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:

_____ _____

Meredith Laree Kanago Date

The Relationship Between Anemia and Biomarkers of Inflammation (CRP and AGP) in Women of Papua New Guinea

By

Meredith Kanago

Master of Science in Public Health

Department of Epidemiology

Approved:

_____

Faculty Advisor (printed)

_____

Faculty Advisor (signature)

_____

Date

The Relationship Between Anemia and Biomarkers of Inflammation (CRP and AGP) in Women of Papua New Guinea

By

Meredith Kanago

B.S., Pacific Lutheran University, 2009

Faculty Advisor: Kevin Sullivan, PhD, MPH, MHA

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 Science in Public Health
in Epidemiology

2011

Abstract

The Relationship Between Anemia and Biomarkers of Inflammation (CRP and AGP) in Women of Papua New Guinea

By Meredith Kanago

Background: Anemia is a major public health problem worldwide, but its burden is greatest where risk factors such as poor nutrition and low socioeconomic status are common. Because biomarkers of inflammation have been shown to affect certain micronutrient measures, including hemoglobin, the primary indicator of anemia status, it is important to understand this relationship in order to properly assess micronutrient status in populations.

Objective: This study seeks to evaluate any possible association between anemia status and the acute-phase protein biomarkers C-reactive protein (CRP) and α-(1) - acid glycoprotein (AGP) and anemia in non-pregnant women aged 18-49 who participated in the 2005 Papua New Guinea National Micronutrient Survey.

Methods*:* The 2005 Papua New Guinea National Micronutrient Survey was a stratified PPS survey with a 2-stage cluster design carried out from May to October 2005. Logistic models were used to analyze data on 662 women to assess the relationship between anemia and elevated CRP and between anemia and elevated AGP.

Results*:* The overall weighted prevalence of anemia in this population was 34.9%. The weighted prevalence of elevated CRP was 10.43%, and the prevalence of elevated AGP was 7.96%. Controlling for region and recent birth, anemia was significantly associated with elevated CRP, with an odds ratio of 2.74 (95% CI: 1.23, 6.15) among those in rural areas. There did not appear to be a similar association among those in urban locations. In addition, after controlling for region and urban/rural location, anemia was significantly associated with elevated AGP, with an odds ratio of 3.98 (95% CI: 1.54, 10.26).

Conclusions: This study found that there is a clear association between anemia status and elevated levels of acute phase proteins in women in Papua New Guinea, suggesting that the presence of infection could have an effect on assessment of anemia in a population. This finding underscores the importance of collecting information on inflammatory biomarkers in nutritional surveys. Future studies should further investigate the geographic factors involved in this association, including the interaction between elevated CRP and urban/rural location.

The Relationship Between Anemia and Biomarkers of Inflammation (CRP and AGP) in Women of Papua New Guinea

By

Meredith Kanago

B.S., Pacific Lutheran University, 2009

Faculty Advisor: Kevin Sullivan, PhD, MPH, MHA

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 Science in Public Health
in Epidemiology

2011

Acknowledgements

I would like to offer my sincere gratitude to my faculty advisor, Dr. Kevin Sullivan, for his constant guidance and support throughout this project. Thank you also to Katie Tripp for sharing this data set and offering technical assistance.

To my family and friends: your unwavering love, patience, and support throughout this process and over the years have been invaluable, and for this I am sincerely thankful.

TABLE OF CONTENTS

LIST OF TABLES

1. Introduction

Anemia is a significant public health issue globally, affecting approximately 2 billion people worldwide, but its burden is greatest where risk factors such as nutritional deficiency and infection are common (1). In the developing world, factors such as inadequate dietary intake of iron, increased bodily iron demand, and infection status appear to be affected by other factors such as geographic location and socioeconomic status (1). The clinical manifestation of anemia, or low blood hemoglobin (<12 g/dL for non-pregnant women, according to WHO), can range from mild asymptomatic cases to more severe conditions; effects can range from poor pregnancy outcomes and reduced cognitive function to economic impacts such as reduced work capacity (2).

Studies have shown that the assessment of micronutrient status in populations can be affected by the presence of inflammation in individuals. Serum concentrations of the so-called "acute phase proteins," particularly C-reactive protein (CRP) and α-(1) - acid glycoprotein (AGP), are known to increase in the event of inflammation (generally a result of infection or trauma) (3). It has been shown that increased levels of these protein may have an effect on measures of micronutrient status, including indicators for vitamin A, zinc, iron, and hemoglobin status (4). Therefore, in order to accurately assess anemia in a population, particularly in developing countries where the burden of infection may be high, it is important to understand the specific relationship between anemia and the biomarkers of inflammation, CRP and AGP.

The Papua New Guinea National Micronutrient Survey was carried out between May and October 2005, with the purpose of assessing overall micronutrient status of

several target groups in Papua New Guinea. Using data from the survey, this study aims to investigate any association between anemia status (indicated by blood hemoglobin measurement) and two acute phase proteins, α-(1)- acid glycoprotein (AGP) and C-reactive protein (CRP) in non-pregnant Papua-New Guinea women aged 18-49 years who participated in the survey.

2. Background

2.1 Anemia Overview

Anemia is a significant public health problem worldwide, affecting people in all countries of the developing world and industrialized nations alike (2, 5, 6). Defined by the World Health Organization as a condition in which the hemoglobin level in the body is lower than normal, resulting in reduced oxygen-carrying capacity of red blood cells, anemia can have severe physiological consequences (2). Hemoglobin cut-off values vary by age, sex, and pregnancy status.

In industrialized countries, an estimated 10.3% of women (aged 18-59) are anemic, while approximately 42.5% of women in non-industrialized countries are affected by anemia. The burden of anemia is high for children in the developing world, with 42% of children less than five years of age and 53% of children 5-14 years of age affected (2).

The causes of anemia are multifactorial, and specific etiologies depend on the setting. For example, most cases of anemia in developed countries result from iron deficiency, but in the developing world, other causes are more common (7). These important factors include other nutritional deficiencies, low socioeconomic status, trauma

resulting in blood loss, childbirth, menstrual losses, genetic conditions, and infection (6, 8). Few data are available on the prevalence of specific anemia etiologies, however, likely because the available indicators do not provide adequate information on their own to provide this information (5).

Iron deficiency occurs in three distinct stages in the body. In the first stage, the body's iron storage is depleted, and as a result, serum ferritin decreases. In the second stage, decreased red blood cell synthesis occurs as a result of diminished iron supply to the erythroid bone marrow, though hemoglobin concentration remains above cut-off levels. In the third and final stage of iron deficiency, hemoglobin concentration drops below the defined cut-off levels for "anemia"(8).

Anemia can have severe medical and social consequences on both an individual and population level, including reduced ability to carry out the activities of daily living and reduced cognitive function (2). Women are of particular concern, because pre-pregnancy anemia is a risk factor for iron deficiency as a pregnancy develops, a state which has been associated with increased maternal and child morbidity and mortality (7). Economically, reduced work productivity in adults is a serious concern, and decreased learning ability in children has been shown as a consequence of anemia (5, 8).To assess anemia status in populations, blood hemoglobin is considered to be a reliable indicator, especially compared to more subjective clinical measures, in addition to being relatively easy and inexpensive to measure (5). Because measurements can vary based on factors such as elevation, smoking status, and ethnicity, correction factors have been developed to account for these variations (6, 9).

2.2 Acute Phase Response

The acute phase response is characterized by changes in the concentration of plasma proteins (known as the acute-phase proteins), and accompanied by larger physiological and biochemical changes. The acute phase proteins are those whose plasma concentration increase or decrease by at least 25 percent during inflammatory disorders, and are referred to as either positive or negative proteins according to the direction of concentration change (i.e., increase or decrease). Many conditions, such as trauma, infection, heatstroke, childbirth, and various immunologically-driven inflammatory conditions can lead to substantial changes in acute-phase protein concentrations (3, 10).

The positive acute-phase proteins alpha-1-acid glycoprotein (AGP) and C-reactive protein (CRP) are of interest here, as their serum concentrations both individually and in combination can be used as indicators of immune response status.

AGP is an acute-phase protein whose plasma concentrations increase in response to injury or infection as part of the inflammatory response. This increase has been correlated with increased protein synthesis in hepatic cells. AGP gene expression is regulated by several cytokines, including interleukin-1 beta, tumor necrosis factor-alpha, and interleukin-6. While the specific inflammatory function is unclear, its ability to bind to steroid hormones and acid drugs is likely an important factor in its physiological role (11). AGP concentrations have been shown to increase 24 hours after the onset of inflammation, and remain at detectable levels for weeks after initial infection; concentrations of AGP have also been shown to be elevated in low grade chronic inflammation (10).

Like AGP, C-reactive protein (CRP) plasma concentrations increase rapidly following inflammatory stimulus. The primary function of CRP is recognition of foreign pathogens and phospholipid constituents of damaged cells by the binding of phosphocholine. These binding abilities, in conjunction with its ability to bind to phagocytic cells, suggest that CRP is a key player in initiating the elimination of targeted cells. In addition, CRP has been implicated in a variety of other inflammatory interactions, suggesting that it plays a role in numerous physiological processes involved in the inflammatory response (3). CRP has been shown to be particularly sensitive to bacterial infections, and plasma concentrations rise within 10 hours of acute inflammation, followed by rapid normalization within one week. Slight elevation of CRP can also be observed during chronic inflammation (10).

It has been shown that categorizing CRP and AGP into four stages of infection category can be useful in assessing the link between micronutrient status and inflammation (4, 12). For example, Thurman et al. found that assessment of plasma retinol concentrations gave results which more closely matched the biologically expected results when four categories of inflammation were used (defined as follows: “No inflammation,“ or no elevation in CRP or AGP values; “incubation, ” or elevated CRP and normal AGP values; “early convalescence, ” or elevated values of both CRP and AGP; “late convalescence” or elevated AGP and normal CRP values) than when a two-group analysis was conducted (4).

2.3 Acute Phase Response and Micronutrient Status

Because the presence of inflammation is known to have an effect on various biomarkers, acute phase proteins such as AGP and CRP are often measured in nutritional studies; for this reason, there is great interest in understanding how these proteins affect micronutrient measures. As such, a number of studies have investigated the link between serum concentration of acute phase proteins and biomarkers for vitamin A status, iron, zinc, and others.

In a double-blind placebo-controlled randomized trial, Wieringa et al. found that elevated CRP and AGP, both independently and in combination, had a significant effect on indicators of iron, vitamin A and zinc, in infants who were randomly supplemented with these nutrients (13). Thurnham et al. conducted a meta-analysis which examined the relationship between serum retinol, an indicator of vitamin A status, which is known to be reduced and therefore overestimate vitamin A plasma concentration in the presence of infection. The authors found that in those with elevated CRP and AGP, retinol levels were indeed reduced, and as such recommend correcting retinol measurements in nutritional studies to account for inflammation effects (4). Little research has been conducted on the specific relationships between the acute phase proteins and inflammation in women. Christian et al. found an inverse association between AGP and CRP and serum retinol concentrations in a group of pregnant women in Nepal (14), but few if any studies have been conducted to examine similar relationships in non-pregnant women.

Clearly, there is much potential for misclassification of micronutrient status when the effects of inflammation are not taken into account (4, 13, 15). While it has been well-

established that the relationships between the acute phase proteins and micronutrient indicators exist, the consequences of these relationships may not be well-understood enough on a population level to be predictable; further, without correction, comparisons of micronutrient deficiencies across populations with differing infection prevalence may not be accurate (13). A number of studies have proposed methods for dealing with these issues. One method would involve excluding from nutritional studies those people with elevated CRP and AGP levels, but this method could reduce sample size and introduce bias (16). Thurnham et al., on the other hand, propose using a correction method to account for the presence of inflammation, as indicated by elevated CRP, AGP, or combined CRP and AGP levels (4, 17).

2.4 Papua New Guinea

Papua New Guinea (PNG) is a developing nation with a population of approximately 6.1 million located in the Southwestern Pacific Ocean (18). This island nation is generally divided into four regions: the Southern region, Highlands region, Momase region, and Islands region, and most of the population lives in rural areas (87%). As only 3% of roads are paved, travel to remote villages can be difficult. Though the official languages of PNG are English, Pidgin, and Motu, more than 800 distinct local languages are spoken in the country. PNG is known for its vast cultural diversity, as each province has distinct sociocultural features (18).

The leading causes of morbidity and mortality in Papua New Guinea are infectious diseases such as malaria, tuberculosis, HIV, and diarrheal diseases, and the average life expectancy at birth is 61 years for males and 64 years for females (18). The

2005 Papua New Guinea National Micronutrient Survey (19) is the first national nutrition survey to be conducted in Papua New Guinea. It was carried out from May to October 2005 by the Department of Health Papua New Guinea, UNICEF PNG, and the University of PNG, with technical support and partial funding by the U.S. Centers for Disease Control and Prevention (CDC).

3. Methods

3.1 Null Hypothesis

There is no association between: 1) anemia status and C-reactive protein (CRP); 2) between anemia status and α-1 acid glycoprotein (AGP); or 3) between anemia status and combinations of CRP and AGP, in non-pregnant women 18-49 years of age in Papua New Guinea.

3.2 Study Design (19)

A stratified 2-stage cluster design with probability proportional to population size (PPS) was used for this study. The sample was stratified on the four main regions of PNG: Momase, Islands, Highlands, and Southern Region.

In the first stage of sampling, 25 primary sampling units (PSUs) were selected using PPS for each region from a list of all census units in PNG. No census unit was selected more than once, and if a census unit that was selected had fewer than 25 households, then the next nearest census unit was combined with the original census unit. PSUs were located in all 20 provinces of PNG, and of the 100 selected, data were collected from 97 PSUs, as 3 of the clusters were inaccessible.

In the second stage of sampling, 20 households were randomly selected from each PSU from a list of all households created by the survey team aided by local leaders. For PSUs containing more than 250 households, maps of the area were drawn and segmented to select 20 households, and segments selected using probability proportional to size, Households were selected using simple or systematic sampling. A household was defined as a group of people who share a common cooking pot and who share household resources such as food and bedding; families who lived in the same room but ate from separately prepared pots were considered to make up different households. Household members were not necessarily related by blood or marriage.

At each household visited, every eligible person was asked to participate in the survey. Though this included other target groups, only non-pregnant women aged 18-49 years will be considered for this analysis. Each eligible woman was assessed for anemia status, blood levels of retinol binding protein, transferrin receptor, CRP, and AGP, urinary iodine levels, height, and weight. Further, each woman was asked about night blindness, tobacco use, last pregnancy, and questions about that child, such as birth weight. Of these variables, this analysis will consider anemia status, acute phase proteins (CRP and AGP), BMI, tobacco use, pregnancy history, household size, and location (urban vs. rural and region). The complete questionnaire can be found in Appendix I.

Sample size calculations for anemia in women aged 18-49 years of age were based on an estimated anemia prevalence of 50%, a precision of ±10%, and a design effect of 2 for each stratum. The resulting sample size was 193 women per region, or a total of 768 women. Additionally, an individual non-response of 20%, household non-response of 10%, and a proportion of eligible women in each household (1.37

women/household), were considered to obtain a final sample size of at least 779 households.

3.3 Data Collection

Data collection took place from May to October of 2005, and was carried out by 6 survey teams, each consisting of 4 members: a team leader, an interviewer/anthropometry assistant, an anthropometrist, and a laboratory technician. Each team included at least one male and one female member in order to ensure the security of the female team members. Where necessary, local people were hired to assist the teams with locating and accessing the PSUs and with troubleshooting in the field.

Team members were chosen after an interview process conducted by the PNG Department of Health and UNICEF. Most survey members selected were Department of Health staff or university students, and many had experience working in healthcare settings or with NGOs. Survey teams were trained for two-weeks in survey methodology, field procedures, selection of households and eligible participants, interview techniques, questionnaire administration, anthropometry, and the collection, storage, and transport of blood samples. Laboratory technicians in each team were trained by a CDC laboratory specialist on capillary blood collection, field analysis of hemoglobin using the HemoCue system, processing dried blood spots, and collection, storage, and transport of urine and stool samples. At the completion of the course, a 2-day pilot survey was conducted in the capital city, Port Moresby.

The survey was approved by a human subjects review board coordinated by the Department of Health PNG and by U.S. CDC. At each household, permission to proceed

with the interview was obtained from the head of the household. Informed consent was obtained verbally from each participant or from the primary caregiver of each child participant.

Questionnaires were used for the interviews (Appendix I), and interviews were conducted in Pidgin where possible; A local translator was used where Pidgin could not be used for the interview. In addition to interview questions, anthropometric measurements and blood and urine samples were collected where appropriate.

Weights and heights of non-pregnant women aged 18-49 years were taken. Height was measured to the nearest 0.1 cm using a Shorr height board with adult extension added. Weight was measured to the nearest 0.1 kg using UNICEF Seca Uniscales. Individuals with disabilities that prevented them from standing up straight or lying down flat, were wearing casts or heavy bandages or were missing one or more limbs were excluded. Women's ages (in years) were based on self-report.

Capillary blood samples were collected into microtainers via finger puncture to the middle or ring finger. Approximately 250-500µl of blood was collected for each participant. Microtainers were inverted ten times once filled, and the blood was then used to assess hemoglobin status and process dried blood spot cards.

Hemoglobin (Hb) levels were measured using the HemoCue system (HemoCue AB, Angelholm, Sweden). Quality control of each HemoCue instrument was performed every morning of the data collection.

Dried blood spots (DBS) were prepared on specially-designed filter paper. Blood remaining in the microtainers after hemoglobin testing was transferred to pre-printed

circles on the filter paper. The DBS cards were transferred to cardboard drying racks, where they were left to dry for the remainder of the day. Every evening, after ensuring that the blood was completely dry, the cards were packed in gas permeable bags, along with desiccant packs and humidity indicator cards. The cards were then sent to Port Moresby, where they were packed in dry ice and shipped to CDC laboratories. From CDC, they were sent to Juergen Erhardt's laboratory in Jakarta, Indonesia, where they were analyzed for CRP and AGP.

Collected data were entered into a CSPRO 3.1 computer database. In order to minimize errors in data entry, the following precautions were embedded into the data entry screens: minimum and maximum allowable values, specified numbers of digits, and skip patterns. All data were entered twice by trained students from the University of PNG. The two data files were then electronically compared to identify data entry errors. Data from the laboratory tests were single-entered by each laboratory into Microsoft Excel spreadsheets. Those data were then cleaned and merged with the questionnaire data by individual or household ID number.

3.4 Analysis

SAS v. 9.1 for Windows (SAS Institute, Cary, NC, USA) was used for this analysis; all weighted analyses were carried using SAS's PROC SURVEYFREQ and PROC SURVEYLOGISTIC programs to account for the complex survey design. Sampling weights were used in the analysis, and weights were attached to each region according to population size. Statistical significance was defined as $\alpha = 0.05$ throughout the analysis.

Data on 850 women were collected in the 2005 PNG National Micronutrient Survey. The final sample size, after excluding those women who were 15, over the age of 49, those who were pregnant, and those who were lacking data on anemia status, and those for whom data on AGP or CRP level was missing, was 662 women. Hemoglobin measurements were adjusted for altitude and tobacco use based on CDC guidelines (20). Anemia was defined as an adjusted hemoglobin measurement below 12g/dl. CRP and AGP variables were dichotomized by defining a CRP level greater than 5mg/L and an AGP measurement greater than 1.2mg/L as elevated, as specified by the laboratory method used. Further, CRP and AGP data were combined to create 4 categories (stages) of infection: normal levels of both CRP and AGP was defined as "no inflammation; elevated CRP alone was defined as "incubation"; elevated CRP and AGP was defined as "early convalescence"; and elevated AGP alone was defined as "late convalescence." Age, a continuous variable, was categorized into 7 groups, and household size was categorized into 3 levels.

Basic descriptive statistics were obtained for the population of interest (non-pregnant women aged 18-49 years), in terms of both weighted and unweighted frequencies. In addition, tables of anemia prevalence and crude prevalence odds ratios with respect to all covariates were created, and similar statistics were computed for CRP status and AGP status.

3.5 Logistic Regression Modeling

Logistic regression models were built to assess: 1) the relationship between anemia status and elevated CRP status; 2) the relationship between anemia status and

elevated AGP status; and 3) the relationship between anemia status and four categories of combined CRP and AGP status (representing four stages of infection: no inflammation, incubation, early convalescence, and late convalescence).

Potential collinearity between study variables was assessed by obtaining and examining condition indices and variance decomposition proportions (VDPs) from the inverse of the information matrix. A condition index >30 was considered indicative of a possible collinearity issue, in which case corresponding VDPs were examined. If two or more variables had corresponding VDPs>0.5, those variables were considered to display collinearity. Collinearity assessment was conducted without accounting for the complex sample design. The SAS macro "COLLIN" from SAS-L by Matthew Zack was used to calculate collinearity diagnostics from variance-covariance matrix in non-linear regression models (21).

Next, first-order interaction between the exposure variables of interest (CRP, AGP, or infection category) and possible predictor variables was assessed by testing the significance of the interaction terms individually, followed by backward elimination of terms from a model containing only those interaction terms found to be individually significant plus all other covariates. For each covariate, a model was created containing the exposure variable of interest, the first-order interaction between the exposure of interest and the covariate of interest, and the individual covariate. Statistical significance was determined by both the Wald and likelihood ratio p-values for the interaction terms. Next, a model was constructed which contained the exposure variable of interest, all covariates, and all first-order interaction terms found to be significant in the previously described process. Interaction terms were then eliminated sequentially on the basis of

least significance, as determined by both the Wald and likelihood ratio test p-values for the interaction terms. This process was repeated until only significant interaction terms remained in the model, along with exposure of interest and all possible predictors, and this model was defined as the gold standard model.

Confounding was assessed by comparing adjusted odds ratios obtained from models in which potential confounders had been dropped sequentially in order of least significance to odds ratios obtained from corresponding gold standard models. In cases when eliminating a covariate resulted in a change in the adjusted odds ratio of greater than 10% from that of the gold standard model, confounding was suspected and that covariate was left in the model. For each model, precision was calculated, and the relative gain or loss in precision resulting from the elimination of a potential confounder was considered in the selection of each final model as well. Details of the model-building process are included in Appendices II and III.

4. Results

4.1 Descriptive Statistics

After excluding ineligible subjects, 662 women remained in the study, comprising 77.9% of the women originally surveyed. Weighted and unweighted descriptive statistics for this study population are presented in Table 1.

The majority of women in the study population were in the 20-24 or 25-29 age groups, with 20.9% and 22.1% falling in these groups, respectively. The majority of

women lived in the Southern region (31.6%) – however, the weighted regional distribution was somewhat different, with 40.1% of women in the Highlands region, followed by 27.2% in the Momase region. A large majority of women surveyed, 74.9%, lived in a rural setting. Roughly half of the study population (51.8%) lived in households of 9 or more people, and 23.8% reported tobacco usage. With regards to pregnancy history, 56.1% reported having given birth in the past 3 years. Most women in the study had normal BMI (68.30%) and the most prevalent level of educational attainment was grade 4-8 (45.36%).

4.2 Prevalence of Anemia

The prevalence of anemia, both overall and with respect to all covariates in the analysis, is presented in Table 2. The overall weighted prevalence of anemia was 34.9%, which according to WHO guidelines, is indicative of a moderate public health problem among the women in the target population (6). In women with elevated CRP, the prevalence of anemia was 44.0%, compared with 33.9% of those without elevated CRP, though the corresponding crude prevalence odds ratio of 1.5 (95% CI: 0.9, 2.6) was not statistically significant at $\alpha = 0.05$. However, prevalence of anemia was significantly higher in individuals with elevated CRP compared with those who had normal CRP levels, with a crude prevalence odds ratio of 2.4 (95% CI: 1.3, 4.7); this result was statistically significant at $\alpha = 0.05$.

The group of individuals whose combined CRP and AGP status indicated that their stage of infection was “no inflammation” exhibited a lower prevalence of anemia than those in the three other infection categories. The only statistically significant crude

prevalence odds ratio for infection category was obtained for those in early convalescence (elevated AGP and CRP), for whom the prevalence odds of anemia was 2.5 times greater than that for the referent group, individuals with no inflammation (95% CI: 1.0, 6.1).

Region and urban/rural location appeared to be significant independent predicators of anemia status. The Momase region had the highest anemia prevalence, at 60.5% with a crude prevalence odds ratio of 15.3 (95% CI: 7.1, 32.9) compared to the referent group, the Highlands region, in which only 9.1% of those surveyed were anemic. The prevalence of anemia in rural locations was 38.9%, with a statistically significant prevalence odds ratio of 2.6 (95% CI: 2.0, 5.4). Further, women who had given birth in the past 3 years had a slightly higher prevalence of anemia than those who had not, with a prevalence odds ratio of 1.6 (95% CI: 1.1, 2.3), and those whose BMI was classified as "underweight" were much more likely to be anemic than those who were obese, with an odds ratio of 5.2 (95% CI: 1.7, 15.5). Household size, smoking status, and education were not statistically significant independent predictors in this crude analysis.

4.3 Prevalence of Elevated CRP and Elevated AGP

Tables 3 and 4 present the prevalence of elevated CRP and AGP, respectively, with respect to anemia status and other covariates. The prevalence of elevated CRP among anemic individuals in the study population was 13.1% (POR=1.5), and the prevalence of elevated AGP among anemic individuals was 12.5% (POR=2.4). The prevalence of elevated AGP among those with elevated CRP was 52.3%, and 39.9% of

those with elevated AGP had elevated CRP (POR=15), suggesting that these two measures were highly correlated.

In this crude analysis, the only statistically significant independent predictor of elevated CRP was BMI, as those in the "normal" range were significantly less likely than those in the referent group to have elevated CRP, with a POR of 0.4 (95% CI: 0.1, 0.7). None of the covariates examined were statistically independent predictors of elevated AGP.

4.4 Association Between Anemia and Elevated CRP

The collinearity assessment revealed a highest condition index value of 41.85. Examining corresponding VDPs revealed a possible collinearity problem, between the interaction terms CRP*URBANRURAL (VDP=0.87) and CRP*REGION (VDP=0.55). However, this was not an issue in the analysis as the latter interaction term was not significant and was dropped from the final model.

After individually assessing all possible first order interaction terms between CRP and covariates, the covariates with statistically significant interaction were smoking status (p=0.029) and urban/rural location (p=0.031). These terms were placed into a full model with all other covariates in order to conduct backward elimination; however, both of the interaction terms were still highly significant in the full model (p=0.0031 and p=0.0057, respectively), so based on a data-based approach, neither was eliminated. Because an interaction between smoking status and CRP was difficult to explain biologically, this term was left out of the final model, and as such was not included in the gold standard model for assessment of confounding.

Covariates which remained in the model after confounding assessment were region and birth in the past three years, along with the urban/rural interaction term and corresponding lower-order term. Thus the final model was as follows:

$$\text{ANEMIA} = \beta_0 + \beta_1\text{CRP} + \beta_2\text{REGION} + \beta_3\text{BIRTH3} + \beta_4\text{URBANRURAL} + \beta_5(\text{CRP*URBANRURAL})$$

Details of the model-building process are included in Appendix II. Parameter estimates, standard errors, and p-values for the terms in the final model appear in Table 5.

Adjusted odds ratios with respect to urban/rural location and controlling for region and recent births are presented in Table 5a. The odds ratio for the group of subjects who lived in a rural location and had elevated CRP was 2.74 (95% CI: 1.23, 6.15) compared to the referent group of those who lived in a rural location and had normal CRP levels. The adjusted odds ratio for those in an urban location with elevated CRP was 0.15 (95% CI: 0.05, 0.45) and for those in an urban setting with normal CRP levels, the odds ratio was 0.35 (95% CI: 0.19, 0.63).

Observed anemia prevalence with respect to urban/rural location is presented in table 8b. For individuals with elevated CRP, the prevalence of anemia was 52.80% among those in rural locations compared to 10.01% among those in urban areas. Among those with normal CRP, the difference was not so extreme, with 37.30% anemia prevalence in rural areas compared to 20.58% in urban areas.

4.5 Association Between Anemia and Elevated AGP

The collinearity assessment revealed a highest condition index value of 13.64, indicating no problems with multicollinearity.

After individually assessing all possible first order interaction terms between AGP and covariates, the only covariate with statistically significant interaction was BMI ($p < 0.0001$) and urban/rural location ($p=0.031$). When this term was placed into a full model, however, it was no longer significant at $\alpha=0.05$ ($p=0.46$), and was eliminated. Therefore, the gold-standard model for assessment of confounding contained no interaction terms.

After assessment of confounding, covariates remaining in the model were region and urban/rural, giving the following final model:

$$\text{ANEMIA} = \beta_0 + \beta_1\text{AGP} + + \beta_2\text{REGION} + \beta_3\text{URBANRURAL}$$

Details of the model-building process are included in Appendix III. Parameter estimates, standard errors, and p-values for the terms in the final model appear in Table 6.

The adjusted odds ratio for elevated AGP controlling for region and urban/rural location is presented in Table 6a and the observed AGP prevalence by region and urban/rural location is presented in Table 6b. The odds ratio for anemia in those with elevated AGP, controlling for region and urban/rural location, was 3.98 (95% CI: 1.54, 10.26).

4.6 Association Between Anemia and Infection Category

A model was created to investigate an association between anemia and the four categories of inflammation described previously. However, because of sparse data in several strata (see Appendix IV), this model was unstable, and the analysis was discontinued.

5. Discussion

5.1 Association Between Anemia and Elevated CRP

This study found a statistically significantly association between anemia and elevated levels of C-reactive protein (CRP), which was confounded by region and recent birth and modified by urban/rural location, in non-pregnant women of child-bearing age in Papua New Guinea. In rural areas, the prevalence odds of anemia were 2.74 times higher in those with elevated CRP compared to those with normal levels. However, this relationship between CRP and anemia did not seem to hold in those individuals who lived in urban areas. This finding should be investigated further, both in terms of biological plausibility (perhaps because of differential infection rates or duration in rural vs. urban populations) and in terms of other possible unmeasured confounders which may differ between these two groups. At the very least, from a public health perspective, these findings suggest that any intervention or hemoglobin adjustment strategy based on CRP measurement would likely be most effectively targeted to rural areas.

It is also important to note that smoking was a statistically significant interaction term with CRP status in the initial assessment; however, because the biological plausibility of this relationship was questionable, especially since hemoglobin values had been corrected for tobacco use, this interaction term was left out of the final model. However, this relationship may warrant further consideration in future study because of its strong statistical significance here.

5.2 Association Between Anemia and Elevated AGP

A statistically significant relationship was found between anemia and elevated α-(1) - acid glycoprotein (AGP), confounded by region and urban/rural location, in non-pregnant women of child-bearing age in Papua New Guinea. No significant interaction was found to modify this relationship. The odds of anemia were found to be 3.98 times higher among those with elevated AGP compared to those with normal AGP (95% CI: 1.54, 10.26). Of note is the fact that this adjusted odds ratio, controlling for region and urban/rural location, showed a much stronger association between AGP and anemia than did the crude odds ratio of 2.43.

5.3 Strengths

The overall large sample size in this study was a clear strength of this study, and allowed for examination of complex interaction across multiple strata. Further, the general methodology of the data collection was strong, including the relatively short timeframe (made possible by a large study team) which likely prevented seasonal fluctuations in infection rates and nutritional factors.

Further, the final models used in the analysis were fairly simple, leading to relatively straightforward interpretation of odds ratios.

5.4 Limitations

In some strata, data were very sparse, and this created some problems. In comparison to the large overall sample size, a relatively small percentage of women in this study had elevated CRP or AGP (10.7% and 7.7%, respectively). Once those groups

were broken down further by anemia status, the numbers became even smaller, creating issues with precision. The sample size issue became especially apparent in an attempt to investigate the relationship between anemia and infection category; there simply were not enough women in each of the four groups to create a stable model for analysis.

The Papua New Guinea National Micronutrient Survey was a cross-sectional survey, meaning its analysis had fundamental limitations, such as the inability to draw causal conclusions. In addition, because information was gathered for each individual at only one point in time, no statements can be made about duration of anemia, or any kind of temporal relationship between anemia and biomarkers of inflammation.

Lastly, 3 of the 100 original clusters were inaccessible at the time of survey administration, and therefore no data were collected from these clusters. It is possible then that some selection bias arose from this exclusion.

5.5 Summary

In women of Papua New Guinea, anemia is a moderate-to-severe public health problem, according to the 2005 National Micronutrient Survey, as an estimated 34.93% of non-pregnant women of childbearing age were found to be anemic. This analysis found an association between the acute phase proteins C-reactive protein and anemia, modified by urban/rural location and confounded by region and recent birth, and an association between α-1 acid glycoprotein (AGP) and anemia, confounded by urban/rural location and region. Those individuals with elevated levels of these acute phase proteins were more likely than those with normal levels to be anemic. This suggests that an overall effective strategy for treating anemia in this population may need to include

addressing the issue of inflammation and infection in addition to nutritional and other factors.

Future studies should investigate the possible interaction between CRP and smoking status, to examine whether there is any biological reason for the statistical association that was observed here. In addition, because geographical location (urban/rural and region) were important factors in every model presented here, it is clear that similar studies are needed in other settings to further understand the relationships observed in this analysis.

REFERENCES

1. World Health Organization. Iron Deficiency Anaemia, Assessment, Prevention and Control: A Guide for Programme Managers. Geneva: World Health Organization, 2001.

2. Tolentino K, Friedman JF. An update on anemia in less developed countries. Am J Trop Med Hyg. 2007;77(1):44-51.

3. Gabay C, Kushner I. Acute-phase proteins and other systemic responses to inflammation. N Engl J Med. 1999;340(6):448-54.

4. Thurnham DI, McCabe GP, Northrop-Clewes CA, Nestel P. Effects of subclinical infection on plasma retinol concentrations and assessment of prevalence of vitamin A deficiency: meta-analysis. Lancet. 2003;362(9401):2052-8.

5. McLean E, Cogswell M, Egli I, et al. Worldwide prevalence of anaemia, WHO Vitamin and Mineral Nutrition Information System, 1993-2005. Public Health Nutr. 2009;12(4):444-54.

6. WHO. Iron Deficiency Anaemia, Assesment, Prevention and Control: A Guide for Programme Managers. Geneva: World Health Organization, 2001 2001. Report No.

7. Yip R, Ramakrishnan U. Experiences and challenges in developing countries. J Nutr. 2002;132(4 Suppl):827S-30S.

8. Huma N, Salim Ur R, Anjum FM, et al. Food fortification strategy--preventing iron deficiency anemia: a review. Crit Rev Food Sci Nutr. 2007;47(3):259-65.

9. CDC criteria for anemia in children and childbearing-aged women. MMWR Morb Mortal Wkly Rep. 1989;38(22):400-4.

10. Wieringa FT, Dijkhuizen MA, West CE, et al. Estimation of the effect of the acute phase response on indicators of micronutrient status in Indonesian infants. J Nutr. 2002;132(10):3061-

11. Fournier T, Medjoubi NN, Porquet D. Alpha-1-acid glycoprotein. Biochim Biophys Acta. 2000;1482(1-2):157-71.

12. Gorstein JL, Dary O, Pongtorn, et al. Feasibility of using retinol-binding protein from capillary blood specimens to estimate serum retinol concentrations and the prevalence of vitamin A deficiency in low-resource settings. Public Health Nutrition. 2007.

13. Wieringa FT, Dijkhuizen MA, West CE, et al. Estimation of the effect of the acute phase response on indicators of micronutrient status in Indonesian infants. J Nutr. 2002;132(10):3061-

14. Christian P, Schulze K, Stoltzfus RJ, West KP, Jr. Hyporetinolemia, illness symptoms, and acute phase protein response in pregnant women with and without night blindness. Am J Clin Nutr. 1998;67(6):1237-43.

15. Stephensen CB, Gildengorin G. Serum retinol, the acute phase response, and the apparent misclassification of vitamin A status in the third National Health and Nutrition Examination Survey.[see comment]. Am J Clin Nutr. 2000;72(5):1170-8.

16. Beard JL, Murray-Kolb LE, Rosales FJ, et al. Interpretation of serum ferritin concentrations as indicators of total-body iron stores in survey populations: the role of biomarkers for the acute phase response. Am J Clin Nutr. 2006;84(6):1498-505.

17. Thurnham DI, McCabe LD, Haldar S, et al. Adjusting plasma ferritin concentrations to remove the effects of subclinical inflammation in the assessment of iron deficiency: a meta-analysis. Am J Clin Nutr. 2010;92(3):546-55.

18. World Health Organization. Country Health Information Profiles: Papua New Guinea. 2007 Revision [November 11, 2007]; Available from: http://www.wpro.who.int/countries/png.

19. Department of Health of Papua New Guinea. National Micronutrient Survey PNG, 2005. Atlanta: Department of Health PNG, UNICEF PNG, University of PNG, CDC, 2006.

20. Sullivan KM, Mei Z, Grummer-Strawn L, Parvanta I. Haemoglobin adjustments to define anaemia. Trop Med Int Health. 2008;13(10):1267-71.

21. Davis CE, Hyde, J.E., Bangdiwala, S.I., Nelson, J.J. An example of dependencies among variables in a conditional of logistic regression. In: S.H. Moolgavkar and R.L. Prentice, editor. Modern Statistical Methods in Chronic Disease Epidemiology. New York: John Wiley & Sons; 1986. p. 140-7.

TABLES

Table 1: Characteristics of the study population

Characteristic	N	Unweighted %	Weighted %
ANEMIC			
Yes	265	40.03	34.93
No	397	59.97	65.07
ELEVATED CRP			
Yes	71	10.73	10.43
No	591	89.27	89.57
ELEVATED AGP			
Yes	51	7.70	7.96
No	611	92.3	92.04
STAGE OF INFECTION			
No inflammation	567	85.65	85.78
Incubation	44	6.65	6.268
Early Convalescence	27	4.08	4.16
Late Convalescence	24	3.63	3.80
AGE GROUP (years)			
15-19	47	7.37	7.81
20-24	133	20.85	20.65
25-29	141	22.10	22.65
30-34	123	19.28	19.50
35-39	80	12.54	12.94
40-44	70	10.97	9.77
45-49	44	6.90	6.68
REGION			
Southern	209	31.57	18.84
Highlands	154	23.26	40.11
Momase	152	22.96	27.18
Islands	147	22.21	13.88
URBAN/RURAL			
Urban	166	25.08	20.53
Rural	496	74.92	79.47
HOUSEHOLD SIZE			
1-5 people	177	26.74	25.52
6-8 people	142	21.45	21.95
9+ people	343	51.81	52.53
SMOKER			
Yes	157	23.75	25.98
No	504	76.25	74.02
BIRTH IN PAST 3 YEARS			
Yes	302	45.62	45.51
No	360	54.38	54.49
BMI			
Underweight	43	6.58	5.38
Normal	446	68.30	70.43
Overweight	116	17.76	18.29
Obese	48	7.35	5.90
EDUCATION			
No formal education	152	23.53	27.82
Grades 1-3	70	10.84	11.26
Grades 4-8	293	45.36	43.30
Grades 9+	131	20.28	17.63
TOTAL	662		

Table 2: Anemia prevalence with respect to covariates

Covariates	N	Anemic (Hb<12g/dl) (#)	Anemic (Hb<12g/dl) (%)**	POR (95%CI)**	X2	p-value**
ELEVATED CRP					2.48	0.1154
Yes	71	33	43.98%	1.53 (0.90, 2.61)		
No	591	232	33.87%	Ref.		
ELEVATED AGP					7.00	0.0081*
Yes	51	30	54.73%	2.43 (1.26, 4.69)*		
No	611	235	33.22%	Ref.		
STAGE OF INFECTION					7.13	0.0679
No inflammation	567	217	32.98%	Ref.		
Incubation	44	18	36.38%	1.16 (0.59, 2.28)		
Early Convalescence	27	15	55.43%	2.53 (1.04, 6.14)*		
Late Convalescence	24	15	53.95%	2.38 (0.97, 5.84)		
AGE GROUP (years)					4.68	0.5850
15-19	47	16	23.73%	0.43 (0.19, 0.97)*		
20-24	133	52	33.61%	0.70 (0.36, 1.38)		
25-29	141	58	34.59%	0.74 (0.39, 1.40)		
30-34	123	50	34.96%	0.75 (0.37, 1.51)		
35-39	80	32	39.29%	0.90 (0.44, 1.86)		
40-44	70	29	40.73%	0.96 (0.46, 2.04)		
45-49	44	20	41.82%	Ref.		
					54.87	<.0001*
REGION						
Southern	209	87	42.65%	7.44 (3.63, 15.22)*		
Highlands	154	14	9.09%	Ref.		
Momase	152	92	60.53%	15.33 (7.14, 32.94)*		
Islands	147	72	48.98%	9.60 (4.62, 19.97)*		
URBAN/RURAL					7.01	0.0081*
Urban	166	40	19.48%	Ref.		
Rural	496	225	38.92%	2.63 (1.29, 5.39)*		
HOUSEHOLD SIZE					3.13	0.2095
1-5 people	177	78	39.40%	Ref.		
6-8 people	142	60	37.29%	1.40 (0.96, 2.04)		
9+ people	343	127	31.76%	1.28 (0.80, 2.04)		
SMOKER					0.0533	0.8173
Yes	157	61	33.83%	Ref.		
No	504	203	35.23%	1.06 (0.63. 1.79)		
BIRTH IN PAST 3 YEARS					5.7205	0.0168*
Yes	302	138	40.57%	1.57 (1.085, 2.27)*		
No	360	127	30.21%	Ref.		
BMI					20.84	0.0001*
Underweight	43	25	62.49%	5.18 (1.727, 15.52)*		
Normal	446	191	37.41%	1.86 (0.88, 3.93)		
Overweight	116	34	20.46%	0.80 (0.35, 1.83)		
Obese	48	12	20.35%	Ref.		
EDUCATION					5.28	0.1522
No formal education	152	64	36.07%	1.29 (0.71, 2.35)		
Grades 1-3	70	36	46.34%	1.98 (1.06, 3.72)		
Grades 4-8	293	115	34.03%	1.18 (0.72, 1.94)		
Grades 9+	131	47	30.36%	Ref.		
OVERALL	662		34.93%			

*****Significant at α=0.05
**Takes into account weighted analysis
Design effect: CRP=1.3, AGP=1.7

Table 3: Elevated CRP prevalence with respect to covariates

Covariates	N	Elevated CRP (CRP>5mg/L) (#)	Elevated CRP (CRP>5mg/L) (%)**	POR (95%CI)**	X2	p-value**
ANEMIC					2.48	0.1154
Yes	265	33	13.13%	1.53 (0.90, 2.61)		
No	397	38	8.98%	Ref.		
ELEVATED AGP					54.44	<0.001
Yes	51	27	52.30%	15.00 (7.31, 30.81)*		
No	611	44	6.81%	Ref.		
AGE GROUP (years)					5.91	0.4337
15-19	47	6	14.31%	3.84 (0.99, 14.86)		
20-24	133	10	7.56%	1.88 (0.49, 7.30)		
25-29	141	16	11.90%	3.11 (0.84, 11.56)		
30-34	123	16	12.96%	3.42 (0.93, 12.64)		
35-39	80	8	10.05%	2.57 (0.59, 11.13)		
40-44	70	10	12.34%	3.24 (0.80, 13.16)		
45-49	44	3	4.17%	Ref.		
REGION					3.30	0.3479
Southern	209	20	9.00%	0.80 (0.36, 1.77)		
Highlands	154	17	11.04%	Ref.		
Momase	152	21	8.55%	0.76 (0.38, 1.50)		
Islands	147	71	14.29%	1.34 (0.65, 2.77)		
URBAN/RURAL					0.0003	0.9865
Urban	166	20	10.47%	Ref.		
Rural	496	51	10.42%	0.99 (0.52, 1.92)		
HOUSEHOLD SIZE					0.070	0.9657
1-5 people	177	17	9.99%	Ref.		
6-8 people	142	15	10.08%	1.01 (0.45, 2.29)		
9+ people	343	39	10.79%	1.10 (0.52, 2.30)		
SMOKER					0.23	0.6320
Yes	157	14	9.55%	Ref.		
No	504	57	10.75%	1.141 (0.67, 1.96)		
BIRTH IN PAST 3 YEARS					0.58	0.4471
Yes	302	33	10.95%	1.23 (0.73, 2.07)		
No	360	38	9.99%	Ref.		
BMI					10.94	0.0120*
Underweight	43	7	19.04%	0.85 (0.31, 2.30)		
Normal	446	38	8.77%	0.35 (0.1, 0.74)*		
Overweight	116	11	9.62%	0.38 (0.14, 1.06)		
Obese	48	12	21.75%	Ref.		
EDUCATION					5.71	0.1266
No formal education	152	12	8.21%	0.97 (0.35, 2.70)		
Grades 1-3	70	13	19.30%	2.60 (1.00, 6.71)*		
Grades 4-8	293	31	10.41%	1.26 (0.53, 3.00)		
Grades 9+	131	14	8.44%	Ref.		
OVERALL	662		10.45%			

*Significant at α=0.05
**Takes into account weighted analysis
Design effect: CRP=1.3, AGP=1.7

Table 4: Elevated AGP prevalence with respect to covariates

Covariates	N	Elevated AGP (CRP>5mg/L) (#)	Elevated AGP (CRP>5mg/L) (%)**	POR (95%CI)**	X2	p-value**
ANEMIC					7.00	0.0081*
Yes	265	30	12.47%	2.43 (1.26, 4.69)*		
No	397	21	5.54%	Ref.		
ELEVATED CRP					54.44	<0.0001*
Yes	71	27	39.90%	15.00 (7.31, 30.81)*		
No	591	24	4.24%	Ref.		
AGE GROUP (years)					5.39	0.4953
15-19	47	4	9.57%	1.77 (0.25, 12.69)		
20-24	133	8	6.33%	1.13 (0.19, 6.61)		
25-29	141	8	6.85%	1.23 (0.21, 7.38)		
30-34	123	7	6.11%	1.09 (0.18, 6.69)		
35-39	80	11	13.11%	2.53 (0.39, 16.18)		
40-44	70	10	12.75%	2.45 (0.40, 14.90)		
45-49	44	2	5.64%	Ref.		
REGION					2.09	0.5546
Southern	209	19	9.38%	1.34 (0.41, 2.60)		
Highlands	154	14	9.09%	Ref.		
Momase	152	10	6.58%	0.70 (0.29, 1.74)		
Islands	147	8	5.44%	0.58 (0.23, 1.44)		
URBAN/RURAL					1.47	0.2248
Urban	166	9	5.27%	Ref.		
Rural	496	42	8.65%	1.70 (0.72, 4.03)		
HOUSEHOLD SIZE					2.18	0.3366
1-5 people	177	10	5.97%	Ref.		
6-8 people	142	14	11.27%	2.00 (0.69, 5.78)		
9+ people	343	27	7.54%	1.29 (0.52, 3.16)		
SMOKER					0.57	0.4516
Yes	157	13	9.21%	Ref.		
No	504	38	7.53%	0.80 (0.45, 1.42)		
BIRTH IN PAST 3 YEARS					0.77	0.3808
Yes	302	27	9.32%	1.33 (0.70, 2.53)		
No	360	24	6.82%	Ref.		
BMI					6.68	0.0828
Underweight	43	5	16.85%	1.702 (0.39, 7.36)		
Normal	446	35	7.79%	0.71 (0.20, 2.57)		
Overweight	116	5	3.93%	0.34 (0.07, 1.76)		
Obese	48	3	10.64%	Ref.		
EDUCATION					4.47	0.2150
No formal education	152	16	10.53%	0.99 (0.35, 2.78)		
Grades 1-3	70	7	10.64%	0.57 (0.24, 1.40)		
Grades 4-8	293	18	6.39%	0.63 (0.18, 2.18)		
Grades 9+	131	10	6.93%	Ref.		
OVERALL	662		7.96%			

*Significant at α=0.05
**Takes into account weighted analysis
Design effect: CRP=1.3, AGP=1.7

Table 5: Logistic regression analysis of factors influencing anemia status, with CRP as main exposure variable

FINAL MODEL:
Anemia = $\beta 0 + \beta_1$CRP + β_2REGION + β_3BIRTH3 + β_4URBANRURAL + β_5(CRP*URBANRURAL)

Observations in analysis: 662

Variable	Parameter Estimate (β_n)	S.E.	Wald χ 2	p-value
Intercept	-2.52	0.31	65.42	<.0001**
CRP	1.01	0.41	6.02	0.0142**
REGION				
1 = Southern	2.45	0.40	38.34	<.0001**
2 = Highlands				
3 = Momase	2.82	0.40	48.73	<.0001**
4 = Islands	2.27	0.37	38.06	<.0001**
BIRTH3	0.44	0.20	4.89	0.027**
URBANRURAL	-1.06	0.30	12.06	0.0005**
CRP*URBANRURAL	-1.85	0.66	7.94	0.0048**

**Significant at α=0.05
All models take into account weighting and complex survey design

Table 5a: Adjusted odds ratios for the association between anemia and CRP, with respect to urban/rural location, controlling for region and birth in past 3 years

	Rural, Elevated CRP	Urban, Elevated CRP	Rural, Normal CRP	Urban, Normal CRP
Adjusted Odds Ratio	2.74	0.15	1.00	0.35
95% CI	(1.23, 6.15)*	(0.05, 0.45)*	-	(0.19, 0.63)*

***Significant at α=0.05**

Table 5b: Anemia prevalence* with respect to combined urban/rural and CRP status

	Rural, Elevated CRP	Urban, Elevated CRP	Rural, Normal CRP	Urban, Normal CRP
Anemia Prevalence N (%)**	30 (52.80%)	3 (10.01%)	195 (37.30%)	37 (20.58%)
Prevalence Ratio	2.57	0.49	1.81	1.00

***Observed prevalence; does not account for age group and region**
****Weighted, taking into account complex survey design**

Table 6: Logistic regression analysis of factors influencing anemia status, with AGP as main exposure variable

FINAL MODEL:
Anemia = $\beta_0 + \beta_1$AGP + β_2REGION + β_3URBANRURAL

Observations in analysis: 516

Variable	Parameter Estimate (β_n)	S.E.	Wald χ^2	p-value
Intercept	-2.36	0.31	57.56	<.0001
AGP	1.38	0.48	8.16	0.0043
REGION				
1 = Southern	2.41	0.38	40.05	<.0001
2 = Highlands	Ref.	-	-	-
3 = Momase	2.89	0.41	50.79	<.0001
4 = Islands	2.37	0.37	40.26	<.0001
URBANRURAL	-1.18	0.29	16.98	<.0001

**Significant at α=0.05
All models take into account weighting and complex survey design

Table 6a: Adjusted odds ratios for the association between anemia and AGP, controlling for region and urban/rural location

	Elevated AGP	Normal AGP
Adjusted Odds Ratio	3.98	1.00
95% CI	(1.54, 10.26)*	-

***Significant at α=0.05**

Table 6b: Anemia prevalence with respect to AGP status, by region and urban/rural location

Anemia Prevalence N* (%)	Elevated AGP	Normal AGP
OVERALL	51 (54.73%)	611 (33.22%)
REGION		
Southern	12 (65.05%)	75 (40.33%)
Highlands	5 (35.71%)	9 (6.43%)
Momase	8 (80.00%)	84 (59.15%)
Islands	5 (62.50%)	67 (48.20%)
URBANRURAL		
Urban	2 (13.27%)	38 (19.82%)
Rural	28 (61.24%)	197 (36.80%)

* Weighted, taking into account complex survey design

APPENDICES

APPENDIX I: Data Collection Forms

WOMEN (15-49 YEARS)

Label

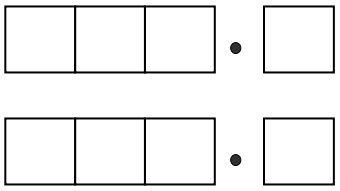
TEAM CODE ☐

VERBAL CONSENT OBTAINED FROM ELIGIBLE WOMAN Yes ☐ No ☐

1.	*Woman's name:*	
2.	*Woman's age*	☐☐ years
3.	WHAT IS YOUR HIGHEST GRADE OF EDUCATION COMPLETED? **YU PINISIM WANEM GRET LONG SKUL?** *(0= No school completed 1-3=Elementary School 4-8= Primary School 9-12=Secondary school)*	Highest grade completed ☐☐ Refused .....7 Other (specify) .....8 Don't know 9
4.	DID YOU SLEEP UNDER A MOSQUITO NET LAST NIGHT? **YU BIN SLIP ANINIT LONG MOSKITO NET O TAUNAM LONG LAS NAIT?**	Yes .....1 No .....2 Refused .....7 Don't know 9
5.	HOW MANY MOSQUITO NETS DOES YOUR HOUSEHOLD HAVE? **HAUS BILONG YU I GAT HAMAS TAUNAM?**	Number of nets ☐☐
6.	DO YOU SMOKE? **YU SAVE SMOK TU?**	Yes.....1 No.....2 2⇨Q.8 Refused .....7 Don't know .....9 9⇨Q.8
7.	HOW MANY STICKS DO YOU SMOKE PER DAY? **HAMASPELA STIK SIMUK YU SAVE SMOKIM INSAIT LONG WANPELA DE?**	Number per day ☐☐
8.	HAVE YOU EVER BEEN PREGNANT?	Yes 1

YU BIN GAT BEL TU? *(Should be asked by female or with female present.)*	No .....2 Refused.....7 Don't know 9	2⇨Q.17 9⇨Q.17
9. HAVE YOU GIVEN BIRTH TO A CHILD IN THE LAST 3 YEARS? **INSAIT LONG LASPELA TRIPELA YIA, YU BIN KARIM WANPELA PIKININI TU?** (*This includes both live births and still births BUT NOT miscarriages*) *(Ask for meri book if available)*	Yes..... 1 No ..... 2 Refused ..... 7 Don't know 9	 2⇨Q.17 9⇨Q.17
10. WHEN YOU WERE PREGNANT WITH YOUR LAST CHILD, DID YOU RECEIVE IRON TABLETS? **TAIM YU BIN BEL LONG LASPELA PIKININI BILONG YU, YU SAVE KISIM AIN TABLET?** *(Show an example of the iron tablet)*	Yes ..... 1 No ..... 2 Refused..... 7 Don't know 9	 2⇨Q.12 9⇨Q.12
11. WHO DID YOU RECEIVE THE IRON TABLETS FROM? **YU BIN KISIM OL AIN TABLET LONG HUSAT?**	Health centre .....1 Health workers on patrol .....2 VBA .....3 VHV.....4 Refused.....7 Other (specify) .....8 Don't know 9	
12. WAS YOUR LAST BORN CHILD WEIGHED AT BIRTH? **OL BIN SKELIM LASPELA PIKININI BILONG YU TAIM YU KARIM?**	Yes .....1 No .....2 Refused.....7 Don't know 9	 2⇨Q.15 9⇨Q.15
13. WHAT WAS THIS CHILD'S WEIGHT **WANEM MAK LONG WEIT O HEVI BILONG EM?** *(Record weight from baby book/health card, if available.)*	☐☐☐☐ grams	
14. Write down where information on the birth weight was obtained from.	From recall .....1 From clinic book .....2 Other (specify) 8	
15. WHEN YOU WERE PREGNANT WITH YOUR LAST CHILD, DID YOU HAVE DIFFICULTY SEEING DURING THE DAY? **TAIM YU BIN BEL WANTAIM LASPELA PIKININI BILONG YU, YU BIN GAT HEVI LONG LUKLUK LONG SAN?**	Yes .....1 No .....2 Refused .....7 Don't know 9	
16. WHEN YOU WERE PREGNANT WITH YOUR LAST CHILD DID YOU HAVE ANY DIFFICULTY SEEING AT DUSK? **TAIM YU BIN BEL WANTAIM LASPELA PIKININI BILONG YU, YU BIN GAT HEVI LONG LUKLUK TAIM EM I LAIK TUDAK?**	Yes .....1 No .....2 Refused .....7 Don't know 9	

17. ARE YOU CURRENTLY PREGNANT? **YU GAGT BEL NAU?** *(If YES end the interview. DO NOT take anthropometric measurements or urine or blood samples)*	Yes .....1 No .....2 Refused .....7 Don't know 9	1⇨END
*Weigh and measure each woman after all questionnaires have been completed. **DO NOT** measure any woman with casts, heavy bandages or disabilities that prevent them being measured. **DO NOT** measure women who are pregnant.*		



ANTHROPOMETRY MODULE	
18. Woman's weight	kg
19. Woman's height	cm
20. *Circle result for height measurement*	Measured .....1 Refused .....7 Other (specify) .....8 Unable 9
CHECK Are there any other women in the household who are eligible for measurement? If not, pass the data collection form on to the laboratory technician.	

SPECIMEN COLLECTION MODULE Do NOT take urine or blood samples from pregnant women	
21. Was urine sample collected from this woman?	Yes .....1 No .....2 Refused .....7 Other (specify).....8
22. *Ask* "WE WOULD LIKE TO TAKE SOME OF YOUR BLOOD FROM YOUR FINGER, FOR TESTING. IS THIS OK? **"MIPELA I LAIK KISIM SAMPELA BLUT LONG PINGA BILONG YU LONG KARIMAUT TES. EM I ORAIT WANTAIM YU?**	Yes .....1 No .....2 Refused .....7 Other (specify) 8
23. Write down the hemoglobin level *(If the Hb is 7 or less then write the result in the space provided and also on a referral sheet and on a referral slip for the health center)*	g/dl
24. *Was finger stick blood sample collected from this woman?*	Yes .....1 Not available .....2 Refused .....7 Other (specify).....8
25. Approximately how many microlitres of finger stick blood were collected from this woman.	microl

FOR NCD CLUSTERS ONLY	
26. Was a venous blood sample collected from this woman?	Yes .....1 Not available .....2 Refused .....7 Other (specify) 8
27. Approximately how many milliliters of venous blood were collected from this woman	ml

THANK *the participant for their cooperation*
CHECK *that all the data collection form has been completed correctly*
CHECK *that the identification numbers are at the top of each page.*

Data Entry Information Panel

(To be completed by the data entry clerks)

First data entry clerk ID number		Second data entry clerk ID number	

HOUSEHOLD QUESTIONNAIRE

TEAM CODE ☐

"WE WOULD LIKE TO TALK TO YOU ABOUT YOUR HOUSEHOLD, THAT IS ALL THE PEOPLE WHO USUALLY SLEEP AND EAT HERE."
"MIPELA I LAIK TOKTOK LONG YU LONG HAUS BILONG YU. DISPELA EM OLGETA PIPEL HUSAT I SAVE SLIP NA KAIKAI HIA."

Read the survey consent form and ask for verbal consent. If consent is not obtained then move on to the next household. If there are no adult household members present in the household schedule another visit when an adult household member will be present.

VERBAL CONSENT OBTAINED FROM ADULT HOUSEHOLD MEMBER Yes ☐ No ☐

1. Day/Month/Year of interview:	Day ☐☐ Month ☐☐ Year ☐☐☐☐	
2. Census Unit		
3. Ward		
4. LLG		
5. District		
6. Province		
7. Region		
8. HOW MANY PEOPLE NORMALLY LIVE IN THIS HOUSEHOLD? **HAMAS PIPEL I SAVE STAP LONG DISPELA HAUS?** *(People who usually eat and sleep in the household)*	☐☐	
9. ARE THERE ANY WOMEN BETWEEN THE AGES OF 15 AND 49 YEARS WHO USUALLY LIVE IN THIS HOUSEHOLD? **I GAT SAMPELA MERI WE KRISMAS BILONG OL I STAP NAMEL LONG 15 NA 49 YIAS I SAVE STAP LONG DISPELA HAUS?**	Yes ..... 1 No ..... 2 Refused ..... 7 Don't know 9	2⇨Q.12 9 ⇨Q.12

10. HOW MANY WOMEN BETWEEN 15 AND 49YEARS LIVE IN THIS HOUSEHOLD? **HAMAS MERI I GAT KRISMAS NAMEL LONG 15 NA 49 YIAS I SAVE STAP LONG DISPELA HAUS?**	☐	
11. COULD YOU PLEASE TELL ME THE NAME AND AGE OF EACH WOMAN AGED 15 TO 49 YEARS WHO LIVES IN THIS HOUSEHOLD EVEN IF THEY ARE NOT HERE RIGHT NOW? **PLIS INAP YU TOKIM MI NEM NA KRISMAS BILONG OL WAN WAN MERI I SAVE STAP LONG DISPELA HAUS NA I GAT KRISMAS NAMEL LONG 15 NA 49 YIAS, MASKI OL I NO STAP LONG HAUS NAU?**	Name _____ Age (Years) 1._____ ☐☐ 2._____ ☐☐ 3._____ ☐☐ 4._____ ☐☐ 5._____ ☐☐	
12. ARE THERE ANY MEN AGED 18 YEARS AND OLDER WHO USUALLY LIVE IN THIS HOUSEHOLD? **I GAT SAMPELA MAN KRISMAS BILONG OL EM 18 NA MOA I SAVE STAP LONG DISPELA HAUS?**	Yes..... 1 No ..... 2 Refused ..... 7 Don't know 9	 2⇨Q.15 9⇨Q.15
13. HOW MANY MEN 18 AND OLDER LIVE IN THIS HOUSEHOLD? **HAMAS MAN WANTAIM KRISMAS NAMEL LONG 18 NA MOA I STAP LONG DISPELA HAUS?**	☐	
14. COULD YOU PLEASE TELL ME THE NAME AND AGE OF EACH MAN AGED 18 YEARS AND OLDER WHO LIVES IN THIS HOUSEHOLD EVEN IF THEY ARE NOT HERE RIGHT NOW? **PLIS INAP YU TOKIM MI NEM NA KRISMAS BILONG WAN WAN MAN I GAT 18 KRISMAS NA MOA, MASKI OL I INO STAP LONG HAUS NAU.**	Name _____ Age (Years) 1. _____ ☐☐ 2._____ ☐☐ 3._____ ☐☐ 4._____ ☐☐ 5._____ ☐☐	
15. ARE THERE ANY CHILDREN AGED 6 MONTHS TO 5 YEARS WHO USUALLY LIVE IN THIS HOUSEHOLD? **I GAT SAMPELA PIKININI I GAT KRISMAS NAMEL LONG 6-PELA MUN NA 5-PELA KRISMAS I STAP LONG DISPELA HAUS?**	Yes..... 1 No ..... 2 Refused ..... 7 Don't know 9	 2⇨Q.18 9⇨Q.18

16. HOW MANY CHILDREN BETWEEN 6 MONTHS TO 5 YEARS LIVE IN THIS HOUSEHOLD? **HAMAS PIKININI I GAT KRISMAS NAMEL LONG 5-PELA MUN NA 5-PELA YIA I STAP LONG DISPELA HAUS?**	☐
17. COULD YOU PLEASE TELL ME THE NAME AND AGE OF EACH CHILD AGED 6 MONTHS TO 5 YEARS WHO LIVES HERE EVEN IF THEY ARE NOT HERE NOW? **PLIS NINAP YU TOKIM MI LONG NEM NA KRISMAS BILONG WAN WAN PIKININI I GAT KRISMAS NAMEL LONG 5-PELA MUN NA 5-PELA KRISMAS I SAVE STAP LONG DISPELA HAUS. MASKI OL I NO STAP LONG HAUS NAU, BAI YU GIVIM NEM NA KRISMAS BILONG OL.** *(Check the clinic book or other document for confirmation of names and ages)*	Name Age in: Years Months 1._____ ☐ ☐☐ 2._____ ☐ ☐☐ 3._____ ☐ ☐☐ 4._____ ☐ ☐☐ 5._____ ☐ ☐☐

18. What type of house is this? *(Observation: Use your own judgment. Do not ask the respondent the answer to this question)*	High cost house ..... 1 Low cost house ..... 2 Flat ..... 3 Duplex ..... 4 Domestic quarters ..... 5 Dormitory ..... 6 Makeshift ..... 10 Traditional ..... 11 Self-help high cost ..... 12 Self-help low cost ..... 13 Other (specify) ..... 8 Don't know 9
19. WHAT IS THE MAIN SOURCE OF DRINKING WATER FOR MEMBERS OF YOUR HOUSEHOLD? **YUPELA LONG HAUS I SAVE KISIM WARA BILONG DRING WE?** *(If necessary confirm this visually)*	Piped into yard or plot ..... 1 Piped into neighborhood (communal) ..... 2 Public well ..... 3 Well in yard ..... 4 Spring ..... 5 River/stream ..... 6 Pond/lake/dam ..... 10 Communal tank ..... 11 Rainwater ..... 12 Tanker-truck, vendor ..... 13 Refused ..... 7 Other (specify) ..... 8 Don't know 9
20. WHAT KIND OF TOILET FACILITY DOES YOUR HOUSEHOLD USE? **WANEM KAIN TOILET YUPELA I YUSIM**?	Flush to sewage system or septic tank ..... 1 Pour flush latrine (water seal type) ..... 2 Improved pit latrine (e.g., VIP) ..... 3 Traditional pit latrine ..... 4 Open pit ..... 5 Bucket ..... 6 No facilities or bush/field/beach ..... 10 Overhang latrine ..... 11 Refused ..... 7 Other (specify) ..... 8 Don't know 9
21. HOW OFTEN DO YOU LISTEN TO THE RADIO? **HAMAS TAIM YU SAVE HARIM REDIO?**	I never listen to the radio ..... 1 Every day ..... 2 Every week ..... 3 Occasionally ..... 4 Other (specify) 8

This next section should be completed by the female head of the household or another person in the household familiar with the salt, flour, oil, sugar and rice used in the household.

"WE ARE INTERESTED IN THE TYPES OF FOOD THAT PEOPLE EAT IN PAPUA NEW GUINEA. I WILL BE ASKING TO SEE THE SALT, FLOUR, OIL, SUGAR AND RICE, AND THEIR PACKAGES, THAT YOU HAVE IN THE HOUSE TODAY. YOU MIGHT WANT TO COLLECT THESE ITEMS BEFORE WE BEGIN THIS PART OF THE INTERVIEW."

"MIPELA I GAT INTRES LONG OL KAIN KAIKAI WE OL PIPEL BILONG PNG I SAVE KAIKAIM. BAI MI ASKIM LONG LUKIM SOL, FLAUA, OIL, SUGA, RAIS, NA OL PEKET BILONG OL BIPO YUMI STATIM DISPELA HAP BILONG ASKIM."

SALT MODULE *If two or more types of salt are available in the household record information on the two main types of salt used in the household.*		
22. DO YOU HAVE ANY SALT CURRENTLY IN YOUR HOUSEHOLD NOW? **YU GAT SAMPELA SOL LONG HAUS BILONG YU NAU?**	Yes ..... 1 No ..... 2 Don't know 9	2 ⇨ Q. 40
23. *If Yes ASK* "MAY I SEE A SAMPLE OF EACH TYPE OF SALT YOU HAVE IN THE HOUSEHOLD" **"INAP MI LUKIM SEMPOL LONG OL KAIN SOL YU GAT LONG HAUS BILONG YU"** *(If there is more than one type of salt record the information for just one type of salt here. Record the information for another type of salt in the Type 2 salt module beginning with question 31.)* *(Observe the type of salt used and circle the appropriate answer)*	Fine table salt ..... 1 Cooking salt ..... 2 Traditional salt ..... 3 Sea water used for cooking ..... 4 Refused ..... 7 Other (specify) ..... 8 Don't know 9	4 ⇨ Q.31
24. *If you DO NOT see the original salt bag or package ask* "COULD I PLEASE SEE THE ORIGINAL SALT BAG OR PACKAGE?" **"PLIS INAP MI LUKIM SOL BEK O PEKET SOL I BIN STAP LONG EN?"**	Yes, original salt bag or package observed ..... 1 No, original salt bag or package not observed .. 2	2 ⇨ Q. 29
25. Write the name of the brand of salt written on the package	Brand name _____	
26. Observe the country where the salt is produced	Papua New Guinea ..... 1 Australia ..... 2 India ..... 3 China…. ..... 4 Thailand…. ..... 5 Other (specify) ..... 8 Don't know 9	
27. Observe the country where the salt is packaged	Papua New Guinea ..... 1 Australia ..... 2 India ..... 3 China…. ..... 4 Thailand…. ..... 5 Other (specify) ..... 8 Don't know 9	

28. <u>Observe</u> – Is the salt iodized?	Yes ..... 1 No or not stated on label ..... 2 Don't know..... 9
29. MAY I ASK WHERE YOU GOT THE SALT FROM? **INAP MI ASKIM YU WE YU BIN KISIM DISPELA SOL?**	Purchased from a shop ..... 1 Purchased from a vendor ..... 2 Mined/collected from the rock ..... 3 Other (specify) ..... 8 Don't know..... 9
30. MAY I TAKE A SAMPLE OF THIS SALT TO THE LABORATORY TO TEST FOR IODINE CONTENT? **INAP MI KISIM SEMPOL LONG DISPELA SOL I GO LONG LEBORETORI LONG TESTIM SAPOS EM MI GAT AIDIN LONG EN?** *(Collect the required amount of salt and replace the salt you have taken with 1 packet of iodized salt)*	Salt sample collected..... 1 Salt sample not collected ..... 2 

TYPE 2 SALT ***If there is a second type of salt used in the household record the information here***		
31. DO YOU HAVE ANY OTHER TYPE OF SALT CURRENTLY IN YOUR HOUSEHOLD NOW? **YU GAT OL SAMPEAL NARAPELA SOL LONG HAUS BILONG YU NAU?**	Yes ..... 1 No ..... 2 Don't know 9	2 ⇨ Q. 40
32. *If Yes ask* "MAY I SEE THIS SALT" **"INAP MI LUKIM DISPELA SOL?"** *(Observe the type of salt used and circle the appropriate answer)*	Fine table salt ..... 1 Cooking salt ..... 2 Traditional salt ..... 3 Sea water used for cooking ..... 4 Refused ..... 7 Other (specify) ..... 8 Don't know 9	
33. *If you DO NOT see the original salt bag or package ask* "COULD I PLEASE SEE THE ORIGINAL SALT BAG OR PACKAGE?" **"PLIS INAP MI LUKIM SOL BEK O PEKET SOL I BIN STAP LONG EN?"**	Yes, original salt bag or package observed ..... 1 No, original salt bag or package not observed .. 2	2 ⇨ Q. 38
34. Write the name of the brand of salt written on the package	Brand……	
35. Observe the COUNTRY where the salt is produced	Papua New Guinea ..... 1 Australia ..... 2 India ..... 3 China…. ..... 4 Thailand…. ..... 5 Other (specify) ..... 8 Don't know 9	
36. Observe the country where the salt is packaged	Papua New Guinea ..... 1 Australia ..... 2 India ..... 3 China…. ..... 4 Thailand…. ..... 5 Other (specify) ..... 8 Don't know 9	
37. Observe – Is the salt iodized?	Yes ..... 1 No or not stated on label ..... 2 Don't know 9	

38. MAY I ASK WHERE YOU GOT THE SALT FROM? **INAP MI ASKIM YU WE YU BIN KISIM DISPELA SOL?**	Purchased from a shop ..... 1 Purchased from a vendor ..... 2 Mined/collected from the rock ..... 3 Other (specify) ..... 8 Don't know 9
39. MAY I TAKE A SAMPLE OF THIS SALT TO THE LABORATORY TO TEST FOR IODINE CONTENT? **INAP MI KISIM SEMPOL LONG DISPELA SOL I GO LONG LEBORETORI LONG TESTIM SAPOS EM MI GAT AIDIN LONG EN?** *(Collect the required amount of salt and replace the salt you have taken with 1 packet of iodized salt)*	Salt sample collected ..... 1 Salt sample not collected ..... 2 Salt Type 2 Label

FLOUR MODULE *If two or more types of flour are available in the household record information on the flour most frequently consumed in the household.*		
40. DID YOU HAVE FLOUR IN THE HOUSEHOLD TODAY? **YU GAT WIT FLAUA LONG HAUS TEDE?**	Yes ..... 1 No ..... 2 Don't know 9	2 ⇨ Q. 49
41. WHERE DID YOU GET THIS FLOUR? **YU BIN KISIM FLAUA WE?**	Shop ..... 1 Other (specify) ..... 8 Don't know 9	8 ⇨ Q. 49
42. PLEASE SHOW US SAMPLES OF THE FLOUR YOU BOUGHT IN THE SHOP? **PLIS SOIM MIPELA SEMPOL BILONG OLGETA WIT FLAUA YU BAIM LONG STOA** *(Observe and circle the type of flour used)*	Whole meal flour ..... 1 White flour (Plain) ..... 2 White (Self Raising) ..... 3 Don't know 9	
43. If you DO NOT see the original bag or package the flour came in *ASK* "COULD I PLEASE SEE THE ORIGINAL BAG OR PACKAGE THE FLOUR CAME IN?" **"PLIS INAP MI LUKIM PEKET FLAUA I BIN STAP INSAIT LONG EM NA YU BAIM?"**	Yes, bag observed ..... 1 No, bag not observed 2	2 ⇨ Q. 48

44. <u>Observe</u> the brand written on the flour package and circle appropriate answer	No label..... 1 Mothers Choice ..... 2 3 Roses..... 3 Flame..... 4 Other (specify) _____.. 8 Don't know..... 9
45. <u>Observe</u> the country where the flour is produced	Papua New Guinea ..... 1 Australia..... 2 India..... 3 Other (specify) _____.. 8 Don't know..... 9
46. *<u>Observe</u> the country where the flour is packaged*	Papua New Guinea ..... 1 Australia..... 2 India..... 3 Other (specify) _____.. 8 Don't know..... 9

47. <u>Observe</u>- Is the flour fortified with vitamins or minerals?	Not fortified or not stated on label ..... 1 Fortified with iron ..... 2 Fortified with folic acid ..... 3 Fortified with iron and folic acid ..... 4 Fortified with other vitamins/minerals (specify) . 5 Enriched with vitamins and minerals ..... 6 Don't know..... 9
48. DO YOU OR OTHERS FROM THIS HOUSEHOLD BUY BREAD THAT IS ALREADY MADE (NOT FROM YOUR OWN DOUGH)? **YU O OL NARAPELA LONG DISPELA HAUS I SAVE BAIM BRET WE OL I BEKIM PINIS (I NO DISPELA YU YET I MEKIM)**	Yes..... 1 No ..... 2 Don't know 9

OIL MODULE *If two or more types of oil are available in the household record information on the cooking oil most frequently consumed in the household.*		
49. DO YOU HAVE ANY OIL IN THE HOUSEHOLD NOW? **YU GAT OIL LONG HAUS NAU?**	Yes..... 1 No ……..... ……..2 Don't know..... 9	2 ⇨ Q. 57
50. WHERE DID YOU GET THIS OIL? **YU BIN KISIM WE?**	Shop ..... 1 Other (please specify) ..... 8 Don't know..... 9	8 ⇨ Q.57
51. PLEASE SHOW US SAMPLE OF THE OIL YOU BOUGHT FROM THE SHOP? **PLIS, SOIM MIPELA SEMPOL LONG OLGETA OIL YU BAIM LONG STOA.**	Observation not possible…… 1 Vegetable oil…… ..... 2 Sunflower oil…… ..... 3 Cooking oil..... 4 Coconut oil..... 5 Palm oil..... 6	

(Observe and circle the type of oil used)	Peanut oil ..... 10 Canola oil ..... 11 Olive oil ..... 12 Soy bean…..…..... ..13 Other (specify) _____.. 8 Don't know.....9	
52. If you DO NOT see the original container the oil came in or package ask "COULD I PLEASE SEE THE ORIGINAL CONTAINER OR PACKAGE THE OIL CAME IN?" **"PLIS INAP MI LUKIM ORIJINEL KONTENA O PEKET OIL I KAM LONG EN?"**	Yes, original container observed ..... 1 No, original container not observed..... 2	2 ⇨ Q. 57
53. Write the name of the brand of oil written on the package	No label or no brand ..... 9 Brand _____	9 ⇨ Q. 57
54. Observe the country where the oil is produced	Papua New Guinea ..... 1 Australia..... 2 Other (specify) _____.. 8 Don't know..... 9	
55. *Observe the* country *where the oil is packaged*	Papua New Guinea ..... 1 Australia..... 2 Other (specify) _____.. 8 Don't know…….... .9	
56. Observe – Is the oil fortified with with vitamin A?	Yes..... 1 No or not stated on label ..... 2 Don't know..... 9	

SUGAR MODULE *If two or more types of sugar are available in the household record information on the sugar most frequently consumed in the household.*		
57. DO YOU HAVE SUGAR IN THE HOUSEHOLD NOW? **YU GAT SUGA LONG HAUS NAU?**	Yes …….. ..... 1 No …….. 2 Don't know..... 9	2 ⇨ Q. 65
58. WHERE DID YOU GET THIS SUGAR? **YU BIN KISIM DISPELA SUGA WE?**	Shop ..... 1 Other (please specify) ..... 8 Don't know..... 9	8 ⇨ Q. 65
59. PLEASE SHOW US SAMPLE OF THE SUGAR YOU BOUGHT IN THE SHOP? **PLIS, SOIM SEMPOL LONG OLGETA SUGA YU BIN BAIM LONG STOA.** (*Observe and circle type of sugar used)*	Observation not possible……..……1 White sugar…….... 2 Brown sugar…… ..... 3 Dont know..... 9	

60. If you DO NOT see the original bag or package the sugar came in *ASK* "COULD I PLEASE SEE THE ORIGINAL BAG OR PACKAGE THE SUGAR CAME IN?" **"PLIS INAP INAP MI LUKIM ORIJINEL BEK O PEKET SUGA I KAM LONG EN?"**	Yes, bag observed ..... 1 No, bag not observed 2	2 ⇨ Q. 65
61. Observe the brand written on the sugar package and circle appropriate answer	No label ..... 1 4 Roses ..... 2 Ramu. ..... 3 CSR. ..... 4 Other (specify) _____ .. 8 Don't know 9	
62. Observe the country where the sugar is produced	Papua New Guinea ....1 Australia…….. .2 Other (specify) _____ .. 8 Don't know…….. .9	
63. Observe the country where the sugar is packaged	Papua New Guinea ....1 Australia…….. .2 Other (specify) _____ ....8 Don't know…….. .9	
64. Observe- Is the sugar fortified with vitamins or minerals?	Not fortified or not stated on label ..... 1 Fortified with vitamin A ..... 2 Fortified with other vitamins/minerals (specify) . 5 Don't know 9	

RICE MODULE *IF TWO OR MORE TYPES OF RICE ARE AVAILABLE IN THE HOUSEHOLD RECORD INFORMATION ON THE RICE MOST FREQUENTLY CONSUMED IN THE HOUSEHOLD.*		
65. DO YOU HAVE RICE IN THE HOUSEHOLD NOW? **YU GAT RAIS NAU LONG HAUS BILONG YU?**	Yes …… ..... 1 No …… ..... 2 Don't know..... ...9	2 ⇨ END
66. WHERE DID YOU GET THIS RICE? **YU BIN KISIM DISPELA RAIS WE?**	Shop ..... 1 Self grown..... 3 Other (specify) ..... 8 Don't know 9	3 ⇨ END 8 ⇨ END

67. PLEASE SHOW US A SAMPLE OF THE RICE YOU BOUGHT IN THE SHOP? **PLIS, SOIM MIPELA OL SEMPOL LONG OL RAIS YU BAIM LONG STOA.** (*Observe and circle type of rice used)*)	Observation not possible……..…….1 White rice…..…… ..... 2 Brown rice…….....3 Don't know..... ..9	

68. If you DO NOT see the original bag or package the rice came in ASK "COULD I PLEASE SEE THE ORIGINAL S BAG OR PACKAGE THE RICE CAME IN?" **"INAP MI LUKIM ORIJINEL BEK O PEKET RAIS I KAM LONG EN"?**	Yes, bag observed ..... 1 No, bag not observed 2	2 ⇨ END
69. Write the brand written on the rice package	No label or no brand ..... 9 Brand _____	9 ⇨ END
70. Observe the country where the rice is produced	Papua New Guinea ....1 Australia…… ... .2 India ..... 3 China ..... 4 Thailand .....5 Other (specify) _____..... ....8 Don't know…… 9	
71. Observe the country where the rice is packaged	Papua New Guinea ....1 Australia…… ... .2 India ..... 3 China ..... 4 Thailand .....5 Other (specify) _____..... ....8 Don't know…… 9	
72. Observe- Is the rice fortified with vitamins or minerals?	Not fortified or not stated on the label ..... 1 Fortified with iron ..... 2 Fortified with riboflavin ..... 3 Fortified with niacin ..... 4 Fortified with iron, riboflavin and niacin ..... 5 Fortified with various vitamins and minerals..... 6 Enriched with vitamins and minerals ..... 10 Don't know 9	

CHILD ONLY HH – Proceed to child (primary care taker data collection form) if there are eligible children (6 months to 5 years of age). If there are no eligible children in the household thank the respondent for his or her time and move on to the next house.

CHILD, MEN AND WOMEN HH – Proceed to the women, children and men data collection forms where applicable. If there are no eligible women, children or men in the household then thank the respondent and move on to the next house.

Data Entry Information Panel

(To be completed by the data entry clerks)

First Data entry clerk ID number		Second Data entry clerk ID number	

APPENDIX II: Assessment of Interaction and Confounding for CRP Model

```
*****************************************************
LOGISTIC REGRESSION FOR anemic10 = CRP
FORWARD-BUILDING OF THE MODEL
  1. Testing interaction terms one at a time
*****************************************************;


*no interaction terms (anemicyn = _crp)
************************************;


      proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            model anemic10 (event='1')= crp10 / technique=newton;
            weight smplwts;
      run;quit;
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4107.6048	1	<.0001
Score	4228.5012	1	<.0001
Wald	2.4784	1	0.1154

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.6690	0.1200	31.0934	<.0001
crp10	1	0.4270	0.2712	2.4784	0.1154

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
crp10	1.533	0.901	2.608

```
********************
Test interaction of crp and age (anemic10 = crp10 age crp10*age)
********************;

      proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class _agecat;
            model anemic10 (event='1')= crp10 _agecat _agecat*crp10/
technique=newton;
            weight smplwts;
      run;quit;

**RESULT= non-significant interaction term (p=0.59);
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	19097.4354	13	<.0001
Score	18672.0134	13	<.0001
Wald	12.5696	13	0.4816

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	2.8242	0.0929
_agecat	6	7.1893	0.3037
crp10*_agecat	6	4.6190	0.5935

```
*************************
Test interaction of crp and region (anemic = crp region crp*region)
*************************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class region;
            model anemic10 (event='1')= crp10 region crp10*region /
technique=newton;
            weight smplwts;
      run;quit;

*Result: insignificant interaction term (p=0.33)
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	250148.527	7	<.0001
Score	222435.949	7	<.0001
Wald	48.6334	7	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	4.0561	0.0440
region	3	44.0038	<.0001
crp10*region	3	3.4189	0.3314

```
********************
Test interaction of crp and urban (anemic = crp urcat10 crp*urcat10)
******************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class urcat10;
            model anemic10 (event='1')= crp10 urcat10
                                        crp10*urcat10 /
technique=newton;
            weight smplwts;
      run;quit;

*Result: significant interaction (p=0.03)
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	0.0984	0.7537
urcat10	1	5.0360	0.0248
crp10*urcat10	1	4.6760	0.0306

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.9348	0.1510	38.3189	<.0001
crp10		1	-0.1072	0.3416	0.0984	0.7537
urcat10	0	1	0.4155	0.1852	5.0360	0.0248
crp10*urcat10	0	1	0.7384	0.3415	4.6760	0.0306

```
**********************
Test interaction of crp and bmi (anemic = crp _bmicat crp*_bmicat);
**********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class _bmicat;
            model anemic10 (event='1')= crp10 _bmicat _bmicat*crp10 /
technique=newton;
            weight smplwts;
      run;quit;

*Result: no significant interaction (p=0.41)
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	49478.8337	7	<.0001
Score	48341.0456	7	<.0001
Wald	29.8671	7	0.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	2.5098	0.1131
_bmicat	3	15.5427	0.0014
crp10*_bmicat	3	2.8560	0.4144

```
********************
Test interaction of crp and hhd (anemic = crp hhdcat crp*hhdcat)
********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class hhdcat;
            model anemic10 (event='1')= crp10 hhdcat hhdcat*crp10/
technique=newton;
            weight smplwts;
      run;quit;

*Result: no interaction (p=0.54)
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	2.0885	0.1484
hhdcat	2	2.8587	0.2395
crp10*hhdcat	2	0.3894	0.8231

Analysis of Maximum Likelihood Estimates

Parameter		DF	Standard Estimate	Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.6152	0.1215	25.6230	<.0001
crp10		1	0.4030	0.2789	2.0885	0.1484
hhdcat	1	1	0.1242	0.1319	0.8869	0.3463
hhdcat	2	1	0.0810	0.1628	0.2479	0.6186
crp10*hhdcat	1	1	0.1820	0.3894	0.2183	0.6403
crp10*hhdcat	2	1	-0.2625	0.4282	0.3757	0.5399

```
********************
Test interaction of crp and smokeyn (anemic = smokeyn crp crp*smokeyn)
********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            model anemic10 (event='1')= crp10 smokeyn10
                                        crp10*smokeyn10 /
technique=newton;
            weight smplwts;
      run;quit;

*Result: significant interaction (p=0.29)
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	10998.5584	3	<.0001
Score	11294.8582	3	<.0001
Wald	7.1346	3	0.0677

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.6920	0.1225	31.9214	<.0001
crp10	1	0.7217	0.3057	5.5734	0.0182
smokeyn10	1	0.0752	0.2637	0.0813	0.7755
crp10*smokeyn10	1	-1.3536	0.6215	4.7433	0.0294

```
************************
Test interaction of crp and birth310 (anemic10 = crp10 birth3
crp*birth3)
************************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            model anemic10 (event='1')= crp10 birth310
                                        crp10*birth310/
technique=newton;
            weight smplwts;
      run;quit;


*Result: no significant interaction (p=0.75)
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15872.6141	3	<.0001
Score	15901.5012	3	<.0001
Wald	7.7681	3	0.0511

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.8913	0.1708	27.2454	<.0001
crp10	1	0.5021	0.3931	1.6317	0.2015
birth310	1	0.4721	0.1938	5.9330	0.0149
crp10*birth310	1	-0.1669	0.5330	0.0980	0.7542

```
************************
Test interaction of crp and educcatnum (anemic10 = crp10 educcatnum
crp*educcatnum)
************************;
      proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class educcatnum (ref='1');
            model anemic10 (event='1')= crp10 educcatnum
                                        crp10*educcatnum
/ technique=newton;
            weight smplwts;
run;quit;
```

```
*Result: no significant interaction (p=0.99)
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	11505.8818	7	<.0001
Score	11749.5205	7	<.0001
Wald	7.1644	7	0.4120

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	2.0831	0.1489
EDUCCATNUM	3	3.9569	0.2662
crp10*EDUCCATNUM	3	0.0955	0.9924

```
*****************************
Backwards Elimination From "Full Model" With All Individually
Significant Interaction Terms
*****************************;

      proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class _agecat _bmicat region (ref='2.00') hhdcat (ref='1')
educcatnum (ref='1')/param=ref;
      model anemic10 (event='1')=crp10 _agecat _bmicat region hhdcat
birth310 educcatnum
                                 smokeyn10 urcat10 crpsmoke crpur /
technique=newton;
      weight smplwts;
      run;quit;

  *Result: All interaction terms significant; however, interaction with
   smoking is not biologically plausible, so it will be left out of the
                            gold standard model
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	310034.089	23	<.0001
Score	264490.034	23	<.0001
Wald	216.6068	23	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	9.1942	0.0024
_agecat	6	4.0489	0.6701
_bmicat	3	4.7779	0.1888
region	3	53.7676	<.0001
hhdcat	2	1.2467	0.5362
birth310	1	2.2965	0.1297
EDUCCATNUM	3	0.7867	0.8526
smokeyn10	1	0.0356	0.8504
urcat10	1	7.3978	0.0065
crpsmoke	1	8.7716	0.0031
crpur	1	7.6418	0.0057

Analysis of Maximum Likelihood Estimates

Parameter		DF	Standard Estimate	Wald Error	Chi-Square	Pr > ChiSq
Intercept		1	-1.8519	0.9352	3.9215	0.0477
crp10		1	1.6302	0.5376	9.1942	0.0024
_agecat	1	1	-0.5846	0.5278	1.2270	0.2680
_agecat	2	1	-0.6921	0.4587	2.2761	0.1314
_agecat	3	1	-0.3542	0.5230	0.4587	0.4983
_agecat	4	1	-0.2613	0.4484	0.3396	0.5601
_agecat	5	1	-0.1420	0.4967	0.0818	0.7749
_agecat	6	1	-0.4487	0.5050	0.7894	0.3743
_bmicat	1	1	0.3977	0.7675	0.2684	0.6044
_bmicat	2	1	0.1134	0.6121	0.0343	0.8530
_bmicat	3	1	-0.4245	0.6028	0.4959	0.4813
region	1.00	1	2.4313	0.4272	32.3917	<.0001
region	3.00	1	2.8263	0.4071	48.1990	<.0001
region	4.00	1	2.2905	0.4066	31.7424	<.0001
hhdcat	2	1	0.0711	0.3220	0.0487	0.8253
hhdcat	3	1	-0.2182	0.2297	0.9023	0.3422
birth310		1	0.3589	0.2368	2.2965	0.1297
EDUCCATNUM	0	1	-0.0936	0.4355	0.0462	0.8299
EDUCCATNUM	2	1	-0.2445	0.3130	0.6104	0.4346
EDUCCATNUM	3	1	-0.1385	0.3747	0.1367	0.7116
smokeyn10		1	-0.0460	0.2437	0.0356	0.8504
urcat10		1	-1.0153	0.3733	7.3978	0.0065
crpsmoke		1	-2.5138	0.8488	8.7716	0.0031
crpur		1	-2.3123	0.8364	7.6418	0.0057

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
crp10		5.105	1.780	14.644
_agecat	1 vs 7	0.557	0.198	1.568
_agecat	2 vs 7	0.501	0.204	1.230

_agecat	3 vs 7	0.702	0.252	1.956
_agecat	4 vs 7	0.770	0.320	1.854
_agecat	5 vs 7	0.868	0.328	2.297
_agecat	6 vs 7	0.638	0.237	1.718
_bmicat	1 vs 4	1.488	0.331	6.699
_bmicat	2 vs 4	1.120	0.337	3.717
_bmicat	3 vs 4	0.654	0.201	2.132
region	1.00 vs 2.00	11.374	4.923	26.274
region	3.00 vs 2.00	16.883	7.602	37.495
region	4.00 vs 2.00	9.880	4.454	21.919
hhdcat	2 vs 1	1.074	0.571	2.018
hhdcat	3 vs 1	0.804	0.512	1.261
birth310		1.432	0.900	2.278
EDUCCATNUM	0 vs 1	0.911	0.388	2.138
EDUCCATNUM	2 vs 1	0.783	0.424	1.446
EDUCCATNUM	3 vs 1	0.871	0.418	1.814
smokeyn10		0.955	0.592	1.540
urcat10		0.362	0.174	0.753
crpsmoke		0.081	0.015	0.427
crpur		0.099	0.019	0.510

```
*****************************
Confounding Assessment

Gold Standard Model     OR: 3.00 (1.27, 7.08)
*****************************;
      proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class _agecat _bmicat region (ref='2.00') hhdcat (ref='1')
educcatnum (ref='1')/param=ref;
      model anemic10 (event='1')=crp10 _agecat _bmicat region hhdcat
birth310 educcatnum
                                 smokeyn10 urcat10 crpur /
technique=newton;
      weight smplwts;
run;quit;
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	6.3128	0.0120
_agecat	6	3.9438	0.6843
_bmicat	3	5.3841	0.1457
region	3	57.7171	<.0001
hhdcat	2	0.8364	0.6582
birth310	1	2.1520	0.1424
EDUCCATNUM	3	0.7977	0.8500
smokeyn10	1	0.9725	0.3241
urcat10	1	6.8553	0.0088
crpur	1	5.5362	0.0186

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
crp10		3.002	1.274	7.078
_agecat	1 vs 7	0.571	0.207	1.572
_agecat	2 vs 7	0.498	0.209	1.187
_agecat	3 vs 7	0.681	0.250	1.858
_agecat	4 vs 7	0.730	0.310	1.719
_agecat	5 vs 7	0.839	0.327	2.157
_agecat	6 vs 7	0.614	0.234	1.611
_bmicat	1 vs 4	1.609	0.297	8.722
_bmicat	2 vs 4	1.185	0.297	4.721
_bmicat	3 vs 4	0.680	0.176	2.629
region	1.00 vs 2.00	10.567	4.642	24.057
region	3.00 vs 2.00	16.367	7.572	35.379
region	4.00 vs 2.00	9.783	4.636	20.645
hhdcat	2 vs 1	1.029	0.540	1.959
hhdcat	3 vs 1	0.830	0.530	1.301
birth310		1.405	0.892	2.213
EDUCCATNUM	0 vs 1	0.922	0.390	2.177
EDUCCATNUM	2 vs 1	0.786	0.422	1.466
EDUCCATNUM	3 vs 1	0.858	0.407	1.810
smokeyn10		0.784	0.483	1.272
urcat10		0.371	0.176	0.779
crpur		0.159	0.034	0.735

```
**Drop educcatnum (p=0.8526);

proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class _agecat _bmicat region (ref='2.00') hhdcat (ref='1')
/param=ref;
      model anemic10 (event='1')=crp10 _agecat _bmicat region hhdcat
birth310
                                 smokeyn10 urcat10 crpur /
technique=newton;
      weight smplwts;
run;quit;

*Result: OR= 2.78 (1.24, 6.27), a 7.3% change from GS, so probably no
confounding
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	6.1083	0.0135
_agecat	6	4.4442	0.6168
_bmicat	3	4.3382	0.2272
region	3	53.3431	<.0001
hhdcat	2	0.7405	0.6906
birth310	1	2.7981	0.0944

smokeyn10	1	1.6791	0.1950
urcat10	1	7.5289	0.0061
crpur	1	5.3317	0.0209

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
crp10		2.784	1.236	6.269
_agecat	1 vs 7	0.520	0.201	1.342
_agecat	2 vs 7	0.506	0.240	1.064
_agecat	3 vs 7	0.653	0.269	1.583
_agecat	4 vs 7	0.677	0.309	1.483
_agecat	5 vs 7	0.835	0.353	1.974
_agecat	6 vs 7	0.646	0.275	1.519
_bmicat	1 vs 4	1.767	0.346	9.033
_bmicat	2 vs 4	1.194	0.315	4.530
_bmicat	3 vs 4	0.747	0.201	2.780
region	1.00 vs 2.00	9.921	4.457	22.085
region	3.00 vs 2.00	14.309	6.559	31.218
region	4.00 vs 2.00	9.306	4.429	19.553
hhdcat	2 vs 1	1.024	0.559	1.879
hhdcat	3 vs 1	0.845	0.553	1.292
birth310		1.445	0.939	2.225
smokeyn10		0.730	0.453	1.175
urcat10		0.367	0.179	0.751
crpur		0.173	0.039	0.767

```
**Drop hhdcat (p=0.69);

proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class _agecat _bmicat region (ref='2.00') /param=ref;
      model anemic10 (event='1')=crp10 _agecat _bmicat region birth310
                                 smokeyn10 urcat10 crpur /
technique=newton;
      weight smplwts;
run;quit;

*Result: OR=2.77 (1.23, 6.24), 7.73% change from GS model, drop
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	6.0535	0.0139
_agecat	6	4.4896	0.6107
_bmicat	3	4.3919	0.2221
region	3	53.6872	<.0001
birth310	1	3.0780	0.0794
smokeyn10	1	1.7666	0.1838
urcat10	1	7.4612	0.0063

crpur	1	5.1922	0.0227

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
crp10		2.770	1.230	6.236
_agecat	1 vs 7	0.522	0.204	1.337
_agecat	2 vs 7	0.508	0.243	1.066
_agecat	3 vs 7	0.657	0.273	1.583
_agecat	4 vs 7	0.676	0.311	1.473
_agecat	5 vs 7	0.847	0.364	1.973
_agecat	6 vs 7	0.649	0.277	1.519
_bmicat	1 vs 4	1.823	0.363	9.149
_bmicat	2 vs 4	1.222	0.327	4.572
_bmicat	3 vs 4	0.766	0.210	2.788
region	1.00 vs 2.00	9.878	4.447	21.943
region	3.00 vs 2.00	14.434	6.599	31.570
region	4.00 vs 2.00	9.379	4.504	19.532
birth310		1.469	0.956	2.258
smokeyn10		0.727	0.454	1.164
urcat10		0.368	0.179	0.754
crpur		0.175	0.039	0.784

```
**Drop _agecat (p=0.61) ;

proc surveylogistic data=thesis.model;
strata region;
cluster cluster;
class _bmicat region (ref='2.00') /param=ref;
model anemic10 (event='1')=crp10  _bmicat region birth310
smokeyn10 urcat10 crpur / technique=newton;
weight smplwts;
run;quit;

*Result: OR=2.96 (1.32, 6.64), 1.57, so no confounding – drop
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	6.8850	0.0087
_bmicat	3	4.1147	0.2493
region	3	53.6372	<.0001
birth310	1	2.3099	0.1286
smokeyn10	1	1.5385	0.2148
urcat10	1	8.1114	0.0044
crpur	1	5.8449	0.0156

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
crp10		2.955	1.315	6.640
_bmicat	1 vs 4	1.791	0.419	7.648
_bmicat	2 vs 4	1.253	0.383	4.096
_bmicat	3 vs 4	0.805	0.257	2.517
region	1.00 vs 2.00	10.426	4.570	23.787
region	3.00 vs 2.00	15.986	7.220	35.395
region	4.00 vs 2.00	9.871	4.694	20.759
birth310		1.369	0.913	2.053
smokeyn10		0.749	0.475	1.182
urcat10		0.363	0.181	0.729
crpur		0.178	0.044	0.722

```
**Drop _bmicat (p=0.25);


proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class region (ref='2.00') /param=ref;
      model anemic10 (event='1')=crp10  region birth310
                                 smokeyn10 urcat10 crpur /
technique=newton;
      weight smplwts;
      run;quit;

       *Result: OR=2.744 (1.23, 6.12), 8.59% change in OR from GS model.
                  Probably no confounding, so dropped.
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	6.0923	0.0136
region	3	59.1920	<.0001
birth310	1	4.7854	0.0287
smokeyn10	1	0.9388	0.3326
urcat10	1	12.0582	0.0005
crpur	1	7.9872	0.0047

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
crp10		2.744	1.231	6.117
region	1.00 vs 2.00	11.355	5.183	24.876
region	3.00 vs 2.00	17.203	8.007	36.962
region	4.00 vs 2.00	9.363	4.534	19.337
birth310		1.543	1.046	2.275
smokeyn10		0.803	0.515	1.252
urcat10		0.346	0.190	0.630
crpur		0.156	0.043	0.566

```
**Drop smokeyn (p=0.33);

      proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class region (ref='2.00') /param=ref;
      model anemic10 (event='1')=crp10  region birth310
                                 urcat10 crpur / technique=newton;
      weight smplwts;
      run;quit;

       *Result: OR = 2.744 (1.23, 6.15), 8.59% change from GS model, so
                                  dropped
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	6.0190	0.0142
region	3	57.4227	<.0001
birth310	1	4.8901	0.0270
urcat10	1	12.0622	0.0005
crpur	1	7.9432	0.0048

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
crp10		2.744	1.225	6.148
region	1.00 vs 2.00	11.635	5.350	25.300
region	3.00 vs 2.00	16.745	7.590	36.939
region	4.00 vs 2.00	9.635	4.691	19.790
birth310		1.551	1.051	2.288
urcat10		0.347	0.191	0.631
crpur		0.158	0.044	0.570

```
proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class region (ref='2.00') /param=ref;
      model anemic10 (event='1')=crp10  region
                                 urcat10 crpur / technique=newton;
      weight smplwts;
      run;quit;

       *Result: OR= 2.70 (1.22, 5.99), 9.96% change, no appreciable gain
                          in precision, so left in;
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
crp10	1	6.0054	0.0143
region	3	58.3798	<.0001
urcat10	1	13.1677	0.0003
crpur	1	7.2200	0.0072

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
crp10		2.703	1.220	5.987
region	1.00 vs 2.00	11.409	5.317	24.482
region	3.00 vs 2.00	16.724	7.672	36.453
region	4.00 vs 2.00	9.737	4.753	19.946
urcat10		0.338	0.188	0.607
crpur		0.175	0.049	0.624

APPENDIX III: Assessment of Interaction and Confounding for AGP Model

```
*****************************************************
          LOGISTIC REGRESSION FOR anemic10 = AGP
              FORWARD-BUILDING OF THE MODEL

          Testing interaction terms one at a time
*****************************************************;


*********************
No interaction terms (anemicyn = agp10)
*********************;
      proc surveylogistic data=thesis.model;
        strata region;
            cluster cluster;
            model anemic10 (event='1')= agp10 / technique=newton;
            weight smplwts;
      run;quit;
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	14271.0590	1	<.0001
Score	15017.8935	1	<.0001
Wald	7.0024	1	0.0081

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.6985	0.1141	37.4506	<.0001
agp10	1	0.8879	0.3355	7.0024	0.0081

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
agp10	2.430	1.259	4.690

```
*********************
Test interaction of agp and age (anemic10 = agp age agp*age)
*********************;

      proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class _agecat;
            model anemic10 (event='1')= agp10 _agecat _agecat*agp10/
technique=newton;
            weight smplwts;
      run;quit;

      **RESULT: not significant (p=0.058);
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	34900.5305	13	<.0001
Score	35131.5175	13	<.0001
Wald	28.7039	13	0.0072

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	3.2909	0.0697
_agecat	6	7.1516	0.3071
agp10*_agecat	6	12.1782	0.0581

```
*********************
Test interaction of agp and region (anemic = agp region agp*region)
*********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class region;
            model anemic10 (event='1')= agp10 region agp10*region /
technique=newton;
            weight smplwts;
      run;quit;



      ***RESULT=not sig (p=0.49)
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	268249.399	7	<.0001
Score	235492.205	7	<.0001
Wald	47.3453	7	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	12.2261	0.0005
region	3	41.7567	<.0001
agp10*region	3	2.4369	0.4868

```
********************
Test interaction of agp and urban (anemic = agp urcat10 agp*urcat10)
*********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class urcat10;
            model anemic10 (event='1')= agp10 urcat10
                                        agp10*urcat10 /
technique=newton;
            weight smplwts;
      run;quit;

**RESULT= not significant (p=0.989);
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	45148.3746	3	<.0001
Score	44143.3852	3	<.0001
Wald	13.0091	3	0.0046

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	0.3374	0.5613
urcat10	1	5.5039	0.0190
agp10*urcat10	1	2.7225	0.0989

```
**********************
Test interaction of crp and bmi (anemic = agp _bmicat agp*_bmicat);
**********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class _bmicat;
            model anemic10 (event='1')= agp10 _bmicat _bmicat*agp10 /
technique=newton;
            weight smplwts;
      run;quit;

      ***RESULT=significant (p<0.001);
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	60240.2404	7	<.0001
Score	56387.2126	7	<.0001
Wald	1283.0581	7	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	80.6714	<.0001
_bmicat	3	19.3508	0.0002
agp10*_bmicat	3	585.0802	<.0001

```
**********************
Test interaction of agp and hhd (anemic = agp10 hhdcat agp*hhdcat)
**********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            class hhdcat;
            model anemic10 (event='1')= agp10 hhdcat hhdcat*agp10/
technique=newton;
            weight smplwts;
      run;quit;

      ***RESULT=not significant (p=0.12);
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	26683.9150	5	<.0001
Score	27679.3808	5	<.0001
Wald	16.2344	5	0.0062

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	6.9958	0.0082
hhdcat	2	3.8358	0.1469
agp10*hhdcat	2	4.2589	0.1189

```
*********************
Test interaction of agp and smokeyn (anemic = smokeyn agp10
agp*smokeyn)
*********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            model anemic10 (event='1')= agp10 smokeyn10
                                        agp*smokeyn10 /
technique=newton;
            weight smplwts;
      run;quit;

      ***RESULT= not significant (p=0.39);
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	16829.2382	3	<.0001
Score	17660.9993	3	<.0001
Wald	8.0295	3	0.0454

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.6942	0.1189	34.0930	<.0001
agp10	1	1.0490	0.3752	7.8174	0.0052
smokeyn10	1	0.7061	0.8181	0.7449	0.3881
smokeyn10*AGP	1	-0.9287	1.0875	0.7292	0.3931

```
*********************
Test interaction of crp and birth3 (anemic10 = agp10 birth3 crp*birth3)
*********************;

proc surveylogistic data=thesis.model;
            strata region;
            cluster cluster;
            model anemic10 (event='1')= agp10 birth310
                                        agp10*birth310/
technique=newton;
            weight smplwts;
      run;quit;


***RESULT=not significant (p=0.29);
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	26959.4723	3	<.0001
Score	27443.8692	3	<.0001
Wald	14.3127	3	0.0025

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9312	0.1626	32.8132	<.0001
agp10	1	1.2051	0.4242	8.0720	0.0045
birth310	1	0.4967	0.1887	6.9248	0.0085
agp10*birth310	1	-0.6547	0.6198	1.1155	0.2909

```
*********************
Test interaction of AGP and educcatnum (amenic10 = agp10 educcatnum
agp10*educcatnum)
*********************;

      proc surveylogistic data=thesis.model;

            strata region;
            cluster cluster;
            class educcatnum (ref='1');
            model anemic10 (event='1')= agp10 educcatnum
                                        agp10*educcatnum/
technique=newton;
            weight smplwts;
      run;quit;

**result: not significant (p=0.54);
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	23762.4622	7	<.0001
Score	24716.5098	7	<.0001
Wald	11.5093	7	0.1179

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	3.8406	0.0500
EDUCCATNUM	3	4.8240	0.1852
agp10*EDUCCATNUM	3	2.1662	0.5386

```
**************************************
      Full model containing interaction term and all other covariates
**************************************;

proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class _agecat _bmicat region (ref='2.00') hhdcat (ref='1')
educcatnum (ref='1');
      model anemic10 (event='1')= agp10 _agecat _bmicat region hhdcat
birth310 educcatnum
                                  smokeyn10 urcat10 agpbmi/
technique=newton;
      weight smplwts;
      run;quit;


*Result: Interaction term (AGP*BMI) is no longer significant in this
model,  so it will not be included in the final, gold standard model;
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	305125.928	22	<.0001
Score	263942.507	22	<.0001
Wald	193.2263	22	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	0.0168	0.8968
_agecat	6	3.7501	0.7105
_bmicat	3	5.1675	0.1599
region	3	55.8213	<.0001
hhdcat	2	1.0014	0.6061
birth310	1	1.7395	0.1872
EDUCCATNUM	3	0.9748	0.8073
smokeyn10	1	1.1355	0.2866
urcat10	1	8.7654	0.0031
agpbmi	1	0.5546	0.4564

```
*****************************
      GOLD STANDARD MODEL
*****************************;

      proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class _agecat _bmicat region (ref='2.00') hhdcat (ref='1')
educcatnum (ref='1')/param=ref;
      model anemic10 (event='1')= agp10 _agecat _bmicat region hhdcat
birth310 educcatnum
                                  smokeyn10 urcat10 / technique=newton;
      weight smplwts;
      run;quit;

*OR= 3.79 (1.36, 10.56)
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	6.4911	0.0108
_agecat	6	3.5737	0.7341
_bmicat	3	4.8389	0.1840
region	3	55.2870	<.0001
hhdcat	2	0.9149	0.6329
birth310	1	1.5206	0.2175
EDUCCATNUM	3	0.9713	0.8082
smokeyn10	1	1.0714	0.3006
urcat10	1	8.8610	0.0029

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
agp10		3.790	1.360	10.562
_agecat	1 vs 7	0.571	0.210	1.552
_agecat	2 vs 7	0.507	0.216	1.190
_agecat	3 vs 7	0.708	0.266	1.887
_agecat	4 vs 7	0.752	0.323	1.750
_agecat	5 vs 7	0.775	0.304	1.978
_agecat	6 vs 7	0.580	0.228	1.480
_bmicat	1 vs 4	1.774	0.409	7.700
_bmicat	2 vs 4	1.190	0.363	3.900
_bmicat	3 vs 4	0.712	0.216	2.343
region	1.00 vs 2.00	10.161	4.564	22.625
region	3.00 vs 2.00	17.525	7.869	39.033
region	4.00 vs 2.00	10.428	4.862	22.365
hhdcat	2 vs 1	0.998	0.532	1.872
hhdcat	3 vs 1	0.812	0.513	1.283
birth310		1.326	0.847	2.077
EDUCCATNUM	0 vs 1	0.824	0.343	1.978
EDUCCATNUM	2 vs 1	0.740	0.394	1.390
EDUCCATNUM	3 vs 1	0.790	0.366	1.707
smokeyn10		0.779	0.485	1.251
urcat10		0.333	0.162	0.687

```
**Remove educcatnum (p=0.8082) ;

proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class _agecat _bmicat region (ref='2.00') hhdcat (ref='1')
/param=ref;
      model anemic10 (event='1')= agp10 _agecat _bmicat region hhdcat
birth310
                                  smokeyn10 urcat10 / technique=newton;
      weight smplwts;
      run;quit;

      *Result: OR= 3.92 (1.43, 10.77), a -3.4% change from GS OR
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	7.0227	0.0080
_agecat	6	3.7427	0.7114
_bmicat	3	4.0110	0.2603

region	3	51.6091	<.0001
hhdcat	2	0.8086	0.6674
birth310	1	1.8823	0.1701
smokeyn10	1	1.8360	0.1754
urcat10	1	9.6813	0.0019

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
agp10		3.920	1.427	10.765
_agecat	1 vs 7	0.529	0.206	1.358
_agecat	2 vs 7	0.522	0.249	1.093
_agecat	3 vs 7	0.700	0.296	1.657
_agecat	4 vs 7	0.707	0.325	1.536
_agecat	5 vs 7	0.772	0.327	1.824
_agecat	6 vs 7	0.607	0.265	1.389
_bmicat	1 vs 4	1.890	0.459	7.790
_bmicat	2 vs 4	1.194	0.381	3.747
_bmicat	3 vs 4	0.781	0.244	2.495
region	1.00 vs 2.00	9.767	4.459	21.391
region	3.00 vs 2.00	15.777	7.004	35.540
region	4.00 vs 2.00	10.184	4.723	21.960
hhdcat	2 vs 1	0.995	0.553	1.792
hhdcat	3 vs 1	0.829	0.538	1.277
birth310		1.352	0.879	2.080
smokeyn10		0.726	0.457	1.154
urcat10		0.332	0.165	0.665

```
      **Remove _agecat (p=0.7114)   ;

proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class  _bmicat region (ref='2.00') hhdcat (ref='1') /param=ref;
      model anemic10 (event='1')= agp10 _bmicat region hhdcat birth310
                                  smokeyn10 urcat10 / technique=newton;
      weight smplwts;
      run;quit;

*Result: OR= 4.135 (1.53,11.15) -9.10% change from GS OR
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	7.8677	0.0050
_bmicat	3	3.9769	0.2640
region	3	53.0406	<.0001
hhdcat	2	0.7130	0.7001
birth310	1	1.5775	0.2091
smokeyn10	1	1.7183	0.1899
urcat10	1	10.4846	0.0012

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
agp10		4.135	1.534	11.151
_bmicat	1 vs 4	1.785	0.495	6.432
_bmicat	2 vs 4	1.199	0.429	3.353
_bmicat	3 vs 4	0.787	0.280	2.213
region	1.00 vs 2.00	10.206	4.607	22.610
region	3.00 vs 2.00	17.150	7.599	38.708
region	4.00 vs 2.00	10.833	5.023	23.364
hhdcat	2 vs 1	1.017	0.563	1.837
hhdcat	3 vs 1	0.847	0.557	1.289
birth310		1.297	0.864	1.948
smokeyn10		0.740	0.472	1.160
urcat10		0.331	0.170	0.647

```
**Drop hhdcat (p=0.70);

      proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class  _bmicat region (ref='2.00') /param=ref;
      model anemic10 (event='1')= agp10 _bmicat region birth310
                                  smokeyn10 urcat10 / technique=newton;
      weight smplwts;
      run;quit;

      *Result: 4.135 (1.54, 11.09)  -9.1% change from GS OR
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	7.9583	0.0048
_bmicat	3	3.9773	0.2639
region	3	53.4197	<.0001
birth310	1	1.9084	0.1671
smokeyn10	1	1.7614	0.1845

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
agp10		4.135	1.542	11.086
_bmicat	1 vs 4	1.823	0.509	6.537
_bmicat	2 vs 4	1.225	0.440	3.409
_bmicat	3 vs 4	0.805	0.288	2.244
region	1.00 vs 2.00	10.183	4.589	22.595

region	3.00 vs 2.00	17.276	7.657	38.981
region	4.00 vs 2.00	10.912	5.095	23.372
birth310		1.324	0.889	1.970
smokeyn10		0.740	0.474	1.154
urcat10		0.332	0.170	0.650

```
      **Drop _bmicat (p=0.23);

proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class  region (ref='2.00') /param=ref;
      model anemic10 (event='1')= agp10 region birth310
                                  smokeyn10 urcat10 / technique=newton;
      weight smplwts;
      run;quit;

      *Result: OR= 3.88 (1.48, 10.18), -2.45% change from GS OR
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	7.6116	0.0058
region	3	59.1520	<.0001
birth310	1	3.8859	0.0487
smokeyn10	1	1.1555	0.2824
urcat10	1	16.0535	<.0001

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
agp10		3.883	1.481	10.178
region	1.00 vs 2.00	10.923	5.111	23.347
region	3.00 vs 2.00	18.481	8.469	40.328
region	4.00 vs 2.00	10.283	4.922	21.485
birth310		1.469	1.002	2.152
smokeyn10		0.788	0.511	1.216
urcat10		0.311	0.176	0.551

```
      **Drop smokeyn10 (p=0.28);

proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class  region (ref='2.00') /param=ref;
      model anemic10 (event='1')= agp10 region birth310
                                  urcat10 / technique=newton;
      weight smplwts;
      run;quit;

      *Result: OR=3.85 (1.46, 10.14), -1.58% change from GS OR
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	7.4476	0.0064
region	3	57.4818	<.0001
birth310	1	3.9466	0.0470
urcat10	1	16.0442	<.0001

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
agp10		3.883	1.481	10.178
region	1.00 vs 2.00	10.923	5.111	23.347
region	3.00 vs 2.00	18.481	8.469	40.328
region	4.00 vs 2.00	10.283	4.922	21.485
birth310		1.469	1.002	2.152
smokeyn10		0.788	0.511	1.216
urcat10		0.311	0.176	0.551

```
**Drop birth310 (p=0.041);

proc surveylogistic data=thesis.model;
      strata region;
      cluster cluster;
      class  region (ref='2.00') /param=ref;
      model anemic10 (event='1')= agp10 region
                                  urcat10 / technique=newton;
      weight smplwts;
      run;quit;

      *Result: OR = 3.98 (1.54, 10.26), -4.96% change from GS OR
```

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
agp10	1	8.1563	0.0043
region	3	58.1483	<.0001
urcat10	1	16.9836	<.0001

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
agp10		3.978	1.542	10.261
region	1.00 vs 2.00	11.147	5.283	23.522
region	3.00 vs 2.00	18.030	8.139	39.941
region	4.00 vs 2.00	10.742	5.159	22.366
urcat10		0.306	0.174	0.538

APPENDIX IV: Infection Category Table

Table 7: Anemia prevalence with respect to infection category, by age group, region, past pregnancy, birth in past three years, and urban/rural location

Anemia Prevalence N* (%)	No Inflammation	Incubation (elevated CRP)	Early Convalescence (elevated CRP & AGP)	Late Convalescence (elevated AGP)
OVERALL	217 (32.98%)	18 (36.38%)	15 (55.43%)	15 (53.95%)
AGE GROUP				
15-19	12 (19.12%)	1 (26.61%)	3 (63.62%)	0
20-24	46 (33.91%)	4 (48.26%)	1 (28.66%)	1 (11.17%)
25-29	48 (33.05%)	5 (31.16%)	3 (80.91%)	2 (45.75%)
30-34	42 (33.58%)	4 (41.26%)	2 (35.06%)	2 (100%)
35-39	25 (35.56%)	0	4 (66.56%)	3 (69.58%)
40-44	22 (38.22%)	0	2 (50.50%)	5 (100%)
45-49	17 (41.25%)	2 (73.11%)	0	1 (27.87%)
REGION				
Southern	72 (40.64%)	3 (33.96%)	5 (50.00%)	7 (80.34%)
Highlands	9 (6.87%)	0	4 (50.00%)	1 (16.67%)
Momase	77 (57.89%)	7 (77.78%)	3 (75.00%)	5 (83.33%)
Islands	59 (47.97%)	8 (50.00%)	3 (60.00%)	2 (66.67%)
BIRTH3				
Yes	113 (39.22%)	9 (40.43%)	8 (58.26%)	8 (47.70%)
No	63 (25.76%)	7 (46.60%)	2 (42.49%)	3 (42.49%)
URBANRURAL				
Urban	37 (21.27%)	1 (4.32%)	2 (29.39%)	0
Rural	180 (36.05%)	17 (47.93%)	13 (58.89%)	15 (63.93%)

* Weighted, taking into account complex survey design

EMORY UNIVERSITY
Institutional Review Board

January 27, 2011

Meredith Kanago
Rollins School of Public Health
1518 Clifton Road
Atlanta, GA 30322

RE: Determination: No IRB Review Required (Not "Human Subjects")
IRB00048148 – ***The Relationship Between Biomarkers of Inflammation and Anemia in Women in Papua New Guinea***
PI: Meredith Kanago

Dear Ms. Kanago:

Thank you for requesting a determination from our office about the above-referenced project. Based on our review of the materials you provided, we have determined that it does not require IRB review because it does not meet the definition(s) of "research" involving "human subjects" or the definition of "clinical investigation" as set forth in Emory policies and procedures and federal rules, if applicable. Specifically, in this project, you propose to analyze previously collected and now de-identified data from the 2005 Papua New Guinea National Micronutrient Survey. With the data set you receive, you will be unable to determine any individuals' identities.

This determination could be affected by substantive changes in the study design, subject populations, or identifiability of data. If the project changes in any substantive way, please contact our office for clarification.

Thank you for consulting the IRB.

Sincerely,

Tom Penna, MTS
IRB Analyst Assistant
This letter has been digitally signed