IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Association of Estrogen Use with Telomere Length: An NHANES Investigation

Kayleigh Fong

St Michael's School, Bryn, Llanelli, Carmarthenshire, SA14 9TU, UK; kayleighfong@gmail.com
Mentor: Saira Luckman

ABSTRACT: A telomere is a repetitive DNA sequence located at the end of human chromosomes, and its length shortens with age. Research has demonstrated that estrogen is associated with the aging of various types of cells, including brain, muscle, and bone cells. This study explored the cross-sectional association between estrogen usage and telomere length, utilizing data from the US National Health and Nutrition Examination Survey (NHANES) 1999-2002. The final number of participants included in this study was 491. This study's bivariate Rao-Scott $\chi 2$ test revealed a significant association between telomere length and gender ($p < 0.01$). Females have longer telomeres than males, indicated by a rate ratio of 1.72. A positive relationship between estrogen usage and telomere length was observed among female participants, yielding a rate ratio of 1.44. The mosaic diagram and Pearson residuals confirmed the statistical significance of this difference ($p < 0.001$). This suggests that estrogen administration may preserve telomeres and contribute to healthy aging. In conjunction with the findings from the literature, this study supports the need for further investigations into the effects of estrogen administration on telomere length.

KEYWORDS: Translational Medical Sciences; Disease Prevention; Anti-ageing; Estrogen; Telomere.

Introduction

Global life expectancy has been rising in recent decades, leading to an increase in age-associated diseases such as cardiovascular disease, Alzheimer's disease, and cancers.[1] These diseases are associated with repetitive nucleotide sequences found at the ends of human chromosomes called telomeres.[1] Telomeres are protective caps on chromosomes that prevent them from unraveling or fusing with other chromosomes. They are composed of repetitive DNA sequences and associated proteins that help maintain the stability and integrity of the genome. Telomeres shorten with each cell division, and when they reach a critical length, the cell enters a state of senescence or apoptosis.[1] Therefore, telomere length (TL) is a marker of cellular aging and biological age. This phenomenon has prompted research into factors that may slow telomere shortening and promote healthy aging.

TL is influenced by both genetic and environmental factors, such as oxidative stress, inflammation, lifestyle, and chronic diseases.[2] Several studies have reported associations between TL and various age-related diseases, such as cardiovascular disease, cancer, dementia, osteoporosis, and diabetes.[3] However, the causal relationship between TL and these diseases is not fully understood, and the direction of the association may vary depending on the tissue type, disease stage, and measurement method.[3]

Several studies have reported an association between TL and life expectancy, suggesting that longer telomeres may confer a survival advantage.[1,4] One of the factors that may affect TL and longevity is gender.[5,6] Women generally live longer than men and have longer telomeres in various tissues. The reasons for this sex difference are unclear, but possible explanations include oxidative stress, immune system function, genetic factors, and hormonal effects.[7] Oxidative stress may cause more damage to the male than the female genome, leading to faster telomere erosion.[8] The immune system may also play a role, as chronic inflammation can accelerate telomere attrition, and women have a more robust immune response than men.[9] Additionally, genetic factors may contribute to the sex difference in TL, as women have two copies of the X chromosome, which contains several genes related to telomere maintenance.[10]

Estrogen may protect telomeres from shortening by enhancing the activity of telomerase, an enzyme that adds DNA to the telomeres. Animal studies have concluded that estrogen can activate telomerase, extending the length of telomeres, and mice with estrogen deficiency had shorter telomeres.[11,12] However, no studies have been found on the association between human estrogen users and TL.

Estrogen is a hormone that plays a vital role in women's health and well-being. Estrogen levels fall naturally during menopause, which can lead to symptoms such as hot flashes, mood swings, vaginal dryness, and osteoporosis. One way to alleviate the adverse effects of low estrogen is to use hormone replacement therapy (HRT), which involves taking synthetic or natural estrogen. In addition to HRT, estrogen is also widely prescribed for contraceptive use. It works by inhibiting ovulation, which means that it can prevent the release of an egg from the ovaries. Therefore, estrogen is a widely prescribed steroid hormone for contraceptive use and the treatment of postmenopausal symptoms. This study investigated the association between estrogen use and TL using the US National Health and Nutrition Examination Survey (NHANES) (https://www.cdc.gov/nchs/nhanes/index.htm).

NHANES is a US national program designed to assess the health and nutritional status of adults and children in the US.

DOI: 10.36838/v6i10.1

The survey combines interviews and physical examinations performed in a two-year continuous cycle. NHANES collects data on various health topics, such as chronic diseases, infectious diseases, nutrition, environmental exposures, medication intake, and genetic factors. One of the genetic factors that NHANES measures is TL. NHANES measured TL in blood samples from participants aged 20 years and older in the two-year cycles of 1999-2000 and 2001-2002.

Several studies have utilized NHANES data to explore the associations between TL and various factors, such as pathogen burden,[13] lipoproteins,[14] cognitive functions, and [15] dietary intake.[16] However, no studies have been conducted to investigate the relationship between estrogen and TL, which may offer an understanding of the potential mechanisms of hormonal effects on telomere dynamics and their implications for health and aging. This study aimed to explore the cross-sectional association between telomere length, estrogen usage, and other covariates in a large sample of female participants using a US national database.

Methods

Dataset:

All the data of this study were obtained from the NHANES, a government-led survey program established in 1971 to assess the well-being of the US population. The survey contains various information categorized into demographics, dietary, examination, laboratory, and questionnaire data. Each data type has multiple subsets; for example, TL is a subset of the laboratory data. Most of these data are freely available to the public, and their quality suits scientific research.[17] This study employed the demographics data, TL data, and dietary supplements and prescription medication subset (DSQ) of the questionnaire data.

Telomere Length:

The TL data from NHANES were obtained from blood samples collected from participants aged 60 and older across the two-year survey cycles. These participants included males and females, thus allowing the investigation into TL differences between genders. Of note, only a proportion of the female participants were administered estrogen. The study exclusively focuses on participants aged 60 or older because telomere shortening is commonly associated with aging, and the shortening rate appears more pronounced in the elderly. Therefore, studying individuals aged 60 and older is crucial for revealing the effects of estrogen on TL. Additionally, the elderly tend to have a higher prevalence of illnesses, and investigating TL in this age group holds greater significance than studying a younger cohort.

The TL assay of NHANES was performed using the quantitative polymerase chain reaction (PCR) method to measure TL relative to standard reference DNA (T/S ratio).[18] The size of a gene, like telomeres, is typically small. PCR facilitates the amplification of targeted DNA segments, thereby enabling a relatively precise measurement of telomere length. The T/S ratio, short for Telomere-to-Single Copy Gene Ratio, is a metric employed to evaluate relative telomere length by comparing it with a single copy gene in the genome. This ratio aids in standardizing measurements, normalizing variations, and providing a more accurate measurement. The mean and standard deviation of the T/S ratio were then calculated for each participant. The TL data are only available for the two NHANES cycles between 1999 and 2001 and 2001 and 2002. The data are publicly available for researchers who wish to explore the genetic or hereditary disease patterns or risks in diverse populations.

Estrogen Usage:

The prescription medicines section of the NHANES dataset includes the names of various types of medicinal estrogen, primarily used for hormonal contraception or HRT. To identify all the estrogens listed in the two cycles of NHANES data, we manually reviewed the dataset and recorded each estrogen's name. These include desogestrel, estradiol, ethynodiol, estropipate, levonorgestrel, mestranol, methyltestosterone, medroxyprogesterone, norethindrone, norgestimate, and progesterone. We selected all the participants taking these estrogens within 90 days of the interview date to ensure sufficient exposure time for the estrogens to produce clinical effects on the participants. While no literature on PubMed or Google Scholar specifies the duration after taking an estrogen that may affect TL, we hypothesized that 90 days is adequate. This assumption is because the TL in the NHANES dataset was measured using leukocytes, which generally have a lifespan ranging from several hours to several days.[19] As the pharmacological mechanism of how estrogen may affect TL remains uncertain, we understand that estrogen may not directly interact with the leukocytes to alter their TL. Thus, setting the 90 days can limit our study, but we believe it is a reasonable estimate.

Covariates Assessment:

In addition to estrogen usage, this study incorporated additional variables identified in the literature as potential influencers of TL. Including these covariates enhances the validity of the study's outcomes, thereby providing a more accurate interpretation of the relationship between estrogen and TL. This study also aimed to find the associations between TL and these covariates. The covariates included in this study were race,[20] education level,[21] poverty income ratio,[22] marital status,[23] body mass index (BMI),[24] smoking status,[25] C-reactive protein level,[26] history of cancer,[27] diabetes,[28] analgesic usage,[29] and the number of cardiovascular diseases.[30]

This study categorized participants into distinct groups based on each covariate (Table 1). For BMI, the participants were divided into four categories: underweight (BMI < 18.5 kg/m^2), normal or healthy weight (BMI between 18.5 and 24.9 kg/m^2), overweight (BMI between 25.0 and 29.9 kg/m^2), and obese (BMI > 30.0 kg/m^2), based on the CDC's reference values.[31] For smoking status, the participants were divided into current smokers, never smokers, and former smokers. Participants who had smoked at least 100 cigarettes in their lifetime but were not currently smoking were classified as former smokers.

The poverty income ratio measures a household's income relative to the poverty threshold. A large value generally indicates a high household income or wealthy household status.

This study categorized the poverty income ratio into three groups with values of <2, 2–4, and >4, serving as a rough indicator of the lower income range, average, and higher income range, respectively.[32]

C-reactive protein is an indicator of inflammation in the body. This study categorized C-reactive protein into four quartiles, with the 4th quartile having a value of >0.54, indicating a high level of inflammation. In contrast, the 1st quartile has a value of 0.0–0.12, indicating a low level of inflammation.[32] For analgesic usage, the interview section of NHANES includes a question asking participants whether they have taken any painkillers daily for a month. If the answer was yes, the participant is considered an analgesic user.

This study also accounted for the total number of specific cardiovascular diseases for each participant, including heart failure, coronary heart disease, angina, heart attack, and stroke. These diseases were consolidated into a single covariate rather than treated as separate ones, as some participants did not provide information on these diseases. Using them as individual covariates could significantly reduce the sample size, as participants with missing data for any of these covariates were excluded from the study, reducing the statistical power of the results.

Sample Size:

The total number of participants in NHANES 1999–2002 was 21,004, and 8223 have data on TL. Only 3937 participants have both data on estrogen usage and TL. After excluding those without information on the demographic data and other covariates, the final sample size was 491 (Figure 1). After applying the weights, these 491 participants represent 6,610,056 American individuals.

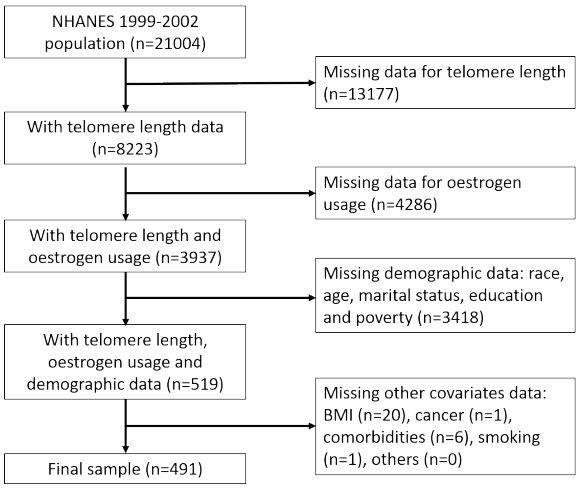

Figure 1: Flow chart of the study population. The NHANES 1999–2002 dataset includes 21,004 participants. Among them, 491 individuals possessed all the essential information for this study, representing 6,610,056 Americans after applying weights.

Statistical Analysis:

R Studio (R version 4.1.2) was used for the data analysis in this study. R Studio is an open-source software implemented with the R programming language for statistical data analysis.[33]

Rao-Scott $\chi 2$ test was used to assess the bivariate association between TL and all other variables. Rao-Scott $\chi 2$ bivariate test is a hypothesis analysis to test whether two nominal variables have a statistically significant relationship. This test suits correlation analysis in complex survey data with sampling weights.[34] A p-value obtained from the test of less than 0.05 is considered statistically significant in this study. A p-value of less than 0.05 indicates less than 5% probability of obtaining the observed results by random chance alone.

In addition to the Rao-Scott $\chi 2$ bivariate test, mosaic diagrams and Pearson residuals analysis were used to examine bivariate associations. A mosaic diagram serves as a visual representation of the relationship between two categorical variables. It employs rectangular tiles to display observed frequencies. Pearson residuals measure the discrepancies between observed and expected frequencies in each tile, which helps to identify statistically significant associations.

Multiple logistic regression analysis was used to calculate the odds ratio (OR) and 95% confidence interval (CI) of TL and estrogen use. Multiple logistic regression analysis is a statistical method that creates a formula to predict the relationship between a single dependent variable (TL) and more than one independent variable. OR measures the likelihood of the outcome occurring under a specific condition. For example, an OR value of 2.2 obtained in this study means that the odds of having a higher TL are 2.2 times higher for the estrogen users than the non-users. The 95% CI was used to measure the certainty of the OR value. For example, the 1.10–4.37 of the 95% CI indicated a 95% chance of certainty that the population's true mean is between the upper (4.37) and lower (1.10) interval range of the OR.

The logistic regression models were constructed using forward stepwise modeling. This implies initiating the model with a single variable, i.e., estrogen usage, and adding other variables individually. A covariate remains in the model only if its inclusion enhances the statistical power; otherwise, it is excluded. This sequential addition of the covariates aids in identifying the key predictors significantly contributing to the outcome variable, i.e., TL. Moreover, this modeling approach yields simpler models with fewer predictors, making interpreting the model more straightforward. This study incorporated the covariates into the model with the following sequence: gender, PIR, race, education, marital status, BMI, smoking, C-reactive protein, history of cancer, number of comorbidities, diabetes, and analgesic usage.

The TL was analyzed as a nominal variable of two categories: above or below the 75th percentile of all the weighted lengths. The range of TL was 0.2 to 2.0, and the 25th, 50th, and 75th percentiles were 0.74, 0.87, and 1.02, respectively. Considering TL as a nominal rather than a continuous variable enabled me to calculate the chances (i.e., the odds ratio) of having TL in the top 25% of the population. Another reason for categorizing TL was to ensure the statistical robustness of the outcomes, as the distribution of TL was not perfectly normal.[35] TL Categorisation can help reduce the impact of outliers and enhance the stability of the associations.

All data in this study were weighted using NHANES measures. NHANES assigned a specific weight to each participant to represent the estimated number of individuals in the U.S.

DOI: 10.36838/v6i10.1

population with similar features. This weighting technique was employed to adjust for oversampling specific groups, including age groups, ethnicities, and socioeconomic strata. Also, during NHANES data collection, not all selected individuals participated in the survey, which may introduce nonresponse bias. Therefore, applying weights allowed the samples of NHANES to reflect the U.S. population more accurately.[36]

Ethical Statement:

The US National Center for Health Statistics (NCHS) Ethics Review Board approved the data collection processes. All participants provided informed consent to participate in the program.

Result and Discussion

Participant Characteristics:

All participants in this study were aged 60 years or above. The majority of them were non-estrogen users, non-Hispanic white, married, with higher education attendance, non-smokers, not regular users of analgesics, and without cancer, diabetes, and major cardiovascular diseases (Table 1). The number of males and females was quite similar, and their PIR was distributed relatively evenly among the three groups.

Table 1: Weighted characteristics of 491 participants by telomere length (above or below the 75th percentile), NHANES 1999–2002. Factors statistically associated with telomere length (p < 0.05) include estrogen usage, gender, and the poverty-income ratio.

Variable	Weighted Participants (%)	Telomere Length[1]		*p*-values[2]
		<75th percentile (%)	>75th percentile (%)	
All	100	75.0	25.0	
Estrogen				0.0065
Yes	24.6	62.2	37.8	
No	75.4	78.9	21.1	
Gender				0.0043
Male	43.9	82.6	17.4	
Female	56.1	69.4	30.6	
Race				0.4588
Non-Hispanic White	81.5	73.4	26.6	
Non-Hispanic Black	5.3	80.9	19.1	
Other	13.2	81.0	19.0	
Education				0.3478
Less than high school	23.7	81.4	18.6	
High school	29.0	70.4	29.6	
More than high school	47.2	74.3	25.7	
Poverty income ratio				0.0061
<2	35.9	86.1	13.9	
2–4	33.7	69.5	30.5	
>4	30.4	67.4	32.6	
Marital Status				0.3683
Married	66.8	72.9	27.1	
Widowed or separated	29.0	77.1	22.9	
Never married	4.20	89.9	10.1	
BMI				0.2688
Underweight	0.6	38.4	61.6	
Normal or healthy weight	30.8	73.3	26.7	
Overweight	41.5	78.7	21.3	
Obese	27.1	71.5	28.5	
Smoking				0.0998
Current smoker	18.4	85.3	14.7	
Previous smoker	35.8	75.2	24.8	
Never smoker	45.8	70.4	29.6	
C-reactive protein				0.2520
1st quartile (0.0–0.12)	25.8	73.3	26.7	
2nd quartile (0.12–0.27)	24.2	74.3	25.7	
3rd quartile (0.27–0.54)	25.7	69.5	30.5	
4th quartile (>0.54)	24.2	82.7	17.3	
History of cancer				0.4325
Yes	17.1	79.8	20.2	
No	82.9	73.8	26.2	
Number of comorbidities[3]				0.6971
0	92.9	75.3	24.7	
1	4.8	62.8	37.2	
2	1.9	76.7	23.3	
>3	0.4	100.0	0.0	
Diabetes				0.5707
Yes	3.0	78.1	21.9	
No	96.2	74.6	25.4	
Borderline	0.8	95.7	4.3	
Analgesic				0.9546
Yes	31.3	75.1	24.9	
No	68.7	74.7	25.3	

[1]Telemere length at 75th percentile is 1.02.

[2]Rao–Scott $\chi 2$ statistics

[3]Comorbidities include congestive heart failure, coronary heart disease, angina, heart attack, and stroke.

Bivariate Association Analysis between TL and Estrogen Usage or Gender:

The variables that demonstrated a statistically significant association ($p < 0.05$) with TL were estrogen usage, gender, and PIR (Table 1). About a quarter of all the participants were taking estrogen, and within them, 37.8% had a longer TL than the 75th percentile. For non-estrogen users, this value was only 21.1%. The rate ratio of having telomeres within the top 75th percentile between the estrogen users and non-users was 1.79. This was calculated by taking the ratio of the estrogen users to non-users, i.e., 37.8% divided by 21.1%. Thus, this cross-sectional study found that estrogen administration is associated with longer TL within all the participants.

Another analysis was performed by subdividing the participants into male and female groups. Figure 2A is a mosaic diagram comparing the TL between the male and female groups. In the mosaic diagram, the size of each tile (rectangle) corresponds to the number of observations in that specific category from the result table. Therefore, the larger the tile, the greater the number of participants in that group. Figure 2A illustrates that the number of participants with TL below the 75th percentile is similar between the male and female groups. In contrast, the female group has a significantly larger number with TL above the 75th percentile. This suggests that females generally have longer telomeres.

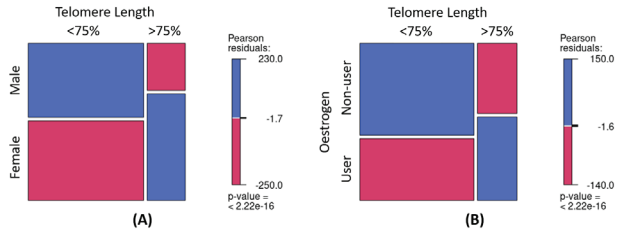

Figure 2: Mosaic diagram comparing the TL between (A) male and female participants and (B) estrogen users and non-users of female-only participants. More female participants had TL above the 75th percentile, with estrogen users showing a higher prevalence compared to non-users.

To explore the impact of estrogen on female participants exclusively, a bivariate association analysis was conducted using the 75th percentile value of female participants as the cutoff point, resulting in a new TL value of 1.056. Among all participants, 56.1% were female, and within this group, 31.8% of estrogen users had TLs longer than the 75th percentile, compared to only 20.9% of non-users. The calculated rate ratio was 1.44. However, this association did not reach statistical significance, as indicated by a p-value of 0.165 from the Rao- Scott $\chi 2$ bivariate test.

A mosaic diagram was created to analyze estrogen's effect on TL further. Figure 2B illustrates a positive deviation towards shorter TL for the non-estrogen users and a positive deviation towards longer TL for the users, suggesting a trend towards higher TL in the users. Pearson residuals confirmed the statistical significance of this difference ($p < 0.001$).

The inconsistency between the Rao-Scott $\chi 2$ bivariate test and Pearson residuals test may be attributed to their methodological differences in modeling association between two or more categorical.[37] The $\chi 2$ test assumes specific distribution-

al properties, but Pearson residuals assess discrepancies from expected counts without the same assumptions. Moreover, unlike the Mosaic diagram, the $\chi 2$ test considers survey sampling weights.[37] Despite these differences, both tests suggest a tendency for higher TL in estrogen users, with one test indicating an almost statistically significant relationship and the other showing statistical significance.

This positive association between estrogen and TL was also found in an animal study investigating the impact of estrogen on TL in a rat model with induced liver fibrosis.[38] The rats administered with estrogen exhibited significantly longer telomeres in hepatocytes than those without estrogen administration. This finding reveals the potential of estrogen in preventing cellular damage caused by liver fibrosis. We understand that the study analyzed the TL of hepatocytes, while our study focused on leukocytes. The former reflects the genomic stability of liver cells, whereas the leukocytes TL represents the overall systemic aging and immune health. Nonetheless, both studies indicate the potential of estrogen in preserving telomeres in both animals and humans.

Bivariate Association Analysis between TL and Covariates:

Among all the covariates, gender and PIR were the only covariates significantly associated with TL. PIR is a measurement of poverty calculated by NHANES, ranging from 0 to 5. A higher value indicates greater family income. This study's results reveal that higher-income participants had longer TLs (Table 1). While no previous studies have shown a direct relationship between PIR and TL, existing studies have established that TL is associated with lifestyle, dietary patterns, and mental states – all of which are also linked with family income, i.e., PIR.[39] Individuals from high-income families tend to engage in more physical activity, adopt healthier diets, and experience a lower risk of mental illness.[39] These factors are positively associated with TL.[39]

Apart from gender and PIR, no other covariates demonstrated a statistically significant association with TL (Table 1). This result may differ from different studies. This inconsistency could be attributed to the variations in sample size, participant characteristics, statistical analysis methods, and measurement techniques for both TL and covariates. For example, the study conducted by Sun *et al.*, 2012 found a negative association between smoking status and TL,[39] which is different from the findings in this study. Sun *et al.*, 2012 included participants of all ages but excluded those with chronic illnesses such as heart disease, stroke, diabetes, and cancer. Their studies did not utilize any sampling weights. In contrast, our study included healthy participants and those with chronic illnesses, limited to individuals aged 60 years or above, and employed the NHANES-provided sampling weights to represent the US population.

Another example study with results that differ from this study was the one performed by Muezzinler *et al.*[40] It suggested a negative association between BMI and TL. Once again, this study did not restrict participant age. Thus, the participant characteristics could vary from those in our study, potentially leading to divergent outcomes. Nonetheless, we believe that the elderly may be less sensitive to factors that could influence TL.

Multiple logistic regression analysis:

The logistic regression models were established for each adjusted covariate for estrogen usage against TL (Table 2). The odds ratios of all the models ranged from the lowest in the gender-adjusted model of 1.81 to the highest in the C-reactive protein-adjusted model of 2.67. Among them, only the gender-adjusted model is statistically insignificant. This suggests a potential existence of multicollinearity in the model. This affects the ability of the model to precisely estimate the individual effects of each covariate, potentially reducing odds ratios and significance.[41] Furthermore, including covariates that do not significantly contribute to explaining the variation in TL may dilute the effects of relevant variables, leading to diminished odds ratios and significance.[42] Many covariates did not exhibit statistical significance in the above $\chi 2$ bivariate association analysis. Consequently, they could not contribute to the multiple logistic regression analysis and thus cannot be utilized to formulate a predictive model of the relationship between TL and other covariates.

Table 2: Weighted associations of individual variables and estrogen use against telomere length (TL) (above the 75th percentile) in 491 participants, NHANES 1999–2002. The C-reactive protein-adjusted model yielded the highest statistically significant ($p < 0.05$) odds ratio among all the variables, followed by diabetes status and analgesic usage.

Variables	Odds ratios (95% confidence intervals) of TL (above the 75th percentile)	
	Covariates	Estrogen (yes vs no)
Race		2.21 (1.29–3.79)
Non-Hispanic White	1.27 (0.62–2.60)	
Non-Hispanic Black	Reference	
Other	0.89 (0.24–3.21)	
Gender		1.81 (0.91–3.59)
Male	Reference	
Female	1.59 (0.89–2.86)	
Education		2.15 (1.21–3.83)
Less than high school	0.64 (0.25–1.62)	
High school	Reference	
More than high school	0.81 (0.44–1.51)	
Poverty income ratio		2.03 (1.15–3.60)
<2	Reference	
2–4	2.65 (1.29–5.40)	
>4	2.57 (1.19–5.55)	
Marital Status		2.22 (1.28–3.84)
Married	1.23 (0.59–2.57)	
Widowed or separated	Reference	
Never married	0.41 (0.08–2.19)	
BMI		2.28 (1.33–3.89)
Underweight	5.16 (0.51–52.1)	
Normal or healthy weight	0.90 (0.54–1.51)	
Overweight	0.71 (0.40–1.27)	
Obese	Reference	
Smoking		2.14 (1.23–3.71)
Current smoker	Reference	
Previous smoker	1.75 (0.71–4.30)	
Never smoker	2.19 (1.07–4.52)	
C-reactive protein		2.67 (1.37–5.20)
1st quartile (0.0–0.12)	2.32 (1.04–5.18)	
2nd quartile (0.12–0.27)	2.21 (0.97–5.01)	
3rd quartile (0.12–0.27)	2.38 (1.11–5.07)	
4th quartile (>0.54)	Reference	
History of cancer		2.29 (1.31–4.03)
Yes	0.69 (0.28–1.67)	
No	Reference	
Number of comorbidities[1]		2.21 (1.29–3.79)
0	Reference	
1	1.53 (0.44–5.41)	
2	0.94 (0.24–3.21)	
>3	<0.01 (NA)[2]	
Diabetes		2.39 (1.37–4.18)
Yes	1.56 (0.09–3.71)	
No	Reference	
Borderline	0.10 (0.01–1.10)	
Analgesic		2.33 (1.37–3.97)
Yes	0.86 (0.42–1.79)	
No	Reference	

[1]Comorbidities include congestive heart failure, coronary heart disease, angina, heart attack, and stroke. [2]NA indicates the values of the 95% confidence intervals are less than 0.0001.

DOI: 10.36838/v6i10.1

Among all the established multiple logistic regression models, the final model was selected based on the highest OR values and their 95% confidence intervals, which were statistically significant (Table 3). An unadjusted model refers to one that includes only the primary predictor variable (estrogen) of interest without controlling for any other covariates. In contrast, an adjusted model includes not only the primary predictor variable but also additional covariates that may influence the outcome variable. Comparing unadjusted and adjusted models aids in understanding the direct effects of predictors on outcomes while addressing the influence of other relevant variables. The unadjusted multiple regression model yielded an OR value of 2.27, signifying that the estrogen users had 2.27 higher odds of having TL in the 75th percentile (Table 3). The adjusted PIR and C-reactive covariates model obtained a slightly higher OR (2.36) than the unadjusted model (Table 3). This indicates a stronger association was observed while controlling for the covariates.

Table 3: Weighted and adjusted odds ratios of estrogen use and covariates against TL (above the 75th percentile) in 491 participants, NHANES 1999–2002. The poverty income ratio and the C-reactive protein-adjusted model exhibited a higher odds ratio than the unadjusted model.

Variable	Odds ratios (95% confidence intervals) of TL (above the 75th percentile)	
	Unadjusted	Poverty income ratio and C-reactive protein-adjusted
Estrogen use (yes vs. no)	2.27 (1.30–3.96)	2.36 (1.18–4.72)
Poverty income ratio		
<2		Reference
2–4		2.54 (1.23–5.27)
>4		2.41 (1.08–5.40)
C-reactive protein		
1st quartile (0.0–0.12)		2.05 (0.93–4.52)
2nd quartile (0.12–0.27)		2.08 (0.91–4.74)
3^{rf} quartile (0.12–0.27)		2.26 (1.07–4.73)
4th quartile (>0.54)		Reference

Limitations:

The data in NHANES are cross-sectional and do not reflect the longitudinal changes in TL over time or the causal relationships between TL and health outcomes. The data are based on a single measurement of TL, which may be influenced by factors such as stress, inflammation, infection, lifestyle, etc. Therefore, the data may not capture the participants' true biological age or cellular senescence. Additionally, the data does not represent the entire US population, as some groups, such as pregnant women, institutionalized persons, and military personnel, were excluded from the survey. Moreover, the sample sizes for some racial/ethnic groups were relatively small, which may limit the statistical power and generalisability of the findings. In addition, this study also includes the common biases of a cross-sectional survey study, i.e., lack of follow-up, control variables, recall bias, and information bias.[42]

Another limitation is the lack of information on estrogen dosage, which is not included in NHANES. This prevented further analysis of the dose-dependent relationship between estrogen and TL. Other variables, such as genetic variation,[43] psychiatric disorders,[44] and oxidative stress level,[45] were found to be associated with TL in previous studies but were not considered covariates here. Again, this is because NHANES does not contain this information.

Conclusion

This study used the US national database to investigate factors influencing telomere length (TL). Our findings revealed that females tend to have longer TL compared to males, providing a potential explanation for their extended longevity. Furthermore, we established a positive association between TL and the usage of estrogen in female participants, with a rate ratio of 1.44. These results highlight the potential of estrogen administration in preserving telomeres and fostering optimal aging. We hypothesize that further controlled trials with more participants are justified to confirm a causal dose-dependent relationship between estrogen and TL. These trials will help validate our findings and provide a more comprehensive understanding of the potential benefits of estrogen in maintaining telomere integrity.

Acknowledgments

I extend my gratitude to my 'Extended Project Qualification' tutor, Ms. Saira Luckman, for her unwavering support and guidance throughout this research project. Additionally, I express my appreciation to my fellow students at St. Michael's School for the invaluable opportunity to engage in collaborative learning and exchange insights on our respective research projects.

References

1. Chakravarti, D.; LaBella, K. A.; DePinho, R. A. Telomeres: History, Health, and Hallmarks of Aging. *Cell* **2021**, *184* (2), 306–322. https://doi.org/10.1016/j.cell.2020.12.028.
2. Lulkiewicz, M.; Bajsert, J.; Kopczynski, P.; Barczak, W.; Rubis, B. Telomere Length: How the Length Makes a Difference. *Mol. Biol. Rep.* **2020**, *47* (9), 7181–7188.
3. Schneider, C. V.; Schneider, K. M.; Teumer, A.; Rudolph, K. L.; Hartmann, D.; Rader, D. J.; Strnad, P. Association of Telomere Length with Risk of Disease and Mortality. *JAMA Intern. Med.* **2022**, *182* (3), 291–300.
4. Dorajoo, R.; Chang, X.; Gurung, R. L.; Li, Z.; Wang, L.; Wang, R.; Beckman, K. B.; Adams-Haduch, J.; Liu, S.; Meah, W. Y. Loci for Human Leukocyte Telomere Length in the Singaporean Chinese Population and Trans-Ethnic Genetic Studies. *Nat. Commun.* **2019**, *10* (1), 2491.
5. Viña, J.; Borrás, C.; Gambini, J.; Sastre, J.; Pallardó, F. V. Why Females Live Longer than Males? Importance of the Upregulation of Longevity-Associated Genes by Estrogenic Compounds. *FEBS Lett.* **2005**, *579* (12), 2541–2545.
6. Zarulli, V.; Barthold Jones, J. A.; Oksuzyan, A.; Lindahl-Jacobsen, R.; Christensen, K.; Vaupel, J. W. Women Live Longer than Men Even during Severe Famines and Epidemics. *Proc. Natl. Acad. Sci.* **2018**, *115* (4), E832–E840.
7. Crimmins, E. M.; Shim, H.; Zhang, Y. S.; Kim, J. K. Differences between Men and Women in Mortality and the Health Dimensions of the Morbidity Process. *Clin. Chem.* **2019**, *65* (1), 135–145.
8. Pinchuk, I.; Weber, D.; Kochlik, B.; Stuetz, W.; Toussaint, O.; Debacq-Chainiaux, F.; Dollé, M. E.; Jansen, E. H.; Gonos, E. S.; Sikora, E. Gender-and Age-Dependencies of Oxidative Stress, as Detected Based on the Steady State Concentrations of Different Biomarkers in the MARK-AGE Study. *Redox Biol.* **2019**, *24*, 101204.

9. Ortona, E.; Pierdominici, M.; Rider, V. Sex Hormones and Gender Differences in Immune Responses. *Front. Immunol*. **2019**, *10*, 1076.
10. Lansdorp, P. M. Sex Differences in Telomere Length, Lifespan, and Embryonic Dyskerin Levels. *Aging Cell* **2022**, *21* (5), e13614.
11. Bayne, S.; Li, H.; Jones, M. E. E.; Pinto, A. R.; van Sinderen, M.; Drummond, A.; Simpson, E. R.; Liu, J.-P. Estrogen Deficiency Reversibly Induces Telomere Shortening in Mouse Granulosa Cells and Ovarian Aging in Vivo. *Protein Cell* **2011**, *2* (4), 333–346.
12. Yamagata, Y.; Nakamura, Y.; Umayahara, K.; Harada, A.; Takayama, H.; Sugino, N.; Kato, H. Changes in Telomerase Activity in Experimentally Induced Atretic Follicles of Immature Rats. *Endocr. J.* **2002**, *49* (6), 589–595.
13. Noppert, G. A.; Feinstein, L.; Dowd, J. B.; Stebbins, R. C.; Zang, E.; Needham, B. L.; Meier, H. C. S.; Simanek, A.; Aiello, A. E. Pathogen Burden and Leukocyte Telomere Length in the United States. *Immun. Ageing* **2020**, *17* (1), 36.
14. Chen, Y.-F.; Zhou, K.-W.; Yang, G.; Chen, C. Association between Lipoproteins and Telomere Length in US Adults: Data from the NHANES 1999–2002. *Lipids Health Dis*. **2019**, *18* (1), 80.
15. Linghui, D.; Shi, Q.; Chi, C.; Xiaolei, L.; Lixing, Z.; Zhiliang, Z.; Birong, D. The Association between Leukocyte Telomere Length and Cognitive Performance among the American Elderly. *Front. Aging Neurosci*. **2020**, *12*, 527658.
16. Tucker, L. A. Fruit and Vegetable Intake and Telomere Length in a Random Sample of 5448 US Adults. *Nutrients* **2021**, *13* (5), 1415.
17. Centers for Disease Control and Prevention, N. C. for H. S. *National Health and Nutrition Examination Survey Data*. National Center for Health Statistics. https://wwwn.cdc.gov/nchs/nhanes (accessed 2023-10-18).
18. Zipf, G.; Chiappa, M.; Porter, K. S.; Ostchega, Y.; Lewis, B. G.; Dostal, J. Health and Nutrition Examination Survey Plan and Operations, 1999-2010. **2013**.
19. Hidalgo, A.; Chilvers, E. R.; Summers, C.; Koenderman, L. The Neutrophil Life Cycle. *Trends Immunol*. **2019**, *40* (7), 584–597.
20. Lynch, S. M.; Peek, M. K.; Mitra, N.; Ravichandran, K.; Branas, C.; Spangler, E.; Zhou, W.; Paskett, E. D.; Gehlert, S.; DeGraffinreid, C. Race, Ethnicity, Psychosocial Factors, and Telomere Length in a Multicenter Setting. *PloS One* **2016**, *11* (1), e0146723.
21. Kingma, E. M.; de Jonge, P.; van der Harst, P.; Ormel, J.; Rosmalen, J. G. The Association between Intelligence and Telomere Length: A Longitudinal Population Based Study. *PloS One* **2012**, *7* (11), e49356.
22. Geronimus, A. T.; Pearson, J. A.; Linnenbringer, E.; Schulz, A. J.; Reyes, A. G.; Epel, E. S.; Lin, J.; Blackburn, E. H. Race-Ethnicity, Poverty, Urban Stressors, and Telomere Length in a Detroit Community-Based Sample. *J. Health Soc. Behav.* **2015**, *56* (2), 199–224.
23. Chen, R.; Zhan, Y.; Pedersen, N.; Fall, K.; Valdimarsdóttir, U. A.; Hägg, S.; Fang, F. Marital Status, Telomere Length and Cardiovascular Disease Risk in a Swedish Prospective Cohort. *Heart* **2019**.
24. Mundstock, E.; Sarria, E. E.; Zatti, H.; Mattos Louzada, F.; Kich Grun, L.; Herbert Jones, M.; Guma, F. T. C. R.; Mazzola In Memoriam, J.; Epifanio, M.; Stein, R. T.; Barbé-Tuana, F. M.; Mattiello, R. Effect of Obesity on Telomere Length: Systematic Review and Meta-Analysis: Effect of Obesity on Telomeres. *Obesity* **2015**, *23* (11), 2165–2174.
25. Astuti, Y.; Wardhana, A.; Watkins, J.; Wulaningsih, W. Cigarette Smoking and Telomere Length: A Systematic Review of 84 Studies and Meta-Analysis. *Environ. Res*. **2017**, *158*, 480–489.
26. Shivappa, N.; Wirth, M. D.; Hurley, T. G.; Hébert, J. R. Association between the Dietary Inflammatory Index (DII) and Telomere Length and C-reactive Protein from the National Health and Nutrition Examination Survey-1999–2002. *Mol. Nutr. Food Res*. **2017**, *61* (4), 1600630.
27. Wentzensen, I. M.; Mirabello, L.; Pfeiffer, R. M.; Savage, S. A. The Association of Telomere Length and Cancer: A Meta-Analysis. *Cancer Epidemiol. Biomarkers Prev*. **2011**, *20* (6), 1238–1250.
28. Cheng, F.; Carroll, L.; Joglekar, M. V.; Januszewski, A. S.; Wong, K. K.; Hardikar, A. A.; Jenkins, A. J.; Ma, R. C. Diabetes, Metabolic Disease, and Telomere Length. *Lancet Diabetes Endocrinol*. **2021**, *9* (2), 117–126.
29. Sibille, K. T.; Langaee, T.; Burkley, B.; Gong, Y.; Glover, T. L.; King, C.; Riley, J. L.; Leeuwenburgh, C.; Staud, R.; Bradley, L. A.; Fillingim, R. B. Chronic Pain, Perceived Stress, and Cellular Aging: An Exploratory Study. *Mol. Pain* **2012**, *8*, 1744-8069-8–12.
30. Boniewska-Bernacka, E.; Pańczyszyn, A.; Klinger, M. Telomeres and Telomerase in Risk Assessment of Cardiovascular Diseases. *Exp. Cell Res.* **2020**, *397* (2), 112361.
31. Centers for Disease Control and Prevention. All About Adult BMI. Centers for Disease Control and Prevention. https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/index.html (accessed 2021-06-21).
32. Fong, P.; Chan, S. T.; Lei, P. N.; Cheong, H. I.; Cheong, I. M.; Hoe, W. L. Association of Suicidal Ideation and Depression with the Use of Proton Pump Inhibitors in Adults: A Cross-Sectional Study. *Sci. Rep*. **2022**, *12* (1), 19539.
33. Kronthaler, F.; Zöllner, S. Describing Data with RStudio. In *Data Analysis with RStudio: An Easygoing Introduction*; Kronthaler, F., Zöllner, S., Eds.; Springer: Berlin, Heidelberg, 2021; pp 35–58.
34. Skinner, C. Analysis of Categorical Data for Complex Surveys. *Int. Stat. Rev*. **2019**, *87* (S1).
35. Nettle, D.; Seeker, L.; Nussey, D.; Froy, H.; Bateson, M. Consequences of Measurement Error in qPCR Telomere Data: A Simulation Study. *PLoS One* ***2019***, *14* (5), e0216118.
36. Centers for Disease Control and Prevention. *Overview of NHANES Survey Design and Weights*. https://www.cdc.gov/Nchs/tutorials/environmental/orientation/sample_design/index.htm (accessed 2020-07-13).
37. Bilder, C. R.; Loughin, T. M. Modeling Association Between Two or More Categorical Variables That Allow for Multiple Category Choices. *Commun. Stat. - Theory Methods* **2007**, *36* (2), 433–451.
38. Sato, R.; Maesawa, C.; Fujisawa, K.; Wada, K.; Oikawa, K.; Takikawa, Y.; Suzuki, K.; Oikawa, H.; Ishikawa, K.; Masuda, T. Prevention of Critical Telomere Shortening by Oestradiol in Human Normal Hepatic Cultured Cells and Carbon Tetrachloride Induced Rat Liver Fibrosis. *Gut* **2004**, *53* (7), 1001–1009.
39. Sun, Q.; Shi, L.; Prescott, J.; Chiuve, S. E.; Hu, F. B.; De Vivo, I.; Stampfer, M. J.; Franks, P. W.; Manson, J. E.; Rexrode, K. M. Healthy Lifestyle and Leukocyte Telomere Length in US Women. *PloS One* **2012**, *7* (5), e38374.
40. Müezzinler, A.; Zaineddin, A. K.; Brenner, H. Body Mass Index and Leukocyte Telomere Length in Adults: A Systematic Review and Meta-analysis. *Obes. Rev*. **2014**, *15* (3), 192–201.
41. Mohammadi, S. A Test of Harmful Multicollinearity: A Generalized Ridge Regression Approach. *Commun. Stat. - Theory Methods* **2022**, *51* (3), 724–743.
42. Solem, R. C. Limitation of a Cross-Sectional Study. *Am. J. Orthod. Dentofacial Orthop*. **2015**, *148* (2), 205.
43. Mirabello, L.; Yu, K.; Kraft, P.; De Vivo, I.; Hunter, D. J.; Prescott, J.; Wong, J. Y. Y.; Chatterjee, N.; Hayes, R. B.; Savage, S. A. The Association of Telomere Length and Genetic Variation in Telomere Biology Genesa. *Hum. Mutat*. **2010**, *31* (9), 1050–1058.
44. Darrow, S. M.; Verhoeven, J. E.; Révész, D.; Lindqvist, D.; Penninx, B. W.; Delucchi, K. L.; Wolkowitz, O. M.; Mathews, C. A. The Association between Psychiatric Disorders and Telomere Length: A Meta-Analysis Involving 14,827 Persons. *Psychosom. Med*. **2016**, *78* (7), 776–787.

DOI: 10.36838/v6i10.1

45. Ahmed, W.; Lingner, J. Impact of Oxidative Stress on Telomere Biology. *Differentiation* **2018**, *99*, 21–27.

Authors

Kayleigh Fong, an A-level student at St. Michael's School in Llanelli, UK, is passionate about medical knowledge and its potential to make a difference in the world. She aims to continue learning and applying her knowledge practically to contribute to the advancement of healthcare and improve the lives of others.