Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

[Farah Srichandra] Date

Micronutrient biomarkers of Nutrition Status Influence the Risk of TB in US Men and Women: NHANES I Epidemiologic Follow-up Study

By

Farah Srichandra

Degree to be awarded: Master of Public Health

Global Health

_____ [Chair's signature]
[Peter Cegielski, MD MPH]
Committee Co - Chair

_____ [Chair's signature]
[Usha Ramakrishnan, PhD]
Committee Co - Chair

3

Micronutrient biomarkers of Nutrition Status Influence the Risk of TB in US Men and Women: NHANES I Epidemiologic Follow-up Study

By

Farah Srichandra

B.S., Loyola Marymount University, 2009

Thesis Committee Co - Chairs: Peter Cegielski, MD MPH and Usha Ramakrishnan PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Health
2011

4

Abstract

Micronutrient biomarkers of Nutrition Status Influence the Risk of TB in US Men and Women: NHANES I Epidemiologic Follow-up Study

By Farah Srichandra

Background: One third of the world's population is thought to be infected with tuberculosis (TB), and new infections occur at a rate of about one per second. The proportion of people who become sick with TB each year is stable or falling in different regions of the world, but, because of population growth, the absolute number of new cases is still increasing. Undernutrition is an important risk factor for TB, but, primarily, studies have been conducted on anthropometric measures for nutritional status and few studies on micronutrient biomarkers for nutritional status.

Objective: This research analyzes the risk of developing TB in relation to micronutrient biomarkers of nutritional status. Our specific interest is whether micronutrient biomarkers of nutritional status at baseline are related to the incidence of TB during the follow-up of 10 to 20 years.

Methods: In the NHANES-1 epidemiological follow-up study,a nationally representative data set of 14,407 adults who were followed up on 3 to 4 occasions over a period of 10 to 20 years. Secondary data analysis was carried out using SAS statistical software, examining the relationship between deficiencies in hemoglobin, albumin, vitamins A, iron and the incidence of TB and then compare with adjusted. The primary dependent variable will be the incidence of TB during the follow - up period.

Results: Anemia and low serum albumin are independent risk factors for physician diagnosed TB, and it's magnitude of effect is 3.02 (1.44, 6.34).As a result, the chances of developing TB increase by 3-fold for persons who are anemic.

Discussion: The follow-up study provides data on a large national sample, it presents a unique opportunity for health researchers to study changes in health status and the factors that contribute to good health as well as illness. It was the first U.S. investigation of its size and scope to follow respondents over a period of years.

5

Micronutrient biomarkers of Nutrition Status Influence the Risk of TB in US Men and Women: NHANES I Epidemiologic Follow-up Study

By

Farah Srichandra

B.S., Loyola Marymount University, 2009

Thesis Committee Co - Chairs: Peter Cegielski, MD MPH and Usha Ramakrishnan PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Health
2011

6

Acknowledgements

I would not have been able to complete this work without the encouragement and support of many people. First I would like to thank Dr. Usha Ramakrishnan and Dr. Peter Cegielski, my thesis advisors, for their support and understanding over the past year. I would like to thank my advisors for their helpful and constructive feedback on the written document. I would not have been able to do all my work without statistical assistance from Dr. Pat Kilgo, who was always willing to make time to discuss SAS with me. NHANES-1 and NHEFS was a tremendous and continual source of information which allowed me database access. Many thanks to my friends and family who gave me unconditional encouragement, love, and support from near and far.

Thank you all!

Table of Contents

I. INTRODUCTION

One third of the world's populations is thought to be infected with tuberculosis (TB), and new infections occur at a rate of about one per second. The proportion of people who become sick with TB each year is stable or falling in different regions of the world, but, because of population growth, the absolute number of new cases is still increasing. Studies have shown that malnutrition is an important risk factor for the development of TB. Malnutrition profoundly affects cell-mediated immunity, which is the principle host defense against TB. Nutrient deficiencies in vitamins A, C, D, protein, and iron are thought to have an impact on cell mediated immunity [1, 2]. Therefore, a decrease in these micronutrients will proportionally decrease cell mediated immunity and subsequently increase the incidence rate of TB.

Among infectious diseases, tuberculosis is the single leading cause of death worldwide. Each year, Mycobacterium tuberculosis causes 8 million new cases and 3.6 million deaths, over 90% of them in developing countries [1, 2]. Mycobacterium tuberculosis is a very successful pathogen that can survive and persist in the human host in the face of a robust immune response. This immune response is sufficient to prevent disease in the majority of infected persons, providing compelling evidence that immunity to tuberculosis is possible. However, it is more striking that the strong immune response is not generally effective at eliminating the organisms, during either initial infection or the persistent or latent phase of infection. Studies in animal models and in humans have demonstrated the wide range of immune components involved in the effective response against M. tuberculosis. These immune responses are directed towards containing or eliminating the tubercle bacillus

9

within the tissues of the host. The estimated eight million new cases of tuberculosis each year clearly demonstrate that these responses are not always effective. M. tuberculosis has obviously evolved a variety of mechanisms to evade destruction by the immune response [3, 4, and 5].

In Cegielski et. al examination of previous studies in humans and animals concluded malnutrition is an important risk factor for the development of TB. [11].

However, laboratory data are available on vitamin A levels, iron indices, and serum albumin levels in NHANES-1have not yet been analyzed. In the past, studies have been conducted on anthropometric measures of nutritional status and TB in NHANES-1 but not on micronutrient biomarkers for nutritional status. Micronutrient biomarkers of nutritional status may indicate an increased risk of TB.

Two hypotheses form the basis of this work:

1. Anemia (low hemoglobin) contributes significantly to the risk of tuberculosis disease in the USA.

2. Low serum albumin and serum iron in the population at risk for TB contributes are predictors of incidence of TB in the USA.

10

II. SPECIFIC AIMS

The following specific aims are designed to test these hypotheses among American adults using data from the National Health and Nutrition Examination Survey-1 (NHANES1) and the NHANES-1 Epidemiologic Follow-up Study (NHEFS).

1. Anemia (low hemoglobin) contributes significantly to the risk of tuberculosis disease in the USA.
2. Low serum albumin and serum iron in the population are risk factors for the incidence of TB in the USA.

Our specific interest is whether micronutrient biomarkers of nutritional status at baseline are related to the incidence of TB during follow-up over a period of 10 to 20 years. In the past, studies have been conducted on anthropometric measures of nutritional status and TB in NHANES-1 but not on micronutrient biomarkers for nutritional status. In addition, previous studies which examined malnutrition status and TB have been cross-sectional in design. In a cross-sectional study, the disease and exposure are determined within a snapshot and it cannot be concluded whether malnutrition status is the cause of TB or vice-versa.

This research analyzes the risk of developing TB in relation to micronutrient biomarkers of nutritional status. The unique strength of longitudinal cohort studies is the ability to measure nutritional status prior to the onset of TB. Lower dietary intakes of key nutrients may have been associated with higher rates of hospitalization rather than (or in addition

to) higher rates of TB. Micronutrient biomarkers of nutritional status may indicate an increased risk of TB.

III. LITERATURE REVIEW

A. Epidemiology of Tuberculosis

The highest prevalence of tuberculosis infection and estimated annual risk of tuberculosis infection are in sub-Saharan Africa and Southeast Asia [4]. Overall, almost 3.8 million cases of tuberculosis were reported in the world in 1990, of which 49% were in Southeast Asia. From the 3 year period 1984 through 1986 to the 3 year period 1989 through 1991, notification rates increased in all World Health Organization regions, except the American and the European regions. In 1990, there were an estimated 7.5 million cases of tuberculosis and 2.5 million deaths worldwide. The human immunodeficiency virus epidemic is causing increases in the number of tuberculosis cases, particularly in Africa. In eastern Europe and the former Soviet Union, drug resistance is a serious problem [6].

TB is transmitted by the airborne route. After being inhaled, once tubercle bacilli reach the alveoli, host defense mechanisms come into play. Bacilli which survive initial phagocytosis by alveolar macrophages reproduce to sufficient numbers that within 1 to 3 months, specific cell-mediated immune responses develop. in 90% of infected persons, these immune responses are sufficient to keep the infection in check locally and keep it from progressing to clinically apparent disease [3]. The bacilli, however, may remain viable in a quiescent state for decades. Of the remaining 10%, about 1/2 (5% of the total), TB will develop within 2 years. In the remaining 5%, TB will develop at some later time

12

in life either through reactivation of the latent infection or exogenous acquisition of new bacilli. If age, illness, malnutrition or other circumstances diminish the strength of the cellular immune system, then viable endogenous bacilli may proliferate until they produce disease [6,7].

Risk factors which have been identified that increase the probability of progressing from primary or latent infection to active disease include: Extremes of age, male sex, genetic susceptibility, immunosuppression due to HIV infection, alcoholism, abnormal chest radiogram showing fibrotic lesions, diabetes mellitus, and being 10% or more below ideal body weight [8].

B. Recent Studies on TB

Recent findings suggest that micronutrient deficiencies, e.g., vitamin A, may affect resistance to TB by altering the function of the respiratory mucosa and the integrity of pulmonary epithelial tissues [16, 17, 18].

Iron plays a critical role in support of cell mediated immunity, however, iron is also critical to the replication of mycobacteria and other pathogens. Thus, the impact of iron on susceptibility to TB is difficult to predict.

Cegielski examined the relationship between undernutrition, as determined in the first National Health and Nutrition Examination Survey (NHANES-1), and TB incidence as ascertained in the NHANES-1 Epidemiological Follow-up Study [2,3,4]. Having body

13

mass index (BMI), average skin-fold thickness, or mid-upper arm cross-sectional muscle area in the lowest decile of the population increased the adjusted hazard of TB from 6- to 10-fold, controlling for other known risk factors for TB. That study, however, did not analyze hemoglobin, albumin, vitamin A, or iron levels.

A unique study of the effect of micronutrient supplementation on TB incidence was reported by Downes in 1949. In a controlled trial among the families of black TB patients in the Harlem ghetto of New York City, families were allocated alternately to receive vitamin and mineral supplements versus no supplements. After 5 years of follow-up, the risk of TB in the control group was 2.8-fold higher than the supplemented group . Compared with those who actually took the vitamin supplements throughout the observation period, the risk of TB in the control group was 5.9-fold higher. As a result, vitamin supplementation substantially reduced the risk of TB among family contacts of active TB cases. [2,3,4].

In the 1940s, Getz et.al enrolled a cohort of 1100 men who were free of TB at baseline, and followed them for up to 5 years with serial clinical, radiographic, and laboratory examinations. Plasma vitamin A levels were low in 13 of 16 men (81%) who developed TB compared to 318 of 1058 (30%) of those who did not.

14

IV. SUBJECTS AND METHODS

A. The first National Health and Nutrition Examination Survey (NHANES-1)

The NHANES-1 was a survey of demographic, socioeconomic, health, and nutrition information based on a representative sample of the U.S. population carried out from 1971-75. The sample was a complex, multistage, stratified, probability sample of the civilian, non-institutionalized population, aged 1-74 years, residing in the 48 coterminous states. Primary sampling units consisted of individual counties or small clusters of contiguous counties. Through successive stages of sampling, a total of 26,322 persons were included in NHANES-1.

The NHANES-1 had 2 distinct components: The main body of the survey was carried out in 65 primary sampling units (PSUs) from 1971 to 1973. An augmentation component, consisting of an additional 35 PSUs, was carried out 1974-75, resulting in 100 total PSUs. Each of these components was nationally representative by itself with its own sample weights and sampling design variables. The present analysis is based on the unweighted data.

The design included oversampling certain population subgroups in the 65 PSU main survey in order to allow more stable estimates of population parameters. The groups which were oversampled included persons living in poverty areas, women of childbearing age (25-44 years), and elderly persons (over 65 years). There was no oversampling in the 35 PSU augmentation survey.

15

At the baseline NHANES-1, participants had an interview, a medical examination, and laboratory tests. Interview data were collected on socioeconomic and demographic variables, dietary intake, and medical history. The medical examination included a systematic physical examination and nutritional anthropometry. Included among the numerous blood and urine laboratory tests, were biochemical markers of protein, iron, vitamin A, thiamine and riboflavin status. Not all data were collected on all participants, however, due to the distinct subsamples.

B. The NHANES-1 Epidemiologic Follow-up Study (NHEFS)

The NHANES-1 cross-sectional data formed the baseline for a large-scale longitudinal follow-up study (NHEFS) which was designed to examine the relationship of baseline clinical, nutritional, and behavioral factors to later morbidity and mortality. Adults from the original cross-sectional survey have been traced and re-interviewed in 4 successive waves carried out from 1982 to 1992.

In the first wave, 1982-84, the NCHS and the National Institute on Aging and other Institutes of the National Institutes of Health attempted to trace and collect data on all 14,407 participants in the NHANES-1 who had been medically examined at baseline and were 25-74 years old at the time. Due to errors in the ascertainment of age at the time of NHANES1, however, the corrected age range of the NHEFS cohort was actually 24-77 years of age at the time of the NHANES-1 examination. For the 1982-84 wave, data collection included personal interviews with those traced (or proxies for those incapacitated or deceased), hospital and nursing home records, and death certificates for

16

decedents. Of the original 14,407 eligible persons, 13,383 (93%) had been successfully traced. Interviews with the subject or a proxy were actually conducted for 91.3% (12,220) of those successfully traced, or 84.8% of the entire original cohort. Of the 11,361 surviving subjects interviews were obtained for 92.3% (10,523), 256 by proxy for incapacitated subjects. Of the 2022 decedents, death certificates were obtained for 95.7% (n=1935) and proxy interviews for 83.9% (n=1697) [14, 16 (p. 3-4)].

The second wave in 1986 was restricted to the 5677 individuals who were 55 years and older at the time of their NHANES-1 examination. For this wave of follow-up, interviewing was conducted primarily by telephone. Of the 3980 subjects not known to be deceased at the time of the 1982-84 wave, 94.6% were traced successfully.

The third and fourth wave in 1987 and 1992 included the total cohort of 14,407 participants, and interviewing was carried out primarily by telephone.

C. The present analytic data set

1. Inclusion/exclusion criteria. The present analytic data set consists of a subset of the entire NHEFS cohort (N=14,407 for 100 PSU) (N=11,348 for 65 PSU) selected with one exclusion criterion. Persons with a history of TB prior to NHANES-1 were excluded (N=218 for 100 PSU) (N=180 for 65 PSU). NHANES-1 participants were asked specifically whether they had ever had TB and whether a doctor had ever told them they had TB. As a result, the 100 PSU and 65 PSU analysis sample consisted of 14,189 and 11,168, respectively.

17

2. Exposure variables. The main exposure of interest was nutritional status as measured in the NHANES-1 survey. Nutritional status was measured by three distinct methods: Anthropometry, biomarkers, and nutrient intake. Only laboratory measure variables were analyzed in the present paper including *blood hemoglobin, serum protein, serum albumin, vitamin A, and serum iron*. The main exposure variables used in this analysis were anemia and serum albumin. Hemoglobin was available from all 100 PSUs (all subsets) of NHANES-1, while serum albumin and vitamin A was available from the 65 PSU main survey.

The criterion for anemia in men was that if hemoglobin levels were less than 130 mg/L, then anemia was considered present. The criterion for anemia in women was if hemoglobin less than 120 mg/L then anemia was considered present.

3. Outcome variables. The outcome of interest was the incident case of tuberculosis as determined from 3 sources: interviews, medical records, and death certificates. Participants in the follow-up surveys were specifically asked details about all overnight health care facility stays, including the reasons for those stays. Up to 4 reasons were recorded in the NHEFS data (as text fields) for each facility stay. These were subsequently categorized by the NCHS staff into 37 condition codes. All of the condition codes which could have included TB and their corresponding text fields were manually searched to find any self-reported occurrences of TB.

18

All health care facilities identified through the interviews were contacted to obtain medical records for all admissions falling within the scope of the survey period (i.e. after NHANES-1). Facilities identified through any of the medical records so obtained were also contacted. Hospital discharge diagnoses and nursing home diagnoses (ICD-9-CM codes) were obtained from these medical record summaries. These included not only those illnesses which occurred at the time of the hospital or nursing home stay, but also any illnesses which were noted in the medical history to have occurred prior to the facility stay and which were coded among discharge diagnoses. In other words, conditions which were recorded in the medical record as "history of" were also captured by the data abstraction process. Death certificates were coded using multiple cause of death coding.

4. Follow-up time. The duration of follow-up was the number of days from each participant's NHANES-1 examination up to the last date at which they were at risk for developing TB. For persons who developed TB, follow-up ended on the date associated with the diagnosis of TB. For persons who were identified as TB cases only on their death certificate, follow-up ended on the date of death. For persons who were identified as TB cases through health care facility records, follow-up ended on the discharge date of the facility stay when the diagnosis was first established. For persons who reported a facility stay for TB but for whom medical records were not received at NCHS, follow-up ended on the reported date of hospitalization.

Follow-up time for persons who did not develop TB was determined in a corresponding manner. For those for whom a death certificate was obtained, follow-up ended on the date

19

of death. For survivors for whom medical records were obtained, follow-up ended on the last discharge date. For survivors who reported a facility stay but for whom medical records were not obtained, follow-up ended on the reported date of the stay.

5. Covariates. The initial exploratory analysis included socioeconomic, demographic, and health variables from the NHANES-1 baseline data which are known to be risk factors for tuberculosis or associated with malnutrition. These included *sex, age, and race*. *Sex* (male, female) and *race* (white, non-white) were dichotomous variables. For the purposes of logistic regression analysis, the linearity of the logit transformation of continuous variables was tested using the SAS macro, EMPTREND. These variables satisfied the linearity assumption after plotting the continuous variables (*serum albumin and age*).

D. Analysis

The analysis consisted of a traditional epidemiologic analysis for a historical cohort study. The analysis was based on the unweighted data from NHANES-1 and NHEFS. Each analysis was conducted in several stages. First, univariate distributions of each variable in the analysis were examined for shape, missing values, and outliers. Univariate analysis was performed using one-way frequency distributions for each of the variables. The center and spread of micronutrient biomarkers, *serum protein, serum albumin, serum iron, hemoglobin, iron binding capacity, and vitamin A* are reported in Table 2. through Table 8.

20

Second, bivariate associations were examined in a cross tabulation of main outcome of TB with the main exposure, anemia, was performed to determine if an association between the main exposure and main outcome (TB status) exists.

To identify confounders and covariates related to the outcome, two-way comparisons of the main outcome of interest (TB status) with covariates were conducted using chi-square test for categorical variables (*sex and race*) and t-tests for continuous variables (*age, serum protein, serum albumin, hemoglobin, and vitamin A).* Known and potential confounders are as follows: *age, gender, and race*. Potential risk factors are *serum albumin, low hemoglobin (anemia), low vitamin A, and iron.* The chi-square-test assumptions that all cells contain more than five subjects as well as the t-test assumption that the sample size is sufficiently large were both met. If the cells contain less than 5 subjects, the Fisher exact p-value was calculated (Tables 12 through 17) The number and percentage of each covariate was determined for participants who developed TB over time of follow up and those who did not. (Tables 9 through 17) P-values were calculated to determine whether there was a statistically significant difference between participants who developed TB compared to those who did not develop TB.

Next, two way comparisons of covariates with our main exposure variable (anemia status) were screened for possible association (Table 9). Covariates which were not associated with either anemia status or TB status were dropped from further analysis. This part of the analysis was repeated for albumin as the exposure variable.

21

Logistic regression was carried out based on covariates associated with TB and anemia. The initial model included all covariates of significant confounding potential. A Wald-chi-squared statistic was used to determine the significance of interaction terms to the model. Interaction was not addressed subsequently by creating new interaction variables: *interact1 = anemia*sex; interact2 anemia*age;interact3 = anemia*serum albumin;interact4 = anemia*race2; interact5 = anemia*serum protein.*

The collinearity among the risk factors were further evaluated based on the variation inflation factors (VIF) with VIF>10 being the criteria for the presence of multicollinearity.

The odds ratios along with p-values were then assessed to statistically evaluate whether or not anemia is an independent risk factor for the main outcome (TB) after adjusting for other covariates and if so, the magnitude of its effect on the main outcome, TB incidence (Table 17).

Modeling proceeded with a stepwise procedure described above. A gold standard model was chosen that retained all potential covariates. In this approach the variable found to be least significant to the model, as determined by the Wald chi square p-value, was removed. If the odds ratio changed significantly upon removal of the variable, it was necessary to control for that confounder by retaining it in the model. However, if removal of the variable had no effect on anemia-Tb status relationship then it was not considered a confounder and could be removed from the model. This process of individually removing the least significant variable was continued until all remaining variables had a significant

22

contribution to the logistic model, indicated by a Wald chi-square p-value of less than 0.05. The final model also included sex and race as confounding variables even though they had p-values>0.05 due to the gender bias of TB in men and predominantly dichotomous race (White versus Black).

V. RESULTS

A. Description of the Study Sample.

The total NHEFS cohort included all 14,407 NHANES-1 participants (from all 100 PSUs) who were aged 25 to 74 years at the time of their NHANES-1 examination. Of this total number, 218 of these had a prior history of TB. The study sample consists of the remaining 14,189 persons (Table 1). Sociodemographic characteristics of the study population at baseline are detailed in Table 1. These unweighted figures reflect the oversampling of women of child - bearing age, the elderly, and persons living in poverty areas. Table 1 reports that the mean age of participants is 48.89 years old. The majority are of white race (83.5%) and female (59.7%). TB incidence is higher in men than women and in blacks than whites. On the other hand, TB incidence increases substantially with increasing age.

B. Bivariate Analysis

In tables 2 through 8, the descriptive statistics with center and spread for continuous variables *(age, serum protein, serum albumin, serum iron, hemoglobin, iron binding capacity, and vitamin A)* are displayed from the study. Age and serum albumin demonstrated significant differences between the TB status groups (P-value <0.05) [Table 2,3]. Next, identified covariates related to TB status from two-way comparisons using chi-

23

square test for categorical variables include sex, low albumin status, and anemia (Tables 4 through 8). Additional covariates from two-way comparisons using t-tests for continuous variables include age and serum albumin (Table 9).

The effect of anemia was not confounded by other covariates for the main outcome (TB). Overall, none of the following variables were identified as confounders: *serum protein, serum albumin, serum iron, hemoglobin, iron binding capacity, low transferring satuaration, and vitamin A* (Tables 10 through 17).

C. Logistic Regression and Modeling

Based on the bivariate analyses above, the covariates associated with the main exposures TB and anemia status (age, anemia, and serum albumin) were carried forward into multivariable analysis. The initial model included all covariates of significant confounding potential including race and sex. A Wald-chi-squared statistic was used to determine the significance of covariates to the model. Baseline logistic regression characteristics includes beta, standard error beta, Wald chi –square statistic, 95% Wald confidence interval, and p-value as shown at the completion of (Table 21).

In addition to the primary outcome variable (TB) and any potential confounders identified (race and sex), the models also contained other covariates considered significant in Table 18 (age and serum albumin).Upon completion of the tests for interaction, no significant interaction between the main exposure (anemia) and the covariates as none were below the pre determined critical p-value of 0.05. In conclusion, none of the

24

variables which interacted with the anemia and TB relationship were statistically significant under investigation (Tables 22, 23).

Ultimately, after testing for interaction our model remains the same: TB = anemia, serum albumin, sex, race, and age. None of the interaction terms were statistically significant. Anemia is an independent predictor of physician diagnosed TB, it's magnitude of effect is 3.02 (1.44, 6.34).As a result, increasing chances of developing TB by 3-fold if anemic. Odds ratio is significant because 95% CI does not contain 1 (Tables 21, 22). There is sufficient statistical evidence to conclude that there is an association between the variables included in the final model because the confidence intervals do not contain 1) [Table 18]. There was no presence of multicollinearity that existed between the covariates based on the variation inflation factors (VIF) with none having a VIF>10.

At the 5% significance level, there is sufficient evidence to conclude that *anemia* independently increases risk of developing TB after adjusting for all other variables in the model (Wald p-value: 0.0004); the odds of developing TB are 3.02 times that of those who were not anemic (Tables 22, 23).

VI. DISCUSSION

Our study shows that anemia status is an independent risk factor for developing TB. Overall, anemia, serum albumin, male sex, and non-white race independently increased risk of developing TB. The large sample size and the relatively good numbers of known and potential confounders taken into consideration are the strength of this study. However,

25

due to over sampling of persons living in high - poverty areas, women of childbearing age (25-44 years), and elderly persons (over 65 years), the findings of the study may not be extrapolated beyond the unique study population.

A. Limitations

This study has several methodological limitations. First, it was not possible to adjust for immigration from developing countries, despite the importance of these factors in TB incidence in the U.S. today, because there were too few people in these categories in the NHEFS cohort.

Second, an important question is the extent to which nutritional status measured at one point in time, 1971-75, impacts the incidence of TB occurring many years later. This is an important issue, but there are partial responses to it. Most of the TB cases in the study cohort were identified in the 1982 wave of follow-up. If nutritional status at the NHANES-1 was highly correlated with nutritional status at the time the TB exposure or reactivation took place then the predictive value of nutrition for TB will be largely maintained. This would tend to minimize the effect of any non-differential secular trends in the population.

A third limitation of this analysis is the method by which incident cases of TB were detected, since it was necessary to extract this information from information pertaining to hospitalizations and deaths. The major concern is that these cases may be differentially distributed between nutritional status groups. A detection bias may have contributed

26

substantially to the magnitude of the effect. In addition, the relationship between loss to follow-up and baseline nutritional status was not examined.

Although not affecting the internal validity of this study, persons who were homeless, who were incarcerated, or in other institutions were excluded from the NHANES-1 survey. Similarly, Alaska, Hawaii, and Indian Reservation lands were excluded. Since these include some of the highest risk groups for TB, these exclusions limit the ability to generalize these findings to the entire U.S. adult population. Similarly, the cohort was a closed cohort and does not represent the current population of adults in the U.S.

B. Strengths

The strengths of this study are first, that it is the first nationally representative study of nutrition and TB in the United States. The sample design allows these results to be generalized to the target population of U.S. adults.

The second strength has to do with the primary data on which this analysis was based. The National Center for Health Statistics, the National Institute on Aging, and the other agencies that were involved in NHANES-1 and NHEFS went to extraordinary lengths to insure the reliability, internal, and external validity of the data.

The basic study design, an individual-level historical cohort study, is a third strength. It enabled ascertainment of nutritional status and the occurrence of TB in the same individuals in correct temporal sequence. This lack has been the major flaw in prior

27

human research on this subject. It also enabled the questions to be addressed with vastly less expense than to assemble a large prospective cohort study or randomized trial designed specifically to address the nutrition-TB question.

C. Future Research

The next step is to repeat this analysis taking full advantage of the sampling weights and sample design variables. In this manner it will be possible to obtain relative risk estimates with confidence intervals that can be inferred to the U.S target population.

The second step will be to analyze the other measures of nutritional status that were obtained in the NHANES-1 including nutrient intake data. Dietary intake data is available as a 24-hour dietary recall and a limited food frequency questionnaire. This will enable us to assess if dietary indicators of nutritional status corroborate and expand, or contravene, the results obtained in this study. The third step would be to incorporate additional data from the NHEFS along with analyzing NHANES-2, NHANES-3 while comparing across the dataset.

28

D. Importance for Public Health

The importance of this issue for the changing epidemiology of TB depends on the ability to intervene and decrease TB morbidity and mortality with improved nutrition. If the nutritional status of certain groups at high risk for TB in the U.S. deteriorates and results in increasing TB incidence (and possibly other morbidity as well), this finding would have significant clinical, public health, and policy implications, especially in light of current trends in federal funding of nutrition assistance programs. Further investigation with diversity in race and immigration may also help improve the efficacy of the study and its applicability to a greater population in the future.

29

VII. LITERATURE CITED

1. McMurray DN (1998) Impact of nutritional deficiencies on resistance to experimental pulmonary tuberculosis. Nutrition Review 56: S147–S152.

2. Cegielski JP, McMurray DN. The Relationship between Malnutrition and Tuberculosis: Evidence from Studies in Humans and Experimental Animals. Internat J Tuberc Lung Dis 2004;8:286-298.

3. Cegielski, J.P., D.P. Chin, M.A. Espinal, T.R Frieden, R.R. Cruz, E.A. Talbot, D.E.C. Weil, R. Zaleskis & M.C. Raviglione: The global tuberculosis situation: progress and problems in the 20th century, prospects for the 21st century. Infect *Dis Clin North Am* 16, 1-58 (2002

4. Wells CD, Cegielski JP, Nelson L, et al. HIV infection and multidrug resistant tuberculosis---the perfect storm. J Infect Dis 2007;196(Suppl 1):S86--107.

5. McMurray DN, Cegielski JP. The influence of nutrition on the risk and outcomes of tuberculosis. ch. 11 in Academy of Sciences of South Africa Consensus Panel on Nutrition, HIV/AIDS, and TB (ed.). HIV/AIDS, TB, and Nutrition: Scientific inquiry into the nutritional influences on human immunity with special reference to HIV infection and active TB in South Africa. Pretoria, South Africa: Academy of Sciences of South Africa, 2007.

6. Flynn JAL. Immunology of tuberculosis and implications in vaccine development. Elsevier 2004;84:105564.

7. Styblo K. Overview and epidemiologic assessment of the current global tuberculosis situation with an emphasis on control in developing countries. Rev Infect Dis 1989;11(suppl 2):S33946.

8. Comstock G. Epidemiology of tuberculosis. Am Rev Respir Dis 1982;125:815.

9. Centers for Disease Control. Tuberculosis morbidityUnited States, 1992. MMWR 1993;42:696704.

10. Frieden TR, Sterling T, Pablos, Mendez A, et al. The emergence of drug resistant tuberculosis in New York city. N Engl J Med 1993;328:5216.

11. Raviglione MC, et al. Global Epidemiology of Tuberculosis

12. Morbidity and Mortality of a Worldwide Epidemic. JAMA. 1995;273:220-226.

13. Bonilla DL, Fan YY, Chapkin RS, McMurray DN. Transgenic mice enriched in omega-3 fatty acids are more susceptible to pulmonary tuberculosis: Impaired resistance to tuberculosis in *fat-1* mice. J Infect Dis 2010; 201: 399-408.

14. Cooper AM. Cell-mediated immune responses in tuberculosis. Annu Rev Immunol 2009; 27: 393-422.

31

15. Garton NJ, Waddell SJ, Sherratt AL, Lee S-M, Smith RJ, et al. Cytological and transcript analyses reveal fat and lazy persister like bacilli in tuberculous sputum. PLoS Med. 2008;5(4):e75. doi:10.1371/journal.pmed.0050075.

16. Leung CC, Lam TH, Chan WM, Yew WW, Ho KS, Leung G, Law WS, Tam CM, Chan CK, Chang KC. Lower risk of tuberculosis in obesity. Arch Intern Med. 2007 Jun 25;167(12):1297-304.

17. Liu PT, Modlin RL. Human macrophage host defense against *Mycobacterium tuberculosis.* Curr Opin Immunol 2008; 20: 371-376.

18. Lonnroth K, Williams BG, Cegielski P, Dye C. A consistent log-linear relationship between tuberculosis incidence and body mass index. Internat J Epidemiol 2010;39:149-155.

19. Cegielski JP, McMurray DN. Tuberculosis: Nutrition and Susceptibility Ch.268 in Enclyclopedia of Human Nutrition, 3rd ed, 2012 Elsevier (in press).

Appendix A.

Table 1. Sociodemographic Characteristics, NHANES -1 AND NHEFS

Characteristic	Number	Percent
Total NHEFS	14,407	100%
TB within 6 months to NHANES-1	- 218	-1.5 (98.5%)
Analysis Cohort	14,189	100%
Sex		
Male	5,811	40.3%
Female	8,596	59.7%
Race		
Non-White	2,371	16.5%
White	12,036	83.5%
Age (Mean)	48.9 years	

Table 2. The Center and Spread of Age in TB Status

Age	n	mean	Sd	Low, High	P-Value
TB					
Yes	61	57.41	14.11	25, 75	<0.0001
No	14218	48.87	15.59	24,77	
Total N= 14,279					

33

Table 3. The Center and Spread of Serum Protein in TB Status

Serum Protein	n	mean	Sd	Low, High	P-Value
TB					
Yes	53	7.21	1.01	5.30, 13.30	0.1620
No	11115	7.12	0.51	4.70, 11.50	
Total N= 11,168					

Table 4. The Center and Spread of Serum Albumin in TB Status

Serum Albumin	n	mean	Sd	Low, High	P-Value
TB					
Yes	53	4.22	0.38	2.80, 4.80	0.0025
No	11115	4.36	0.33	2.70, 5.70	
Total N= 11,168					

Table 5. The Center and Spread of Serum Iron in TB Status

Serum Iron	n	mean	Sd	Low, High	P-Value
TB					
Yes	51	96.02	40.57	33.00, 253	0.3472
No	10488	100.9	37.22	17.00, 396	
Total N= 10,539					

Table 6. The Center and Spread of Hemoglobin in TB Status

Hemoglobin	n	mean	Sd	Low, High	P-Value
TB					
Yes	59	140.8	15.10	110.0, 173.0	0.1365
No	13561	143.7	15.06	50.00, 224.0	
Total N= 13,620					

Table 7. The Center and Spread of Iron Binding Capacity in TB Status

Iron Binding Capacity	n	mean	Sd	Low, High	P-Value
TB					
Yes	51	335.1	73	175.0, 543.0	0.0010
No	10488	362.8	0.51	112.0, 717.0	
Total N= 10,539					

Table 8. The Center and Spread of Vitamin A in TB Status

Vitamin A	n	mean	Sd	Low, High	P-Value
TB					
Yes	51	61.12	26.63	25.0, 180.0	0.3541
No	10738	58.71	18.43	9.0, 279.0	
Total N= 10,789					

Table 9: Covariates, by TB Status (Continuous Variables)

	Overall mean (SD)	**TB** mean (SD)		**P-Value**
		Yes	No	
Risk Factors (continuous)				
Age	48.89 (15.59)	57.41 (14.11)	48.86 (15.59)	<.0001
Serum Protein	7.11(0.51)	7.21 (1.01)	7.11 (0.51)	0.1620
Serum Albumin	4.35 (0.33)	4.22 (0.38)	4.36 (0.33)	0.0025
Hemoglobin	143.73(15.05)	140.8 (15.1)	143.7 (15.0)	0.1365
Folate	80.84(62.12)	107(54.18)	80.7 (62.13)	0.2330
Vitamin A	58.73(18.47)	61.12(26.63)	58.71(18.43)	0.3541

Table 10. The Association of Sex and TB Status

	TB	**No TB**	**P-Value**
Sex			
Male	34 (0.59%)	5690(99.41%)	0.0140
Female	27(0.32%)	8438(99.68%)	
Total N= 14189	61 (0.43%)	14128(99.57%)	

Table 11. The Association of Race and TB Status

	TB	**No TB**	**P-Value**
Race			
Non-White	15(0.64%)	2318 (99.36%)	0.0853
White	46(0.39%)	11810(99.61%)	
Total N= 14189	61 (0.43%)	14128(99.57%)	

36

Table 12. The Association of Anemia and TB Status

	TB	No TB	P-Value
Anemia			
Yes	8 (1.19%)	664(98.81%)	0.0022
No	51 (0.39%)	12897(99.61%)	
Total N= 13620	59 (0.43%)	13561(99.57%)	

Table 13. The Association of Low Transferrin Saturation and TB Status

	TB	No TB	P-Value
Low Transferrin Saturation			
Normal > or = 200	41 (0.50%)	8205(99.50%)	0.7092
Low <200	10(0.44%)	2283 (99.56%)	
Total N= 10,539	51 (0.48%)	10488(99.52%)	

Table 14. The Association of Low Albumin and TB Status

	TB	No TB	Fisher's Exact P-Value
Low Albumin			
Normal > 3.5	50 (0.45%)	10963(99.55%)	0.0371
Low < or = 3.5	3(1.94%)	152(98.06%)	
Total N= 11,168	53 (0.47%)	11115(99.53%)	

37

Table 15. The Association of Low Serum Iron and TB Status

	TB	No TB	Fisher's Exact Test
Low Serum Iron			
Normal > or = 50	48(0.48%)	9956 (99.52%)	0.2241
Low <50	3(0.56%)	532(99.44%)	
Total N= 10,539	51 (0.48%)	10488(99.52%)	

Table 16. The Association of High Iron Binding Capacity and TB Status

	TB	No TB	Fisher's Exact P-Value
High Iron Binding Capacity			
Normal	42 (0.51%)	8135(99.49%)	0.1021
High > 400	9(0.38%)	2353(99.62%)	
Total N= 10,539	51 (0.48%)	10488(99.52%)	

Table 17. The Association of Iron Deficiency and TB Status

	TB	No TB	Fisher's Exact Test
Iron Deficiency			
Normal	51(0.45%)	10253 (99.55%)	0.3157
Abnormal	0(1.46%)	235(98.54%)	
Total N= 10,539	51 (0.48%)	10488(99.52%)	

38

Table 18: Covariates, by Anemia Status (Continuous Variables)

	Overall mean (SD)	**Anemia** mean (SD)		**P-Value**
		Yes	No	
Risk Factors (continuous)				
Age	48.89 (15.59)	49.03 (16.08)	48.88 (15.54)	0.8168
Serum Protein	7.11 (0.51)	7.14 (0.69)	7.11 (0.50)	0.937
Serum Albumin	4.35 (0.33)	4.15 (0.39)	4.37(0.33)	<0.001
Hemoglobin	143.73 (15.05)	113.3 (9.86)	145.3 (13.52)	<.0001
Vitamin A	58.73(18.47)	51.22 (19.03)	59.05 (18.32)	<.0001

Table 19. The Association of Sex and Anemia Status

	TB	**No TB**	**P-Value**
Sex			
Male	165 (2.99%)	5351(97.01%)	<0.001
Female	507(6.26%)	7597(93.74%)	
Total N= 14189	672(4.93%)	12948(95.07%)	

39

Table 20. The Association of Race and Anemia Status

	TB	No TB	P-Value
Race			
Non-White	289(13.31%)	1882 (86.69%)	<0.001
White	383(3.35%)	11066(96.65%)	
Total N= 14189	672 (4.93%)	12948(95.07%)	

Table 21. Logistic Regression Baseline Characteristics

Covariables	Beta	Std. Error Beta	Wald Chi-Square	95% Wald CI	P-Value
Serum Protein	0.703	0.2478	2.01	(0.43,1.14)	0.1558
Serum Albumin	3.217	0.3820	9.35	(1.52,6.80)	0.0022
Serum Iron	1.004	0.0040	0.8846	(0.99, 1.01)	0.3469
Hemoglobin	1.013	0.0085	2.23	(0.99,1.03)	0.1356
Iron Binding Capacity		0.00271	11.12	(0.99, 1.01)	0.0009
Vitamin A	0.994	0.007	0.8674	(0.98,1.01)	0.3517

40

Table 22. Magnitude of Effect for Variables Included in Final Model

Final Model Variables	Odds Ratio	Confidence Interval	P-value
ANEMIA	**3.02**	**(1.44, 6.34)**	**0.0004**
SEX	**1.86**	**(1.12, 3.08)**	**0.0140**
RACE	**1.66**	**(0.93, 2.96)**	**0.0853**
SERUM ALBUMIN	**3.22**	**(1.52, 6.80)**	**<0.001**
AGE	**0.97**	**(0.95, 0.98)**	**<0.001**

Table 23. The Various Logistic Models tested in comparison to the Gold Standard Model, NHANES.

VARIABLES INCLUDED IN MODEL
TB2009 = ANEMIA, SALBNH, SEX, RACE, AGE