunsaturated fat intake among 40 men [29]. Another small population-based study confirmed that 43 healthy men had high concentrations of total T among those men who ate a high-fat diet [36]. Confirming these latter observations, two experimental studies conducted in rat models showed positive associations between T levels and unsaturated fatty acids [32, 38].

Fourth, it has been suggested that dietary protein levels influence circulating T levels. However, the direction of this influence remains unclear due to contradicting results found in the literature. A few experimental studies have reported that a low-protein diet can lead to low levels of T in male rat models

[39, 40], but an inverse association was found between protein and T levels in a number of population-based studies [25, 41]. Interestingly, more recent literature has focused mainly trying to compare and assess the effects on T levels by different types of proteins (e.g., soy vs whey protein which is the most common type of protein investigated). For instance, Kraemer et al. 2013 identified lower levels of T among 10 young men following a supplementation with soy protein, but not with whey protein [42]. However, a recent meta-analysis [43, 44] showed that soy protein does not alter T profile, and it is not inferior to the positive effects of whey protein [45]. This finding was recently confirmed by a double-blind randomized clinical trial reporting no alteration in the T profile after soy or whey protein supplementation among young men with resistance exercise training [46].

Finally, there has been a small number of studies investigating the link between phosphorous and T levels. However, this link merits further scrutiny because high levels of phosphorous have been associated with a high risk of cardiovascular diseases [47], but low levels of T have been linked to a higher risk of cardiovascular disease [4, 48]. A community-dwelling study of 1346 older men reported that high serum T levels were associated with lower serum phosphorus levels in a model adjusted for age, race, and estradiol levels [49]. This seems to concur with the aforementioned statements that low T levels increased cardiovascular diseases, but high phosphorous levels have the same effect. Interestingly, a recent cross-sectional study of 1899 Korean men showed a non-significant correlation between serum phosphorus and T [50].

Our study has a number of strengths. NHANES is a program of studies that is representative of the civilian non-institutionalized US population, which aids in the generalizability of these results. In addition, NHANES follows a rigorous protocol with extensive quality control procedures for the collection of the exposures, outcome of interest, and potential confounding factors analyzed and adjusted in this study. We also mutually adjusted for testosterone, SHBG, and estradiol. Robustness of the statistical associations between investigated markers and T deficiency was ascertained through replication analysis and adjustment for presence of major confounders. The systematic screening was able to eliminate factors with small effects which may be more prone to bias. Additionally, this method overcomes the limitation of selective reporting, which may be an issue in studies focusing on individual exposures.

Our findings should be interpreted in the context of the study design due to its inherent limitations. First, both NHANES III and NHANES 1999–2004 are cross-sectional studies; therefore, we cannot infer causality. Further, due to the cross-sectional nature of the data, we cannot rule out the issue of reverse causality between T deficiency and several dietary and lifestyle factors. Second, although in our study we followed a systematic approach that can provide a list of T deficiency correlates with strong statistical power, it will be necessary to conduct other

study designs/models (e.g., causal inference modeling and randomized trials) to determine which of these correlates are the most important. Third, even though we took into account potential confounders and intercorrelation between replicated factors, residual confounding may have occurred because the information of these confounders was obtained by self-reporting. We were unable to exclude the potential role of other relevant factors apart from those assessed in NHANES III. Therefore, obtaining an equivalent definition of “genome-wide significance” as one would be able to claim in a genome-wide association studies analysis may be impractical or else requires more rigorous and thorough assessments of nutrition and lifestyle determinants. However, this study has the potential to contribute to the current body of literature which seeks to formulate a clinical definition of T deficiency.

## Conclusion

This is the first study to use an EWAS approach to identify nutritional factors that were robustly associated with T deficiency. This study has the potential to contribute to the current body of literature that aims to formulate a clinical definition of T deficiency after taking into account nutritional and lifestyle factors.

**Funding information** David S. Lopez was supported by the National Institutes of Health and National Institute on Aging, Grant #: 1P30AG059301-01

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** The protocols for the conduct of NHANES III and NHANES 1999–2004 were approved by the Institutional Review Board of the National Center for Health Statistics, US Centers for Disease Control and Prevention. Informed consent was obtained from all participants.

**Informed consent** We conducted a secondary data analysis using data from NHANES, which can be accessed by the public. NHANES obtained informed consents from participants.

**Submission declaration and verification** This work has not been published previously and it is not under consideration for publication elsewhere. The publication is approved by all authors and, if accepted, will not be published elsewhere in the same form, in English or in any other language.

## References

1. Khera M, Adaikan G, Buvat J, Carrier S, El-Meliegy A, Hatzimouratidis K, McCullough A, Morgentaler A, Torres LO, Salonia A (2016) Diagnosis and treatment of testosterone



- deficiency: recommendations from the fourth international consultation for sexual medicine (ICSM 2015). *J Sex Med* 13:1787–1804
2. Buvat J, Maggi M, Guay A, Torres LO (2013) Testosterone deficiency in men: systematic review and standard operating procedures for diagnosis and treatment. *J Sex Med* 10:245–284
  3. Corona G, Rastrelli G, Di Pasquale G, Sforza A, Mannucci E, Maggi M (2018) Endogenous testosterone levels and cardiovascular risk: meta-analysis of observational studies. *J Sex Med* 15:1260–1271
  4. Lopez DS, Canfield S, Wang R (2016) Testosterone replacement therapy and the heart: friend, foe or bystander? *Transl Androl Urol* 5:898–908
  5. Lopez DS, Advani S, Tsilidis KK, Wang R, Canfield S (2017) Endogenous and exogenous testosterone and prostate cancer: decreased-, increased- or null-risk? *Transl Androl Urol* 6:566–579
  6. Platz EA (2008) Low testosterone and risk of premature death in older men: analytical and preanalytical issues in measuring circulating testosterone. *Clin Chem* 54:1110–1112
  7. Mulligan T, Frick MF, Zuraw QC, Stenham A, McWhirter C (2006) Prevalence of hypogonadism in males aged at least 45 years: the HIM study. *Int J Clin Pract* 60:762–769
  8. Araujo AB, Esche GR, Kupelian V, O'Donnell AB, Travison TG, Williams RE, Clark RV, McKinlay JB (2007) Prevalence of symptomatic androgen deficiency in men. *J Clin Endocrinol Metab* 92:4241–4247
  9. Paduch DA, Brannigan RD, Fuchs EF, Kim ED, Marmar JL, Sandlow JI. The laboratory diagnosis of testosterone deficiency. Available from <http://university.aunet.org/common/pdf/education/clinical-guidance/Testosterone-Deficiency-WhitePaper.pdf>. White Paper 2017
  10. Malik RD, Lapin B, Wang CE, Lakeman JC, Helfand BT (2015) Are we testing appropriately for low testosterone?: characterization of tested men and compliance with current guidelines. *J Sex Med* 12:66–75
  11. Rohrmann S, Shiels MS, Lopez DS, Rifai N, Nelson WG, Kanarek N, Guallar E, Menke A, Joshi CE, Feinleib M, Sutcliffe S, Platz EA (2011) Body fatness and sex steroid hormone concentrations in US men: results from NHANES III. *Cancer Causes Control* 22:1141–1151
  12. Patel CJ, Rehkopf DH, Leppert JT, Bortz WM, Cullen MR, Chertow GM, Ioannidis JP (2013) Systematic evaluation of environmental and behavioural factors associated with all-cause mortality in the United States national health and nutrition examination survey. *Int J Epidemiol* 42:1795–1810
  13. National Center for Health Statistics (1994) Plan and operation of the Third National Health and Nutrition Examination Survey, 1988–94. Series 1: programs and collection procedures. *Vital Health Stat* 1:1–407 [https://www.cdc.gov/nchs/data/series/sr\\_01/sr01\\_032.pdf](https://www.cdc.gov/nchs/data/series/sr_01/sr01_032.pdf)
  14. Nyante SJ, Graubard BI, Li Y, McQuillan GM, Platz EA, Rohrmann S, Bradwin G, McGlynn KA (2012) Trends in sex hormone concentrations in US males: 1988–1991 to 1999–2004. *Int J Androl* 35:456–466
  15. Loucks EB, Rehkopf DH, Thurston RC, Kawachi I (2007) Socioeconomic disparities in metabolic syndrome differ by gender: evidence from NHANES III. *Ann Epidemiol* 17:19–26
  16. Ma Z, Gingerich RL, Santiago JV, Klein S, Smith CH, Landt M (1996) Radioimmunoassay of leptin in human plasma. *Clin Chem* 42:942–946
  17. Wulaningsih W, Van Hemelrijck M, Tsilidis KK, Tzoulaki I, Patel C, Rohrmann S (2017) Investigating nutrition and lifestyle factors as determinants of abdominal obesity: an environment-wide study. *Int J Obes (Lond)* 41:340–347
  18. Lopez DS, Qiu X, Advani S, Tsilidis KK, Khera M, Kim J, Morgentaler A, Wang R, Canfield S (2018) Double trouble: co-occurrence of testosterone deficiency and body fatness associated with all-cause mortality in US men. *Clin Endocrinol (Oxf)* 88:58–65
  19. Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100:9440–9445
  20. Benjamini Y, Hechtlinger Y (2014) Discussion: an estimate of the science-wise false discovery rate and applications to top medical journals by Jager and Leek. *Biostatistics* 15:13–16 discussion 39–45
  21. Lee JY, Chen HI, Urban C, Hojat A, Church E, Xie SX, Farrar JT (2010) Development of and psychometric testing for the brief pain inventory-facial in patients with facial pain syndromes. *J Neurosurg* 113:516–523
  22. Timmerman ME, Kiers HA, Smilde AK (2007) Estimating confidence intervals for principal component loadings: a comparison between the bootstrap and asymptotic results. *Br J Math Stat Psychol* 60:295–314
  23. Lumley T. Survey: analysis of complex survey samples. R Package version 3.14 (2009). Available at: <https://cran.r-project.org/web/packages/survey/index.html>. (Accessed: 15th April 2015)
  24. Nayyar T, Mukherjee S, Das SK (2000) Alterations in binding characteristics of peripheral benzodiazepine receptors in testes by vitamin A deficiency in guinea pigs. *Mol Cell Biochem* 211:47–50
  25. Bishop DT, Meikle AW, Slattery ML, Stringham JD, Ford MH, West DW (1988) The effect of nutritional factors on sex hormone levels in male twins. *Genet Epidemiol* 5:43–59
  26. Livera G, Rouiller-Fabre V, Pairault C, Levacher C, Habert R (2002) Regulation and perturbation of testicular functions by vitamin A. *Reproduction* 124:173–180
  27. Yang Y, Luo J, Yu D, Zhang T, Lin Q, Li Q, Wu X, Su Z, Zhang Q, Xiang Q, Huang Y (2018) Vitamin A promotes leydig cell differentiation via alcohol dehydrogenase 1. *Front Endocrinol (Lausanne)* 9:644
  28. Howie BJ, Shultz TD (1985) Dietary and hormonal interrelationships among vegetarian Seventh-Day Adventists and nonvegetarian men. *Am J Clin Nutr* 42:127–134
  29. Key TJ, Roe L, Thorogood M, Moore JW, Clark GM, Wang DY (1990) Testosterone, sex hormone-binding globulin, calculated free testosterone, and oestradiol in male vegans and omnivores. *Br J Nutr* 64:111–119
  30. Belanger A, Locong A, Noel C, Cusan L, Dupont A, Prevost J, Caron S, Seigny J (1989) Influence of diet on plasma steroids and sex hormone-binding globulin levels in adult men. *J Steroid Biochem* 32:829–833
  31. Raben A, Kiens B, Richter EA, Rasmussen LB, Svenstrup B, Micic S, Bennett P (1992) Serum sex hormones and endurance performance after a lacto-ovo vegetarian and a mixed diet. *Med Sci Sports Exerc* 24:1290–1297
  32. Marra CA, de Alaniz MJ (1989) Influence of testosterone administration on the biosynthesis of unsaturated fatty acids in male and female rats. *Lipids* 24:1014–1019
  33. Grundy SM, Vega GL (1990) Causes of high blood cholesterol. *Circulation* 81:412–427
  34. Huff T, Jialal I (2018) Physiology, cholesterol. In: Anonymous StatPearls. StatPearls Publishing LLC, Treasure Island (FL)
  35. Zhang X, Li J, Zhou X, Guan Q, Zhao J, Gao L, Yu C, Wang Y, Zuo C (2017) Simvastatin decreases sex hormone levels in male rats. *Endocr Pract* 23:175–181
  36. Dorgan JF, Judd JT, Longcope C, Brown C, Schatzkin A, Clevidence BA, Campbell WS, Nair PP, Franz C, Kahle L, Taylor PR (1996) Effects of dietary fat and fiber on plasma and urine androgens and estrogens in men: a controlled feeding study. *Am J Clin Nutr* 64:850–855
  37. Nagata C, Takatsuka N, Kawakami N, Shimizu H (2000) Relationships between types of fat consumed and serum estrogen and androgen concentrations in Japanese men. *Nutr Cancer* 38:163–167

38. Gromadzka-Ostrowska J, Przepiorka M, Romanowicz K (2002) Influence of dietary fatty acids composition, level of dietary fat and feeding period on some parameters of androgen metabolism in male rats. *Reprod Biol* 2:277–293
39. Hanai M, Esashi T (2007) The interactive effect of dietary protein and vitamin levels on the depression of gonadal development in growing male rats kept under disturbed daily rhythm. *J Nutr Sci Vitaminol (Tokyo)* 53:138–144
40. Hanai M, Esashi T (2012) Effect of dietary protein levels on sex hormones in growing male rats kept under constant darkness. *Exp Anim* 61:555–561
41. Anderson KE, Rosner W, Khan MS, New MI, Pang SY, Wissel PS, Kappas A (1987) Diet-hormone interactions: protein/carbohydrate ratio alters reciprocally the plasma levels of testosterone and cortisol and their respective binding globulins in man. *Life Sci* 40:1761–1768
42. Kraemer WJ, Solomon-Hill G, Volk BM, Kupchak BR, Looney DP, Dunn-Lewis C, Comstock BA, Szivak TK, Hooper DR, Flanagan SD, Maresh CM, Volek JS (2013) The effects of soy and whey protein supplementation on acute hormonal responses to resistance exercise in men. *J Am Coll Nutr* 32:66–74
43. Hamilton-Reeves JM, Vazquez G, Duval SJ, Phipps WR, Kurzer MS, Messina MJ (2010) Clinical studies show no effects of soy protein or isoflavones on reproductive hormones in men: results of a meta-analysis. *Fertil Steril* 94:997–1007
44. Messina M, Hamilton-Reeves J, Kurzer M, Phipps W (2007) Effect of soy protein on testosterone levels. *Cancer Epidemiol Biomarkers Prev* 16:2795
45. Kalman D, Feldman S, Martinez M, Krieger DR, Tallon MJ (2007) Effect of protein source and resistance training on body composition and sex hormones. *J Int Soc Sports Nutr* 4:4–2783–44
46. Reidy PT, Borack MS, Markofski MM, Dickinson JM, Deer RR, Husaini SH, Walker DK, Igbini S, Robertson SM, Cope MB, Mukherjee R, Hall-Porter JM, Jennings K, Volpi E, Rasmussen BB (2016) Protein supplementation has minimal effects on muscle adaptations during resistance exercise training in young men: a double-blind randomized clinical trial. *J Nutr* 146:1660–1669
47. Foley RN, Collins AJ, Herzog CA, Ishani A, Kalra PA (2009) Serum phosphorus levels associate with coronary atherosclerosis in young adults. *J Am Soc Nephrol* 20:397–404
48. Kloner RA, Carson C 3rd, Dobs A, Kopecky S, Mohler ER 3rd (2016) Testosterone and cardiovascular disease. *J Am Coll Cardiol* 67:545–557
49. Meng J, Ohlsson C, Laughlin GA, Chonchol M, Wassel CL, Ljunggren O, Karlsson MK, Mellstrom D, Orwoll ES, Barrett-Connor E, Ix JH, Osteoporotic Fractures in Men (MrOs) Study Group (2010) Associations of estradiol and testosterone with serum phosphorus in older men: the osteoporotic fractures in men study. *Kidney Int* 78:415–422
50. Min SK, Choi K, Kim SK, Lee GI, Cho IC (2016) Phosphorus as predictive factor for erectile dysfunction in middle aged men: a cross sectional study in Korea. *Investig Clin Urol* 57:442–448

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.