GENUS, **LXVIII** (No. 2), 1-27

ILIANA V. KOHLER* – PHILIP ANGLEWICZ**
HANS-PETER KOHLER*** – JOHN F. McCABE[ξ]
BEN CHILIMA[†] – BETH J. SOLDO[♀]

# Evaluating health and disease in Sub-Saharan Africa: minimally invasive collection of plasma in the Malawi Longitudinal Study of Families and Health (MLSFH)

## 1. INTRODUCTION

The collection of biomarker-based indicators of health as part of population-based samples has only recently become available for large-scale survey research, and it represents a potentially important addition to socioeconomic surveys as they can provide valuable insights into biological functions and the complex causal pathways between socioeconomic environments and health (Boerma *et al*., 2001; Crimmins and Vasunilashorn, 2011; Ewbank, 2008; Finch *et al*., 2001; Wachter, 2001; Weinstein and Willis, 2001; Weir, 2007). Numerous studies have established the relevance of biomarker collections for measuring the efficacy of immune system functions (Danesh *et al*., 2000; Kiecolt-Glaser *et al*., 1987; Ridker *et al*., 2000), malnutrition (Halterman *et al*., 2001; Nokes *et al*., 1998; van den Broek., 1998), diabetes (Rohlfing *et al*., 2000), morbidity and mortality outcomes (Crimmins and Vasunilashorn, 2011). Increasingly panel studies on aging in industrialized countries[1] are among the large-scale studies collecting biomarker-based health indicators (Crimmins *et al*., 2010; Crimmins and Vasunilashorn, 2011; McDade, 2011; Weinstein and Willis, 2001). Very few population-based studies in the developing world have followed suit, despite the considerable relevance of biomarker-based assessments of health in the context of low income countries. Exceptions include several extant isolated populations that

* Population Aging Research Center and Population Studies Center, University of Pennsylvania, Philadelphia, USA.
** School of Public Health & Tropical Medicine and Department of Global Health Systems and Development, New Orleans, USA.
*** Department of Sociology and Population Studies Center, University of Pennsylvania, Philadelphia, USA.
[ξ] Programmer/Data Analyst.
[†] Ministry of Health and Population Community Health Sciences Unit, Lilonge, Malawi, Africa.
[♀] Population Studies Center, University of Pennsylvania, Philadelphia, USA.
Corresponding author: Iliana V. Kohler; e-mail: iliana@pop.upenn.edu.

[1] Examples are the Health and Retirement Study (HRS) in the U.S. and the English Longitudinal Study of Aging (ELSA) in the UK.

are under intensive study - including the Cebu in the Philippines (McDade *et al*., 2010) and the Tsimane in Bolivia (Gurven *et al*., 2007, 2008) - to investigate the evolution of human immune response to various physiologic challenges and its relevance for human health (McDade, 2003). There are almost no population-based sources of general biomarker-based health indicators in sub-Saharan Africa (SSA). Most population-based biomarker collections in the region - for instance those conducted as part of Demographic and Health Surveys (DHS) (Measure DHS, 2011a,b) - have focused on HIV, other STDs, malaria and nutrition status (e.g., anemia). While infectious diseases - such as HIV and malaria - often attract the majority of research and NGOs attention in sub-Saharan Africa (Behrman *et al*., 2011), there is an important need to collect and analyze biomarkers of general health since chronic diseases will increasingly become relevant for understanding the health of sub-Saharan African adult populations that over the next decades will age rapidly in ways that will likely be distinct from other developing countries due to the exposure to the AIDS epidemic (Cohen and Menken, 2006; Heuveline, 2004; Merli and Palloni, 2006; Zaba *et al*., 2004). In this context, biomarker-based health indicators can provide important information on normal physiological processes and signal the presence of pathogenic processes and/or diseases (Crimmins and Vasunilashorn, 2011; McDade, 2011). They can clarify how social and behavioral factors influence health (Crimmins and Vasunilashorn, 2011), and can help to identify pathways that are important for predicting morbidity and mortality trends (McDade *et al*., 2007). Because biomarker-based health information is not subject to the same measurement errors as self-reported health indicators, biomarker-based measures of adult health may be particularly useful in surveys in developing countries where the measurement of health is problematic and little is known about disease incidence, duration, and prevalence or cause-specific morbidity and mortality among adult and elderly individuals (Boerma *et al*., 2001).

The Malawi Longitudinal Study of Families and Health (MLSFH), a longitudinal panel study in rural Malawi that provides a rare record of more than a decade's duration of demographic, socioeconomic and health conditions in one of the world's poorest countries, has augmented its data collection with biomarker-based indicators (biomarkers) of adult health in 2009. We use the term "biomarker" as defined by Crimmins and colleagues (Crimmins *et al*., 2008), that is, as an objectively measured trait that prior research has shown to be a reliable indicator of normal biologic or pathologic processes that are common to human aging. The MLSFH biomarker collection focused on blood serum indicators pertinent to the study of health and aging in the sub-Saharan African context. There are numerous biomarkers of population health that could potentially inform our understanding of adult health in rural Malawi. Those that we chose have demonstrated analytic utility in studies of both, in developed countries (Crimmins and Vasunilashorn,

2011) - notably in the U.S. (Alley *et al*., 2006; Khera *et al*., 2005; Kuo *et al*., 2006; Ridker *et al*., 2002, 2001) and Japan (Saito *et al*., 2003) - and in developing countries, including resource-poor countries such as Nigeria (Adeyinka *et al*., 1995; Drain *et al*., 2007), Amazonian Bolivia (Gurven *et al*., 2008; McDade *et al*., 2005), Mexico (Han *et al*., 2002), the Philippines (McDade *et al*., 2008), South Africa and other sub-Saharan populations (Masemola *et al*., 2007; Walker *et al*., 2004), Aboriginal Australia (McDonald *et al*., 2004), and indigenous Africans (Parkin *et al*., 2008; Sitas *et al*., 2008). The MLSFH biomarker sample is a subset of 906 MLSFH respondents and provides 12 biomarker-based indicators of general health, focusing specifically on cardiovascular risks (lipids), metabolic processes (glucose and HbA1C), organ function (creatinine, albumine, total protein, uric acid, urea/blood urea nitrogen (BUN)), and inflammation (wide-range CRP). The aims of this paper are to document the protocol for this biomarker data collection and to provide descriptive information about the study population and the collected biomarker-based indicators of health among adults in a low income SSA population. We also discuss the feasibility of biomarker data collections in less-developed contexts such as SSA. Collecting such biomarkers in surveys inevitably involves balancing the ease and cost with specimen stability and assay reliability. These concerns are exacerbated in developing countries with poor infrastructures for health care and transportation. Overcoming such obstacles is necessary if we are to improve our understanding of population health in resource-poor countries that are transitioning from an acute disease regime to one increasingly dominated by chronic conditions (Murray *et al*., 2003; Murray and Lopez, 1997). These MLSFH biomarkers provide new opportunities to study aspects of health and disease in a context characterized by a mature HIV/AIDS epidemic, high levels of poverty and high levels of morbidity and mortality.

## 2. DATA

### 2.1 *Malawi Longitudinal Study of Families and Health (MLSFH)*

The biomarker data collection is a pilot study conducted within the Malawi Longitudinal Study of Families and Health (MLSFH)[2]. MLSFH is a longitudinal panel data collection with survey waves in 1998, 2001, 2004, 2006, 2008, 2010 and 2012 that is currently focused on studying the mechanisms that individuals, families, households, and communities develop and use in a poor rural setting to cope with the impacts of high morbidity and mortality in their immediate living environment (Anglewicz *et al*., 2009; Boileau

[2] Formerly known as the Malawi Diffusion and Ideational Change Project, MDICP.

*et al*., 2009; Castro *et al*., 2010; Clark *et al*., 2009; Delavande and Kohler, 2009, 2012; Helleringer and Kohler, 2005; Kohler *et al*., 2007; Kohler and Thornton, 2012; Kohler *et al*., forthcoming; Obare *et al*., 2009). The MLSFH is implemented in three sites in rural Malawi: Rumphi (in the northern region), Mchinji (in the central region), and Balaka (in the southern region). These rural regions are similar in terms of their overall economic context that is based on subsistence agriculture; however, the study regions also reflect interesting heterogeneity in terms of marriage patterns (Reniers, 2003), religious affiliations (Trinitapoli and Regnerus, 2006), schooling (Grant, 2008; Helleringer and Kohler, 2005), patrilineal vs. matrilineal inheritance and landownership, and HIV prevalence, thus reflecting a range of socioeconomic/health/demographic conditions and permitting the evaluation of contextual effects, including risk environments, on estimated relations pertaining to health and its behavioral and socioeconomic determinants.[3] Detailed descriptions of the MLSFH sample selection, data collection, and data quality are provided on the project website http://www.malawi.pop.upenn.edu, in a Special Collection of the online journal Demographic Research that is devoted to the MLSFH (Watkins *et al*., 2003), and in a recent follow-up publication that incorporates the 2004 and 2006 MLSFH data (Anglewicz *et al*., 2009). Comparisons

[3] The MLSFH started in 1998 with a sample of 1,541 ever-married women aged 15-49 and 1,065 of their spouses. In 2001, respondents were re-interviewed, along with any new spouses since 1998. In 2004, the study added two new components to the data-collection: a new additional sample of approximately 1,500 adolescents, and free HIV testing and a voluntary counselling on the HIV test results for all respondents. The MLSFH returned for a fourth wave of survey data collection and a second round of HIV testing in 2006, and it followed-up in 2008 and 2010 with two additional rounds of wide-ranging survey data for about 4,000 respondents aged between 21 ($10^{th}$ percentile) and 67 ($90^{th}$ percentile). In 2012, the MLSFH focused on mature and elderly individuals and interviewed a sub-sample of respondents age 45 and up. The three regions studied in MLSFH are rural, and subsistence agriculture is the predominant economic activity among study participants. Important heterogeneity, however, exists with respect to marriage pattern, HIV prevalence and sexual behaviours, economic activities and education. Rumphi District, located in the northern region of the country, follows the patrilineal system of kinship and lineage where residence is ideally patrilocal, inheritance is traced through sons, and parents of a groom pay bride wealth. The northern district, inhabited primarily by Tumbukas, is predominantly Protestant. Mchinji District, located in the central region, follows a less rigid matrilineal system whereby residence may be matrilocal or patrilocal. The Center is primarily inhabited by Chewas, with almost equal proportions of Catholics and Protestants. The study population for the biomarker sample is from the southern region of Malawi (Balaka). This southern region is primarily inhabited by Lomwes and Yaos, representing a predominantly Muslim population. The region follows a matrilineal system of kinship and lineage system where residence is ideally matrilocal, although it is not uncommon for wives to live at least some period of time in their husband's village. The southern region exhibits lower ages of sexual debut and larger numbers of lifetime sexual partners than the other MLSFH study regions. Moreover, residents of southern Malawi tend to be less educated and poorer than those living in the north, and they tend to be more likely to migrate. Perhaps not surprisingly, HIV/AIDS prevalence in the southern region is significantly higher than in the northern and central region.

with the Malawi Demographic and Health Survey (DHS) showed that the MLSFH sample population is reasonably representative of the rural Malawi population (Anglewicz *et al*., 2009; Bignami-Van Assche *et al*., 2003, Watkins *et al*., 2003).

### 2.2 *Study population for MLSFH biomarker collection*

Due to our particular interest in the interaction between HIV and the biomarker measures, we chose among the three MLSFH study sites Balaka, the southern District with the highest HIV prevalence among the three MLSFH sites.[4] The MLSFH biomarker sample was collected in two stages. First, all respondents who were found HIV positive in a previous MLSFH round were included in the sample. Next, in addition we drew a random sample of approximately 1,500 respondents (aged $\geq$ 18 years) from the 2300 total respondents in the 2008 MLSFH Balaka sample. Because of weather obstacles and failed attempts to find respondents, we were able to re-contact 1,031 individuals. Of these, 49 respondents (4.7%) refused to participate, and we collected biomarker specimens for 982 respondents, of which approximately 60 cases had previously tested positive for HIV. The biomarker data collection was approved by the IRB at the University of Pennsylvania (May 9th, 2008) and by the Malawi National Health Sciences Research Council (NHSRC) (December 8th, 2008).

### 2.3 *MLSFH biomarker collection: minimally invasive collection of plasma in the field*

To date, the primary method for obtaining biomarkers in developing countries has been dried blood spots (DBS). McDade pioneered these techniques for collecting and storing small blood samples and developing assays for important biomarkers, such as hsCRP, a marker of inflammation (McDade *et al*., 2004, 2007; McDade, 2006). The infrastructure for processing DBS assays has not kept pace with the large volume of DBS obtained in large, nationally representative surveys. Moreover, the biomarkers that can be obtained from DBS remain restricted. To avoid the complications associated with DBS, the MLSFH has tested a new approach for collecting measures of population health and their adaptability to extreme conditions in tropical zones. Our results indicate the reproducibility of biomarkers obtained from the LabAnywhere (previously Demecal) system (LabAnywhere, Haarlem, The Netherlands) (Gootjes *et al*., 2009), a new system for the collection of blood plasma that has been used in other large-scale biomarker collections in developed countries (de Beer *et al*., 2009; van Dijk *et al*., 2010; van Stralen *et al*., 2011).

[4] Logistical and monetary considerations precluded conducting this data collection in all three sites.

The LabAnywhere system requires only a few drops of blood harvested from a lancet puncture of a sanitized fingertip. A sponge device is used for absorbing the drop of blood. After the sponge turns completely red, it is dropped into a container with buffer fluid. A gentle swinging motion for 40 seconds is necessary to release the dilution buffer. A filter is used to separate the red blood cells from the plasma. The distinctive feature of this system is that the blood is pressed through a patented filter that separates out plasma. Unlike a clinic based procedure for obtaining blood plasma, the LabAnywhere system does not require the use of a centrifuge. The reliability, sensitivity, and specificity of the test kits have been demonstrated by LabAnywhere in the Netherlands, and the applications of test specific recovery factors yielded a good correlation with results of venous blood samples (Gootjes *et al*., 2009). In general, LabAnywhere plasma samples are stable for 4 days at 4 °C , 2-3 days at room temperature and 1 day at 37 °C.

An important advantage of using the LabAnywhere kits for blood sampling is the minor discomfort it causes to study participants. The kits offer an ideal combination providing viable blood samples for analysis and a non-invasive means of collection. As such, the LabAnywhere method offers several advantages over the other common means of collecting blood samples such as DBS or venipuncture (McDade, 2011).[5] The LabAnywhere kit allows for the preparation of plasma from just one single drop of blood at any location, at any time (e.g., at the homes of MLSFH respondents). Up to 16 assays can be done with this plasma. Because the plasma has been diluted, the LabAnywhere analyzer technology will measure very small amounts permitting this quantity of assays. The disadvantages of the LabAnywhere method to collect and analyze blood plasma compared to other approach are summarized and discussed in more details by McDade (2011).

### 2.3.1 MLSFH biomarkers

The collected biomarkers include: a lipids panel consisting of cholesterol, LDL, HDL, and triglycerides, as measures for risk factors for cardiovascular disease; circulating blood glucose and HbA1c (only in cases when the blood glucose was above the normal range) as markers of the metabolic function; markers of organ, specifically renal function and clearance (total protein, uric acid, albumin, urea/blood urea nitrogen (BUN), and creatinine) and wide-range CRP (wrCRP) as a non-specific marker of inflammation and the immune function (Maharshak *et al*., 2008). Few, if any, biomarkers are free-standing reli-

[5] Blood sample collection through DBS has several disadvantages including potential damage and loss of viability caused by the drying of blood, and insufficient blood for testing provided on the filter paper. While intravenous blood collection avoids the problems of DBS, it involves a much more invasive procedure and requires collection by trained phlebotomists.

able diagnostic tools, and neither are the ones listed above. Although the biomarkers collected as part of the MLSFH are well-known, we briefly discuss our reasons for their selection, the critical levels used for obtaining indicators of health risks, mostly for the U.S. and similar developed contexts, and the anticipated relations of each biomarker to others we measure.

### 2.3.2 Lipids

*Total cholesterol* (*TC*)*, high-density lipoprotein* (*HDL*)*, low density lipoprotein* (*LDL*)*, and triglycerides* (*TG*): Lipids are fats that store energy for quick release, and to varying degrees, all lipids are recognized risk factors for cardiovascular disease in the developed world. In the absence of other risk factors, the American Heart Association considers a total cholesterol reading of less than 200mg/dl desirable, 200-230 mg/dl borderline, and in excess of 240mg/dl as conveying a high risk for cardiovascular disease. The optimal level of HDL in the U.S. is 50 mg/dl, but not less than 35 mg/dl, and less then 100mg/dl is recommended for LDL. Normal fasting triglyceride levels in the U.S. are below 150 mg/dl; 150-199 mg/dl is considered borderline high, 200-499 mg/dl high, and 500 mg/dl and greater very high (Adult Treatment Panel III, 2002). Because respondents in Balaka live in rural environments, with food being produced at relatively high energy expenditures but yielding few protein-dense calories, we expected that the distribution of lipids in our biomarker sample will be more concentrated at the lower end of distributions than observed in European or U.S. populations.

### 2.3.3 Metabolic processes

*Glucose and HbA1c*: Random blood glucose, also known as a non-fasting blood sugar, is a biomarker for the efficiency of the metabolic system. Glucose is the main source of energy for the body. Insulin, the hormone that cells use to metabolize the glucose, is produced in the pancreas. It is released into the blood in response to levels of circulating glucose. A random blood glucose (RBG) tests has two advantages: it does not require respondent fasting and it is less expensive. But because fasting is not a prerequisite for the test, the RBG measure is less precise. The normal range for a random blood sugar test is 70-100 mg/dl. HbA1c measures average blood sugar level for the past two to three months rather than measuring blood sugar levels at one point of time. HbA1c below 5 percent is seen as normal level and a target, although it can range from 4.5 to 6 percent. People with diabetes are characterized by elevated HbA1c levels and for them a level of about 7 percent is a target. In our data collection, HbA1c was not measured for the entire sample, but only for respondents who showed elevated blood glucose levels (i.e., 12 study participants with a mean value of HbA1c of 5.53 and 0.71 std. dev.).

### 2.3.4 Biomarkers of organ function

*Creatinine*: creatinine is one of the waste products in the blood created by the normal breakdown of muscles and circulating levels of creatinine are fairly reliable indicator of the efficacy of kidneys. Normal levels of creatinine in the blood are approximately 0.6 to 1.2 mg/dl in adult males and 0.5 to 1.1 mg/dl in adult females. In malnourished persons or those who experience weight loss or wasting, such as common in persons with HIV/AIDS or cancer, creatinine levels may be surpassed. Crimmins and colleagues also note that serum creatinine is generally less reliable than urinary creatinine (Crimmins *et al*., 2008). Any condition that impairs the function of the kidneys will increase creatinine level in the blood.

*Albumin*: like creatinine, serum albumin is used to assess renal and liver function. Albumin is the protein of highest concentration in the blood and maintains oncotic pressure of blood to prevent its leakage into tissue. The normal (U.S.) range for of albumin is 3.5 to 5.5 mg/dl. A low albumin level is correlated with inflammation and malnutrition while high levels signal dehydration. Low concentrations of albumin, even within the normal range, have been positively related to coronary disease (Djousse *et al*., 2003).

*Total protein*: unlike fats and carbohydrates, proteins are not stored in the body. They are continuously broken down (metabolized) into amino acids that are used as building blocks for other proteins. The LabAnywhere test is a rough measure of all the proteins found in the plasma, principally albumin and globulin. The normal range of the test is 6.0 to 8.3mg/dl.

*Uric Acid*: uric acid is produced in the body from purine metabolism and excreted by the kidneys. Elevated uric acid is associated with gout, starvation, metabolic syndrome or kidney stones, and decreased uric acid is associated with multiple sclerosis. Normal values of uric acid range between 3.5 and 7.2 mg/dl.

*Urea/Blood Urea Nitrogen* (*BUN*): blood carries proteins for use by cells throughout the body. After the cells use the protein, the remaining waste products are returned to the blood as urea, a compound containing nitrogen. Healthy kidneys take urea out of the blood and send it to the bladder for excretion. If kidneys are not working well, the urea stays in the blood. Normal blood contains 7 to 20 milligrams of urea per deciliter of blood. A BUN result of more than 20 mg/dl indicates that kidneys are not functioning normally. Other possible causes of an elevated BUN include dehydration, gastrointestinal bleeding or heart failure.

### 2.3.5 Biomarkers of the immune system

*C-reactive protein* (*CRP*): CRP is the most commonly used non-specific marker of inflammation and infection. As an acute-phase response protein, CRP can increase as much as 1000-fold in 24 hours. At elevated levels CPR

indicates systemic infection or tissue damage, and levels above 3.0 mg/l are generally considered as indicating a high risk for cardiovascular disease. We assayed only this biomarker of immune function because of budgetary constraints and we used the wide-range CRP (wrCRP) assay since it detects levels of CRP in the range of 0.012-16.0 mg/l, and thus is sensitive to and measures both very low and very high levels. There are few studies of CRP in low-resource countries against which to compare the distribution of CRP. In Malawi, we expected median and mean levels of CRP to be classified as high-risk range ($\geq$ 3.0 mg/l) for cardiovascular disease and related events, such as myocardial infarction, or more relevantly, advancing HIV infection (Feldman *et al*., 2003).

### 2.4 *Field procedures, training and data management*

The LabAnywhere test kits were delivered directly to Malawi in September 2008. A test-run of all field procedures for about 70 cases was completed prior to the full-scale data collection in early January 2009, including the freezing and the transport protocol, and the assays themselves. In this trial we were interested in how well the integrity of the plasma samples was maintained and the reliability of the assays themselves. The actual field work commenced in mid January and was completed by early February, 2009.

The MLSFH biomarker collectors team consisted of 25 individuals who had previously been trained by the Malawian Government in finger prick blood collection as part of HIV voluntary counseling and testing. They underwent one-week of training in the use of the LabAnywhere kits prior to beginning the biomarker collection. Training included the specific procedures for collecting samples using the LabAnywhere kit and for the proper storage and transportation of the samples to the fieldwork headquarters; training also included the ethical considerations of biomarker collection in the field, informed consent, and the requirements for protecting respondents privacy and the confidentiality of the collected data. All trained biomarker collectors were required to collect test samples that were sent to LabAnywhere for assessment. Only biomarker collectors who produced usable samples from their test kits during the training were certified by LabAnywhere and hired for the MLSFH biomarker collection on the main sample.

The specimen collection was performed after obtaining informed consent from the study participants, which involved a detailed discussion of the data collection process, the use of the biomarkers for research purposes, the risks of study participation and the limited value of the collected biomarkers for identifying any specific diseases the respondent might have. To protect the respondent's confidentiality, the specimen were marked with a special identification number known only to the MLSFH Biomarker Coordinator, and only the coor-

dinator was able to link the plasma samples with the respondents and their personal information. While in the field during the day, the collected specimen were stored in a cooler. Upon returning from the field each day, the biomarker coordinator checked all samples to verify that they were collected and labeled properly; all plasma samples were stored in a -20°C freezer until they were shipped to LabAnywhere. At the end of each week, all biomarker samples were cross-checked with field records, and sent via DHL from Malawi to the LabAnywhere laboratory in the Netherlands for testing. The samples were packed in a special cooler with ice packs provided by LabAnywhere, which were designed specifically for transporting the frozen blood samples, including minimum/maximum thermometers to monitor the cooling conditions. LabAnywhere was able to analyze 92.7%, or 910 of the 982, samples they received. None were discarded because of inadequate temperature control. Each shipment also included a list of identification numbers, so that the entry of test results by LabAnywhere preserved confidentiality of participants. After the biomarker processing was completed by LabAnywhere[6], the results were sent to MLSFH. LabAnywhere also prepared a database with the assayed values and individual IDs were mapped to the extant database maintained by the MLSFH. After eliminating outliers that are likely to be erroneous measurements (see below), the final sample includes 906 respondents with at least one valid biomarker measurement.

Upon receiving the test results, MLSFH convened an information session in all participating villages during which potential health concerns identified by the tests were discussed. Individual respondents were given the option to discuss privately their results with a health care counselor. The MLSFH also worked with local health clinics to follow up on any potential health issues that were identified by the biomarker tests. However, except for referrals to local health clinics, no specific treatments were provided as part of the MLSFH biomarker study.

### 2.5 *Analytical approach*

Some of the biomarker measures fell outside the plausible range, possibly as a result of measurement problems or instabilities of the blood samples. We identified outliers as values that were 2 interquartile ranges below the 25$^{th}$ or about the 75$^{th}$ percentile, and replaced these with missing values. In addition, CRP values of zero were replaced with missing. In total, the following outliers were identified: 3 for total cholesterol (TC), 16 for HDL, 10 for LDL, 4 for triglecerides (TG), 1 for glucose (RBG), 6 for creatinine, 23 for albumin, 14 for total protein (TP), 2 for uric acid, 5 for urea (BUN), 48 for CRP (for CRP, all

[6] Less than one week after arriving in Amsterdam, the Netherlands.

of which were zero values). Log transformations of all biomarkers are calculated since they are preferable for some of the very skewed distributions - such as for CRP - to reduce the influence of extreme values on the results.

## 3. RESULTS AND DISCUSSION

Descriptive statistics for the study population are reported in Table 1, including to standard measures such as age, sex, marital status and education, also BMI, HIV+ status, subjective health and household characteristics because they measure important heterogeneity in this population. The mean age is about 42-43 years, with about 50% of the study population older than 40. The study population is therefore relatively old, especially when compared to the median age of the Malawi population that is around 17 years. This overrepresentation of elderly individuals, which is an attractive aspect for analyses using the MLSFH biomarkers, is the result of the aging of the MLSFH panel population recruited in 1998 and the addition of a MLSFH *parent sample* in 2008 that includes all respondents' living parents. About 80% of the study population is married, and because the original MLSFH sample included women and their current husbands, slightly more men than women are married. About 70% of respondents are Muslim, with the remainder mostly belonging to various Christian dominations and only a small fraction reporting no religion. More than 1/2 of women, and close to 1/3 of men, have not attended school, and only a very small fraction has attended secondary school.

About 12-14% of the study population is underweight with a BMI below 18.5; and the vast majority has a normal BMI between 18.5 and 25. A small fraction of respondents, somewhat more for females than for males, are somewhat overweight with a BMI between 25 and 30, and virtually no respondents are obese with a BMI $\geq$ *30*. Height and/or weight, and therefore BMI, is missing for about 20% of female and 40% of male respondents.

Among women 8% and among men 5% are HIV positive (Obare, 2010; Obare *et al*., 2009). Despite the relatively high levels of morbidity and mortality experienced by rural Malawians, about 50% of women and 70% of men described their subjective health as either "very good" or "excellent", a pattern that is consistent with other low income populations (Curtis and Lawson, 2000; Idler and Benyamini, 1997; Idler *et al*., 1999; Ng *et al*., 2010; Subramanian *et al*., 2009, 2010).

More than 80% of respondents have access to potable water - mostly through a covered well or bore hole - and about 80% of respondents live in households with a pit latrine. About 15% of the houses have a metal roof, which is an indicator of relative wealth. While the majority of respondent's households has at least one mosquito net, only around 65% of these are treated with insecticide.

Table 1 – *Summary statistics for the study population*

| | Females mean (St.d.) | Males mean (St.d.) | Total mean (St.d.) |
|---|---|---|---|
| Number of observations | 571 | 335 | 906 |
| Age (in 2008) | 42.17 | 43.54 | 42.68 |
| | (17.75) | (16.87) | (17.43) |
| *Age group* | | | |
| < 30 | 0.296 | 0.307 | 0.3 |
| 30–39 | 0.205 | 0.131 | 0.178 |
| 40–49 | 0.186 | 0.152 | 0.173 |
| 50–59 | 0.144 | 0.209 | 0.168 |
| 60–69 | 0.0823 | 0.14 | 0.104 |
| 70+ | 0.0876 | 0.0597 | 0.0773 |
| Married (in 2008) | 0.762 | 0.892 | 0.809 |
| Muslim (vs Christian/other/none) | 0.691 | 0.706 | 0.696 |
| *Schooling attainment* | | | |
| No school | 0.575 | 0.32 | 0.483 |
| Primary level | 0.399 | 0.618 | 0.478 |
| Secondary level | 0.0256 | 0.0615 | 0.0386 |
| *Body Mass Index* (*BMI*) (2008) | | | |
| Underweight (BMI < 18.5) | 0.143 | 0.118 | 0.135 |
| Normal (18.5 ≤BMI < 25) | 0.75 | 0.837 | 0.777 |
| Overweight ( 25≤BMI < 30) | 0.0915 | 0.0345 | 0.0737 |
| Obese (BMI ≥ 30) | 0.0156 | 0.00985 | 0.0138 |
| BMI unknown | 0.215 | 0.396 | 0.282 |
| HIV positive | 0.0835 | 0.0496 | 0.0715 |
| *Subjective health* | | | |
| Fair/Poor | 0.158 | 0.1 | 0.136 |
| Good | 0.307 | 0.195 | 0.265 |
| Very good | 0.279 | 0.298 | 0.286 |
| Excellent | 0.256 | 0.407 | 0.312 |
| *Resp.'s houshold has* | | | |
| Access to potable water | 0.843 | 0.88 | 0.857 |
| Metal roof on house | 0.144 | 0.159 | 0.149 |
| Pit latrine | 0.782 | 0.838 | 0.802 |
| Mosquito nets | 0.816 | 0.828 | 0.821 |
| Mosquito nets treated with insecticide | 0.652 | 0.662 | 0.655 |

*Note*: Descriptive statistics are calculated for respondents with at least one valid biomarker measure.

Table 2 documents the means, std. deviations and percentiles of the biomarker-based health indicators collected as part of the MLSFH. Some noteworthy patterns of the biomarker distributions include:

*Total Cholesterol* (*TC*): the overall mean of total cholesterol is 110 mg/dl (median = 108), with an interquartile range of 88.8-131. Only the four largest observations in our sample are considered as having an elevated risk for heart disease by U.S. standards (i.e., TC > 200 mg/dl).

*High-density lipoprotein* (*HDL*): The overall mean of HDL is 32 mg/dl (median = 31), with an interquartile range of 23–39. Only about 5% of women or men have optimal levels HDL levels, defined as HDL ≥ 50 mg/dl, and the majority of the sample has levels of HDL that are too low based on U.S. clinical standards.

*Low-density lipoprotein* (*LDL*): The overall mean LDL in the overall sample is 59 mg/dl (median = 58), with an interquarile range of 42-73. By U.S. standards, only the top 5% of sampled individuals had elevated levels of LDL (LDL > 100 mg/dl).

*Triglycerides* (*TG*): The mean level of triglycerides in the MLSFH biomarker sample is 59,5 (median = 53), with an interquartile range of 35-71. The vast majority of respondents have TG levels well within the normal range, and only the top 1.4% of respondents women had triglycerides above 150 mg/dl.

*Glucose* (*RBG*): the overall mean RBG in the overall sample is 75 mg/dl (median = 68.5), with an interquarile range of 61-85. Only about 10% of respondents participants had RBG above the 100mg/dl cut-off.

*Creatinine*: the mean level of creatinine is 0.73 (median = 0.71), with an interquartile range of 0.60-0.83. All but the lowest quartile of the distribution is within the normal range by U.S. clinical standards.

*Albumin* (*ALB*): the mean and median of albumin is 3.6g/dl, with an interquartile range of 3.4-3.9. About 38% of the sample have below normal levels of albumin.

*Total protein* (*TP*): the mean for total protein is 6.89g/dl, with a median of 6.86 and an interquartile range of 6.36-7.40. All but the bottom 12% and the top 4% of the MLSFH biomarker respondents are within the normal range for TP.

*Uric Acid*: the mean level of uric acid in the MLSFH biomarker sample is 4.45 (median = 4.37), with an interquartile range of 3.70-5.21. The normal range for uric acid varies by laboratory standards, but typically is in the range of 3.6-8.3 mg/dl. In our sample, about 25% of the values are in the range 2.7-3.7, and only one observation (9.91) measured above 8.3.

*Urea/Blood Urea Nitrogen* (*BUN*): the overall mean of blood urea nitrogen in the MLSFH biomarker sample is 10.7 mg/dl (median = 10.4), with an interquarile range of 6.2-12.3. None of the MLSFH respondents has an elevated level above 20mg/dl, suggesting kidney problems.

*C-reactive protein* (*CRP*): the distribution of CRP is very skewed, with the mean level of 4.5 exceeding the median level of .70 by a factor of 6.4. The interquarile range of CRP is 0.20-2.80 mg/l. Contrary to our expectation of

Table 2 – *Summary statistics for the biomarker-based health indicators*

| | | | | Percentiles | | | | |
|---|---|---|---|---|---|---|---|---|
| | N | Mean | std. | 5th | 25th | 50th | 75th | 95th |
| Total cholesterol (TC) (mg/dl) | 904.0 | 110.4 | 29.6 | 65.6 | 88.8 | 108.1 | 131.3 | 162.2 |
| High-density cholesterol (HDL) (mg/dl) | 891.0 | 32.0 | 10.8 | 15.4 | 23.2 | 30.9 | 38.6 | 50.2 |
| Low-density cholesterol (LDL) (mg/dl) | 897.0 | 59.0 | 22.3 | 27.0 | 42.5 | 57.9 | 73.4 | 96.5 |
| Triglecerides (TG) (mg/dl) | 902.0 | 59.5 | 29.6 | 26.5 | 35.4 | 53.1 | 70.8 | 115.0 |
| Glucose (RBG) (mg/dl) | 904.0 | 75.0 | 19.5 | 52.3 | 61.3 | 68.5 | 84.7 | 113.5 |
| Creatinine (mg/dl) | 901.0 | 0.73 | 0.19 | 0.45 | 0.60 | 0.71 | 0.83 | 1.06 |
| Albumin (ALB) (g/dl) | 884.0 | 3.64 | 0.44 | 2.90 | 3.36 | 3.63 | 3.92 | 4.34 |
| Total protein (TP) (mg/dl) | 893.0 | 6.89 | 0.83 | 5.52 | 6.36 | 6.86 | 7.40 | 8.28 |
| Uric Acid (mg/dl) | 905.0 | 4.45 | 1.18 | 2.69 | 3.70 | 4.37 | 5.21 | 6.56 |
| Urea (BUN) (mg/dl) | 902.0 | 10.7 | 3.13 | 6.16 | 8.68 | 10.4 | 12.3 | 16.5 |
| C-reactive protein (CRP) (mg/dl) | 845.0 | 4.50 | 11.8 | 0.10 | 0.20 | 0.70 | 2.80 | 25.0 |

finding widespread elevated level of CRP, 21% of the respondents have CRP levels that are considered normal by the U.S. standards of 3.0 mg/l.

Table 3 describes the correlation coefficients between the different biomarkers. The correlation patterns are consistent with observations from developed countries indicating that there are significant and positive correlations among all lipid measures (TC, HDL, LDL, TG), except for TG and HDL that are negatively correlated. Glucose is positively correlated with all the lipids, except HDL, and all indicators of renal and liver function. Albumin and creatinine are positively correlated, although only albumin is inversely correlated with CRP. This is as expected because albumin is a negative acute-phase protein, and the albumin concentration falls approximately 20 percent during the inflammatory process. CRP is negatively correlated with the lipids TC, HDL and LDL, and it is positively correlated with TG. It has a weak correlation with the other biomarkers, except for albumin, with which it is slightly negatively correlated. With the exception of lipids, we observe generally a weak correlation between the biomarkers, which suggests that they are indicators of different biological processes and reflect different dimensions of the respondent's health.

Very few of the MLSFH biomarkers fall in the high risk categories as defined by western standards, and this is particularly the case of biomarkers associated with cardiovascular disease (CVD) such as total cholesterol, LDL, triglycerides, albumin and CRP. HDL is the exception, where a sizable proportion of the MLSFH sample is in the high risk group. To evaluate these findings in a comparative context, Table 4 therefore compares both the mean/median levels of total cholesterol, HDL, LDL, triglycerides, albumin and CRP in the MLSFH biomarker sample, and the percentage in the high risk category for each of these biomarkers, with the Tsimane of Bolivia, an extensively studied low income population (Gurven *et al*., 2007, 2008, 2009; McDade *et al*., 2005). The top part of the table compares the mean/median levels, for each age group, between these two populations, and the bottom part compares the percentage in the high risk categories. Despite differences in the collection of blood samples and in their processing (venous blood for the Tsimane), the age-specific patterns of the biomarkers shown above are remarkably similar. This finding lends credibility to using the LabAnywhere system rather than collecting venous blood as was done in the Tsimane study. The majority in either population fell into the U.S. high-risk range only with respect to HDL, with most respondents in both populations having too low levels.

For CRP, we can also compare the MLSFH biomarker sample to another low-CRP population of "ultramarathon runners". In this population, CRP can be markedly suppressed, independent of adiposity, with median CRP levels being less than half of the control median (Tomaszewski *et al*., 2003). The Yakut, a subsistence population in Siberia, have also been shown to have low

Table 3 – *Correlation coefficients for biomarker measure*

| | TC | HDL | LDL | TG | RBG | Creat. | ALB | TP | Uric A. | BUN | CRP |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Total cholesterol (TC) (mg/dl) | 1.00* | | | | | | | | | | |
| High-density cholesterol (HDL) (mg/dl) | 0.58* | 1.00* | | | | | | | | | |
| Low-density cholesterol (LDL) (mg/dl) | 0.89* | 0.30* | 1.00* | | | | | | | | |
| Triglecerides (TG) (mg/dl) | 0.24* | -0.14* | 0.20* | 1.00* | | | | | | | |
| Glucose (RBG) (mg/dl) | 0.11* | 0.01 | 0.11* | 0.10* | 1.00* | | | | | | |
| Creatinine (mg/dl) | 0.09* | -0.09* | 0.07* | 0.10* | 0.26* | 1.00* | | | | | |
| Albumin (ALB) (g/dl) | 0.26* | 0.23* | 0.18* | 0.00 | 0.19* | 0.26* | 1.00* | | | | |
| Total protein (TP) (mg/dl) | 0.17* | 0.03 | 0.18* | 0.17* | 0.15* | 0.22* | 0.61* | 1.00* | | | |
| Uric Acid (mg/dl) | -0.01 | -0.15* | 0.03 | 0.15* | 0.18* | 0.47* | 0.11* | 0.24* | 1.00* | | |
| Urea (BUN) (mg/dl) | 0.02 | -0.09* | -0.01 | 0.06 | 0.23* | 0.47* | 0.12* | 0.10* | 0.34* | 1.00* | |
| C-reactive protein (CRP) (mg/l) | -0.17* | -0.24* | -0.12* | 0.18* | 0.00 | 0.02 | -0.13* | 0.04 | 0.07* | 0.04 | 1.00* |

*Legend*: TC = Total cholesterol (mg/dl); HDL = High-density cholesterol (mg/dl); LDL = Low-density cholesterol (mg/dl); TG = Triglecerides (mg/dl); RBG = Glucose (mg/dl); Creat. = Creatinine (mg/dl); ALB = Albumin (g/dl); TP = Total protein (mg/dl); Uric A. = Uric Acid (mg/dl); BUN = Urea (mg/dl); CRP = C-reactive protein (mg/l).
*Note*: * p-value < 0.05.

Table 4 – *CVD risk factors*

| | *Malawi* | | | | | | *Tsimane* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Mean/median levels* | | | | | | *Mean/median levels* | | | | |
| | Age Group | | | | | | Age Group | | | | |
| | 40–49 | 50–59 | 60–69 | 70+ | Total | *N* | 40–49 | 50–59 | 60–69 | 70+ | *N* |
| *Mean values by age:* | | | | | | | | | | | |
| Body Mass Index (BMI) | 21.7 | 21.5 | 21.2 | 20.3 | 21.3 | 381 | 23.9 | 24.4 | 23.2 | 22.1 | 477 |
| Total cholesterol (TC) (mg/dl) | 111.7 | 116.7 | 123.8 | 123.1 | 117.4 | 472 | 144 | 144 | 136 | 134 | 203 |
| High-density cholesterol (HDL) (mg/dl) | 32.8 | 32.3 | 33.9 | 30.5 | 32.5 | 465 | 37 | 37 | 37 | 35 | 172 |
| Low-density cholesterol (LDL) (mg/dl) | 58.5 | 63.5 | 66.0 | 68.6 | 63.1 | 471 | 80 | 79 | 76 | 72 | 170 |
| Triglecerides (TG) (mg/dl) | 56.8 | 63.5 | 66.6 | 70.3 | 62.9 | 471 | 137 | 142 | 116 | 121 | 203 |
| Albumin (ALB) (g/dl) | 3.60 | 3.60 | 3.60 | 3.46 | 3.58 | 460 | – | – | – | – | – |
| C-reactive protein (CRP) (mg/l) | 4.52 | 5.57 | 4.93 | 6.06 | 5.17 | 448 | 9.9 | 6.8 | 7.2 | 15.1 | 205 |
| *Median values by age:* | | | | | | | | | | | |
| C-reactive protein (CRP) (mg/l) | 0.75 | 0.80 | 0.80 | 1.15 | 0.90 | 448 | 2.7 | 2.7 | 4.0 | 3.4 | 205 |
| | *Prevalence of high risk (%)* | | | | | | *Prevalence of high risk (%)* | | | | |
| | Age Group | | | | | | Age Group | | | | |
| | 40–49 | 50–59 | 60–69 | 70+ | Total | *N* | 40–49 | 50–59 | 60–69 | 70+ | *N* |
| Body Mass Index: BMI $\geq$ 30kg/m$^2$ | 3.1 | 1.7 | 2.6 | 0.0 | 2.1 | 381 | 3.1 | 8.3 | 2.3 | 1.9 | 477 |
| Total cholesterol: TC $\geq$ 240mg/dl | 0.0 | 0.7 | 0.0 | 0.0 | 0.2 | 472 | 0.0 | 0.0 | 0.0 | 0.0 | 203 |
| High-density cholesterol: HDL $<$ 40mg/dl | 81.9 | 82.4 | 73.6 | 84.3 | 80.8 | 465 | 59.7 | 68.1 | 57.7 | 68.2 | 172 |
| Low-density cholesterol: LDL $\geq$ 160mg/dl | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 471 | 0.0 | 0.0 | 0.0 | 0.0 | 170 |
| Triglecerides: TG $\geq$ 200mg/dl | 0.0 | 0.0 | 0.0 | 1.4 | 0.2 | 471 | 15.5 | 12.0 | 13.3 | 3.8 | 203 |
| Albumin: ALB $<$ 3.5g/dl | 40.4 | 44.5 | 43.5 | 49.2 | 43.6 | 460 | – | – | – | – | – |
| C-reactive protein: CRP $\geq$ 3mg/l | 22.6 | 23.8 | 26.7 | 32.4 | 25.3 | 448 | 49.0 | 45.1 | 60.0 | 53.7 | 205 |

*Notes:* Source for Tsimane data: (Gurven et al., 2009); totals for biomarkers across all age groups are not available for Tsimane data.

CRP levels (Sorensen *et al*., 2006). In the Yakut, the median CRP is 0.76, compared to a median CRP of 0.70 in the MLSFH biomarker sample. Comparing the MLSFH biomarker sample CRP levels to a range of countries from very modern such as the US and UK to the Yakut and Brazil (Snodgrass *et al*., 2007), the MLSFH median CRP level of 0.70 is clearly the lowest, except for Japan with a median of 0.16 mg/L for men and 0.09 among women.

## 4. CONCLUSION

The MLSFH biomarker sample makes a potentially important contribution to the different populations in low income countries for which biomarker-based health indicators are available. The present study confirms that the collection of such biomarkers using the LabAnywhere system is feasible in rural sub-Saharan contexts: refusal rates to the biomarker collection was very low in the MLSFH, which is in sharp contrast to developed countries; following the procedures described above, only a small fraction of the biomarker samples could not be analyzed by LabAnywhere. Providing a broader range of biomarkers and being logistically easier, our approach represents an attractive alternative to DBS and venous blood samples.

Several important questions arise from the MLSFH biomarker collection and analyses that need to be more carefully addressed in future research. For example, the generally very low proportion of respondents that are classified as "high risk" based on any of the collected biomarkers is very low. The implications of this finding are currently not well understood. Do, for instance, these inter-country differences in CRP indicate that Malawians are in better health then the Yakuts, or Aborigines (McDonald *et al*., 2004), or in worse health than the Japanese populations (Yamada *et al*., 2001)? More likely, the results suggest that the conventional critical values, which are mostly validated for developed countries, are not applicable for low income SSA contexts. Many of the MLSFH biomarkers are standard blood work commonly ordered for mid-life adults. These are important measures for comparative purposes, but the environment in which participants live can amend their interpretation. In the context of Malawi, low levels of total protein are more likely to signal inadequate nutrition, rather than renal disease. Enteric infections in Malawi, rare in the U.S., can impair intestinal transport of nutrients regardless of diet or anemia in Malawi. Low levels of circulating glucose may be an important adaptation to unpredictable energy inputs rather than a symptom of pathology, per se. For example, very low HDL has been observed with acute or chronic inflammatory response, as indicated by elevated levels of mean hsCRP and reduced albumin (Schifferli, 2007). Changes in lipoproteins are noted to occur during the acute-phase reaction to inflammation. Similarly, inflammation and acute phase proteins may alter/reverse cholesterol transport by HDL (van der Westhuyzen *et al*., 2007). Increasingly, HDL is also thought of as a component of

the innate immune system because of its capacity to efflux cholesterol during the acute phase response (Jahangiri *et al*., 2009). In addition, a possible survivor effect may play important role: in the Malawian environment characterized by high disease and poverty loads, individuals surviving into their 40s in a context, where period life expectancy is about 47, may have a health advantage that is reflected in the collected measures. Future research using the MLSFH biomarker data will help to answer these questions, as well as help to better understand health and disease, and their relation to social, demographic and environmental risk factors, in Malawi and sub-Saharan Africa generally.

Limitations of the MLSFH biomarker data collection include that the present study was conducted in only one area of rural Malawi, Balaka in the south of the country. This region is predominantly Muslim and has the highest HIV/AIDS prevalence among the three MLSFH sites. Without evidence from other regions in Malawi, it is difficult to assess to which extent these data reflect an overall health pattern for the rural population in the country. For instance, Muslim populations are characterized by different dietary and life style habits, and evidence suggests that they are characterized by different health and mortality patterns compared to Christian populations (Kohler and Preston, 2011). Thus, it is reasonable to expect differences in biomarker characteristics between the different Malawian regions and ethnic groups. Despite existing cultural, dietary and life style differences, people in rural Malawi live at large in high poverty and have been exposed to a high risk disease environment since birth. These latter factors may be more powerful determinants of health patterns in a poor context compared to cultural, dietary and life style norms. Based on this pilot biomarker data collection, we are not able to distinguish between the determinants of health as reflected by biomarkers in Malawi.

*Competing Interests*: the authors have no competing interests (financial, personal or otherwise).

## Acknowledgements

We gratefully acknowledge the generous support that this study has received through the National Institute of Child Health and Development (grant numbers R03HD058976, R21HD050652, R01HD044228, R01HD053781) and the Population Aging Center (NICHD P30-AG-012836) and the Population Studies Center (NIA R24-HD-044964) at the University of Pennsylvania.

## References

ADEYINKA A.O., OHAERI J.U. OSUNTOKUN B.O. (1995), "The correlates of cortical sulci atrophy and central atrophy in a cohort of schizophrenia and manic subjects in Nigeria", *Canadian Journal of Psychiatry*, 40(4): 200-2004.

ADULT TREATMENT PANEL III (2002), "Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Final Report", *Circulation* 106(25): 3143-3421.

ALLEY D.E., SEEMAN T.E., KI KIM J., KARLAMANGLA A., HU P., CRIMMINS E.M. (2006), "Socioeconomic status and C-reactive protein levels in the US population: NHANES IV", *Brain, Behavior, and Immunity* 20(5): 498-504.

ANGLEWICZ P., ADAMS J., OBARE F., KOHLER H-P., WATKINS S. (2009), "The Malawi Diffusion and Ideational Change Project 2004-06: Data collection, data quality and analyses of attrition", *Demographic Research,* 20(21): 503-540. doi:10. 4054/DemRes.2009.20.21.

BEHRMAN J.R., BEHRMAN J.A., PEREZ N. (2011), "Out of sync? Social science research on health and development and health conditions in developing countries", *Demographic Research,* 24(2):45-78, doi:/10.4054/DemRes.2011.24.2.

BIGNAMI-VAN ASSCHE S., RENIERS G., WEINREB A.A. (2003), "An assessment of the KDICP and MDICP data quality: Interviewer effects, question reliability and sample attrition", *Demographic Research,* Special Collection 1(2): 31-76. http://www.demographic-research.org.

BOERMA J.T., HOLT E., BLACK R. (2001), "Measurement of biomarkers in surveys in developing countries: Opportunities and problems", *Population and Development Review*, 27(2): 303-314.

BOILEAU C., CLARK S., POULIN M., BIGNAMI-VAN ASSCHE S., RENIERS G., WATKINS S.C., KOHLER H-P., HEYMANN J. (2009), "Sexual and marital trajectories and HIV infection among ever-married women in rural Malawi", *Sexually Transmitted Infections,* 85 (Supplement 1): i27-i33, doi:10.1136/sti.2008.033969.

CASTRO R., BEHRMAN J.R., KOHLER H-P. (2010), "Perception of HIV risk and the quantity and quality of children: The case of rural Malawi", University of Pennsylvania, Population Studies Center, PSC Working Paper No.23 (2010-08), http://repository.upenn.edu/psc_working_papers/23/.

CLARK S., POULIN M., KOHLER H-P. (2009), "Marital aspirations, sexual behaviors, and HIV/AIDS in rural Malawi", *Journal of Marriage and the Family,* 71(2): 396-416. doi:10.1111/j.1741-3737.2009.00607.x.

COHEN B., MENKEN J. (2006), *Aging in Sub-Saharan Africa: Recommendations for Further Research*. National Academy Press, Washington, D.C.

CRIMMINS E., KIM J., VASUNILASHORN S. (2010), "Biodemography: New approaches to understanding trends and differences in population health and mortality", *Demography*, 47 (Supplement 1), S41-S64, doi:10.1353/dem.2010.0005.

CRIMMINS E., VASUNILASHORN S., KIM J.K., AALLEY D. (2008), "Chapter 5: Biomarkers related to aging in human populations", in MAKOWSKI G. S., *Advances in Clinical Chemistry*, Amsterdam The Netherlands, Elsevier, vol. 46, 161-216.

CRIMMINS E.M., VASUNILASHORN S. (2011), "Links between biomarkers and mortality", in ROGERS R.G., CRIMMINS E.M., *International Handbook of Adult Mortality*, Amsterdam, Netherlands: Springer, *International Handbooks of Population*, vol. 2, 381-398, doi:10.1007/978-90-481-9996-9_18.

CURTIS S., LAWSON K. (2000), "Gender, ethnicity and self-reported health: the case of African-Caribbean populations in London", *Social Science & Medicin,* 50(3): 365-385.

DANESH J., WHINCUP P., WALKER M., LENNON L., THOMSON A., APPLEBY P. (2000), "Low grade inflammation and coronary heart disease: Prospective study and updated meta-analysis", *British Medical Journal*, 321: 199-204.

de BEER M., van EIJSDEN M., VRIJKOTTE T., GEMKE R. (2009), "Early growth patterns and cardiometabolic function at the age of 5 in a multiethnic birth cohort: the ABCD study", *BMC Pediatrics,* 9(1): 23, doi:10.1186/1471-2431-9-23.

DELAVANDE A., KOHLER H-P. (2009), "Subjective expectations in the context of HIV/AIDS in Malawi", *Demographic Research*, 20(31): 817-874, doi:10.4054/ DemRes.2009.20.31.

DELAVANDE A., KOHLER H-P. (2012), "The impact of HIV testing on subjective expectations and risky behavior in Malawi", *Demography,* 49(3): 1011-1036, doi:10. 1007/s13524-012-0119-7.

DJOUSSE L., ROTHMAN K.J., CUPPLES L.A., ARNETT D.K., ELLISON R.C. (2003), "Relation between serum albumin and carotid atherosclerosis: the NHLBI Family Heart Study", *Stroke,* 34(1): 53.

DRAIN P.K., KUPKA R., MSAMANGA G.I., URASSA W., MUGUSI F., FAWZI W.W. (2007), "C-reactive protein independently predicts HIV-related outcomes among women and children in a resource-poor setting", *AIDS* 21, 2067-2075.

EWBANK D.C. (2008), "Biomarkers in social science research on health and aging: A review of theory and practice", in WEINSTEIN M., VAUPEL J.W., WACHTER K.W., *Biosocial Surveys*, National Academies Press, Washington, D.C,156-171.

FELDMAN J., GOLDWASSER P., HOLMAN S., DEHOVITZ J., MINKOFF H. (2003), "C-reactive protein is an independent predictor of mortality in women with HIV-1 infection", *JAIDS Journal of Acquired Immune Deficiency Syndromes,* 32(2): 210-214.

FINCH C.E., VAUPEL J.W., KINSELLA K. (eds.) (2001), *Cells and Surveys: Should Biological Measures Be Included in Social Science Research,* National Research Council (U.S.), Committee on Population, Washington, D.C.

GOOTJES J., TEL R.M., BERGKAMP F.J.M., GORGELS J.P.M.C. (2009), "Laboratory evaluation of a novel capillary blood sampling device for measuring eight clinical chemistry parameters and HbA1c", *Clinica Chimica Acta,* 401(1-2): 152-157.

GRANT M.J. (2008), "Children's school participation and HIV/AIDS in rural Malawi: The role of parental knowledge and perceptions", *Demographic Research*, 19(45): 1603-1634. http://www.demographic-research.org.

GURVEN M., KAPLAN H., SUPA A.Z. (2007), "Mortality experience of Tsimane Amerindians of Bolivia: Regional variation and temporal trends", *American Journal of Human Biology*, 19(3): 376-398, doi:10.1002/ajhb.20600.

GURVEN M., KAPLAN H., WINKING J., FINCH C., CRIMMINS E.M. (2008), "Aging and inflammation in two epidemiological worlds", *Journal of Gerontology Series A: Biological Sciences and Medical Sciences,* 63(2): 196-199.

GURVEN M., KAPLAN H., WINKING J., RODRIGUEZ D., VASUNILASHORN S., KIM J., FINCH C., CRIMMINS E. (2009), "Inflammation and infection do not promote arterial aging and cardiovascular disease risk factors among lean horticulturalists", *PLoS One* 4(8): e6590, doi:doi:10.1371/journal.pone.0006590.

HALTERMAN J.S., KACZOROWSKI J.M., ALIGNE C.A., AUINGER P., SZILAGYI P.G. (2001), "Iron deficiency and cognitive achievement among school-aged children and adolescents in the United States", *Pediatrics,* 107(6): 1381-1386.

HAN TS., SATTAR N., WILLIAMS K., GONZALES-VILLALPANDO C., LEAN M. E.J., HAFFNER S.M. (2002), "Prospective Study of C-Reactive Protein in Relation to the Development of Diabetes and Metabolic Syndrome in the Mexico City Diabetes Study, *Diabetes Care*, 25(11): 2016-2021.

HELLERINGER S., KOHLER H-P. (2005), "Social networks, risk perceptions and changing attitudes towards HIV/AIDS: New evidence from a longitudinal study using fixed-effect estimation", *Population Studies* 59(3): 265-282. doi:10.1080/ 00324720500212230.

HEUVELINE P. (2004), "Impact of the HIV epidemic on population and household structure: The dynamics and evidence to date", *AIDS* 18(Supplement 2): S45-S53.

IDLER E.L., BENYAMINI Y. (1997), "Self-rated health and mortality: a review of twenty-seven community studies", *Journal of Health and Social Behavior,* 38, 21-37.

IDLER E.L., HUDSON S.V., LEVENTHAL H. (1999), "The meaning of self-ratings of health: A qualitative and quantitative approach", *Research on Aging,* 21(3): 458-476.

JAHANGIRI A., de BEER M.C., NOFFSINGER V., TANNOCK, L.R., RAMAIAH C., WEBB N.R., van der WESTHUYZEN D.R., de BEER F.C. (2009), "HDL remodeling during the acute phase response", *Arteriosclerosis, Thrombosis and Vascular Biology,* 29: 261-267, doi:10.1161/ATVBAHA.108.178681.

KHERA A., MCGUIRE D.K., MURPHY S.A., STANEK H.G., DAS S.R., VONGPATANASIN W., WIANS JR F.H., GRUNDY S.M., DE LEMOS J.A. (2005), "Race and gender differences in C-reactive protein levels", *Journal of the American College of Cardiology,* 46(3): 464-469.

KIECOLT-GLASER J.K., GLASER R., DYER C., SHUTTLEWORTH E., OGROCKI P., SPEICHER C.E. (1987), "Chronic stress and immunity in family caregivers for Alzheimer's disease victims", *Psychosomatic Medicine,* 49: 523-535.

KOHLER H-P., BEHRMAN J.R., WATKINS S.C. (2007), "Social networks and HIV/AIDS risk perceptions", *Demography,* 44(1): 1-33. doi:10.1353/dem.2007.0006.

KOHLER H-P., THORNTON R. (2012), "Conditional cash transfers and HIV/AIDS prevention: Unconditionally promising?", *World Bank Economic Review,* 26(2): 165-190. doi:10.1093/wber/lhr041.

KOHLER I.V., KOHLER H-P., ANGLEWICZ P., BEHRMAN J.R. (forthcoming), "Intergenerational transfers in the era of HIV/AIDS: Evidence from rural Malawi", *Demographic Research,* Earlier version is available as PSC Working Paper #27 (2011-03) at http://repository.upenn.edu/psc_working_papers/27/.

KOHLER I.V., PRESTON S.H. (2011), "Ethnic and religious differentials in bulgarian mortality, 1993–98", *Population Studies,* 65(1): 91-113. doi: 10.1080/00324728.2010. 535554.

KUO H.K., BEAN J.F., YEN C.J., LEVEILLE S.G. (2006), "Linking C-reactive protein to late-life disability in the National Health and Nutrition Examination Survey (NHANES) 1999–2002", *Journal of Gerontology Series A: Biological Sciences and Medical Sciences,* 61(4): 380.

MAHARSHAK N., ARBEL Y., GAL-OZ A., ROGOWSKI O., SHAPIRA I., BERLINER S., VERED Y., CANAANI J., DOTAN I. (2008), "Comparative analysis of Bayer wide-range C-reactive protein (WR-CRP) and the Dade-Behring high sensitivity C-reactive protein (hs-CRP) in patients with inflammatory bowel disease", *Journal of Digestive Diseases,* 9(3): 140-143.

MASEMOLA M.L., ALBERTS M., URDAL P. (2007), "Apolipoprotein E genotypes and their relation to lipid levels in a rural South African population", *Scandinavian Journal of Public Health*, 35: 60-65.

McDADE T. (2011), "The state and future of blood-based biomarkers in the Health and Retirement Study", *Forum for Health Economics & Policy*, 14(3), http://www.bepress.com/fhep/14/3/5.

McDADE T., BURHOP J., DOHNAL J. (2004), "High-sensitivity enzyme immunoassay for C-reactive protein in dried blood spots", *Clinical Chemistry*, 50(3): 652-654.

McDADE T., RUTHERFORD J., ADAIR L., KUZAWA C. (2010), "Early origins of inflammation: microbial exposures in infancy predict lower levels of c-reactive protein in adulthood", *Proceedings of the Royal Society B: Biological Sciences*, 277(1684): 1129.

McDADE T., WILLIAMS S., SNODGRASS J.J. (2007), "What a drop can do: dried blood spots as a minimally invasive method for integrating biomarkers into population-based research", *Demography*, 44(4): 899-925.

McDADE T.W. (2003), "Life history theory and the immune system: steps toward a human ecological immunology", *American Journal of Physical Anthropology*, Suppl 37, 100-125, doi:10.1002/ajpa.10398.

McDADE T.W. (2006), "Measuring immune function: Markers of cell-mediated immunity and. inflammation in dried blood spots", in ICE G.H., JAMES G.D., *Measuring Stress in Humans: A Practical Guide for the Field*, Cambridge University Press, Cambridge, forthcoming.

McDADE T.W., LEONARD W.R., BURHOP J., REYES-GARCIA V., VADEZ V., HUANCA T., GODOY R.A. (2005), "Predictors of C-reactive protein in Tsimane 2 to 15 year-olds in lowland Bolivia", *American Journal of Physical Anthropology*, 128(4): 906-13.

McDADE T.W., RUTHERFORD J.N., ADAIR L., KUZAWA C. (2008), "Adiposity and pathogen exposure predict C-reactive protein in Filipino women", *Journal of Nutrition*, 138(12): 2442.

McDONALD S., MAGUIRE G., DUARTE N., WANG X., HOY W. (2004), "C-reactive protein, cardiovascular risk, and renal disease in a remote Australian Aboriginal community", *Clinical Science*, 106: 121-128.

MEASURE DHS (2011a), Biomarker's corner: Using biomarkers to collect health data. Measure DHS, Calverton, MD 20705, http://www.measuredhs.com/aboutsurveys/biomarkers.

MEASURE DHS (2011b). Demographic and Health Surveys (DHS). Measure DHS, Calverton, MD 20705, http://www.measuredhs.com/aboutsurveys/dhs/.

MERLI G.M., PALLONI A. (2006), "The HIV/AIDS epidemic, kin relations, living arrangements and the elderly in South Africa", in COHEN, B., MENKEN J., *Aging in Sub-Saharan Africa: Recommendations for Furthering Research*, National Academy Press, Washington, D.C.

MURRAY C.J.L., FERGUSON B.D., LOPEZ A.D., GUILLOT M., SALOMON J.A., AHMAD O., (2003), "Modified logit life table system: principles, empirical validation, and application", *Population Studies*, 57(2): 165-182.

MURRAY C.J.L., LOPEZ A.D. (1997), "Mortality by cause for eight regions of the world: Global Burden of Disease Study" *The Lancet*, 349: 1269-1276.

NG N., KOWAL P., KAHN K., NAIDOO N., ABDULLAH S., BAWAH A., BINKA F., CHUC N.T.K., DEBPUUR C., EGONDI T., XAVIER GOMEZ-OLIVE F., HAKIMI M., HIRVE S., HODGSON A., JUVEKAR S., KYOBUTUNGI C., Van MINH H., MWANYANGALA M.A., NATHAN R., RAZZAQUE A., SANKOH O., KIM STREATFIELD P., THOROGOOD M., WALL S., WILOPO S., BYASS P., TOLLMAN S.M., CHATTERJI S. (2010), "Health inequalities among older men and women in Africa and Asia: evidence from eight Health and Demographic Surveillance System sites in the INDEPTH WHO-SAGE study", *Global Health Action*, 3: 1654-9880.

NOKES C., van den BOSCH C., BUNDY D.A.P. (1998), "The effects of iron deficiency and anemia on mental and motor performance, educational achievement, and behavior in children: An annotated bibliography", A Report of the International Nutritional Anemia Consultative Group, http://www.idpas.org/pdf/119AEffectsofIronDeficiency.pdf.

OBARE F. (2010), "Nonresponse in repeat population-based voluntary counseling and testing for HIV in rural Malawi", *Demography*, 47(3): 651–665, doi:10.1353/dem.0. 0115.

OBARE F., FLEMING P., ANGLEWICZ P., THORNTON R., MARTINSON F., KAPATUKA A., POULIN M., WATKINS S. C., KOHLER H-P. (2009), "Acceptance of repeat population-based voluntary counseling and testing for HIV in rural Malawi", *Sexually Transmitted Infections*, 85(2): 139-144, doi:10.1136/sti.2008.030320.

PARKIN D.M., SITAS F., CHIRENJE M., STEIN L., ABRATT R., WABINGA H. (2008), "Part I: Cancer in indigenous Africans–burden, distribution, and trends", *Lancet Oncology,* 9(7): 683-692.

RENIERS G. (2003), "Divorce and remarriage in rural Malawi", *Demographic Research*, Special Collection 1(6): 175-206, http://www.demographic-research.org.

RIDKER P., HENNEKENS C., BURING J., RIFAI N. (2000), "C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women". *New England Journal of Medicine*, 342: 836-843.

RIDKER P.M., RIFAI N., ROSE L., BURING J.E., COOK N.R. (2002), "Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events", *New England Journal of Medicine*, 347(20): 1557-1565.

RIDKER P.M., STAMPFER M., RIFAI N. (2001), "Novel risk factors for systemic atherosclerosis", *JAMA, the Journal of the American Medical Association*, 285(19): 2481.

ROHLFING C.L., LITTLE R.R., WIEDMEYER H.M., ENGLAND J.D., MADSEN R., HARRIS M., FLEGAL K., EBERHARDT M., GOLDSTEIN D. (2000), "Use of GHb (HbA1c) in screening for undiagnosed diabetes in the US population", *Diabetes Care*, 23: 187-191.

SAITO M., ISHIMITSU T., MINAMI J., ONO H., OHRUI M., MATSUOKA H. (2003), "Relations of plasma high-sensitivity C-reactive protein to traditional cardiovascular risk factors", *Atherosclerosis*, 167(1): 73-79.

SCHIFFERLI J.A. (2007), "Case 40-2006: Anemia and Low HDL Cholesterol", *New England Journal of Medicine*, 356: 1893-1895, http://www.nejm.org/doi/full/10.1056/NEJMc070139#t=article.

SITAS F., PARKIN D.M., CHIRENJE M., STEIN L., ABRATT R., WABINGA H. (2008), "Part II: Cancer in indigenous Africans–causes and control", *Lancet Oncology*, 9(8): 786-795.

SNODGRASS J.J., LEONARD W.R., TARSKAIA L.A., MCDADE T.W., SORENSEN M.V., ALEKSEEV V.P., KRIVOSHAPKIN V.G. (2007), "Anthropometric correlates of C-reactive protein among indigenous Siberians", *Journal of Physiological Anthropology*, 26(2): 241-246.

SORENSEN M., LEONARD W., TARSKAYA L., IVANOV K., SNODGRASS J., ALEKSEEV V., KRIVOSHAPKIN V., RIFAI N. (2006), "High-sensitivity C-reactive protein, adiposity, and blood pressure in the Yakut of Siberia", *American Journal of Human Biology*, 18(6): 766-775.

SUBRAMANIAN S., SUBRAMANYAM M., SELVARAJ S., KAWACHI I. (2009), "Are self-reports of health and morbidities in developing countries misleading? evidence from India", *Social Science & Medicine*, 68(2): 260-265.

SUBRAMANIAN S.V., HUIJTS T., AVENDANO M. (2010), "Self-reported health assessments in the 2002 World Health Survey: how do they correlate with education?", *Bulletin of the World Health Organization*, 88(2): 131-138, doi:{10. 2471/BLT.09.067058}.

TOMASZEWSKI M., CHARCHAR F.J., PRZYBYCIN M., CRAWFORD L., WALLACE A.M., GOSEK K., LOWE G.D., ZUKOWSKA-SZCZECHOWSKA E., GRZESZCZAK W., SATTAR N., *et al*. (2003), "Strikingly low circulating CRP concentrations in ultramarathon runners independent of markers of adiposity: How low can you go?", *Arteriosclerosis, Thrombosis and Vascular Biology*, 23(9): 1640.

TRINITAPOLI J., REGNERUS M.D. (2006), "Religious involvement and HIV/AIDS risk in rural Malawi", *Journal for the Scientific Study of Religion*, 45: 505-528.

van den BROEK N.R., WHITE S.A., NEILSON J.P. (1998), "The association between asymptomatic HIV infection and the prevalence and severity of anemia in pregnant Malawian women", *American Journal of Tropical Medicine and Hygiene*, 59: 1004-1007.

van der WESTHUYZEN D., de BEER F., WEBB N. (2007), "HDL cholesterol transport during inflammation", *Current Opinion in Lipidology*, 18(2): 147-151.

van DIRK A., van EIJSDEN M., STRONKS K., GEMKE R., VRIJKOTTE T. (2010), "Cardio-metabolic risk in 5-year-old children prenatally exposed to maternal psychosocial stress: the abcd study", *BMC Public Health,* 10(1): 251, doi:10.1186/ 1471-2458-10-251.

van STRALEN M., TE VELDE S., SINGH A., DE BOURDEAUDHUIJ I., MARTENS M., van der SLUIS M., MANIOS,Y., GRAMMATIKAKI E., CHINAPAW M., MAES L., BERE E., JENSEN J., MORENO L., JAN N., MOLNAR D., MOORE H., BRUG J. (2011), "European Energy Balance Research to prevent excessive weight gain among youth (ENERGY) project: Design and methodology of the ENERGY cross-sectional survey", *BMC Public Health*, 11(1): 65. doi:10.1186/1471-2458-11-65.

WACHTER K. (2001), "Biosocial opportunities for Surveys", in FINCH C.E., VAUPEL, J.W., KINSELLA K., *Cells and Surveys: Should Biological Measures Be Included in Social Science Research*, National Research Council (U.S.), Committee on Population, Washingon, D.C., 329-336.

WALKER A.R.P., WALKER B.F., SEGAL I. (2004), "Some puzzling situations in the onset, occurrence and future of coronary heart disease in developed and developing populations, particularly such in sub-Saharan Africa", *Journal of the Royal Society for the Promotion of Health*, 124(1): 40.

WATKINS S., BEHRMAN J.R., KOHLER H-P., ZULU E.M. (2003), "Introduction to "Research on demographic aspects of HIV/AIDS in rural Africa", *Demographic Research*, Special Collection 1(1): 1-30, doi:10.4054/DemRes.2003.S1.1.

WEINSTEIN M., WILLIS R. (2001), "Stretching Social Surveys to Include Bioindicators: Possibilities for the Health and Retirement Study, Experience from the Taiwan Study of the Elderly", in FINC C.E., VAUPEL J.W., KINSELLA K., *Cells and Surveys: Should Biological Measures Be Included in Social Science Research*, National Research Council (U.S.), Committee on Population, Washingon, D.C., 250-275.

WEIR D. (2007), "Elastic powers: The integration of biomarkers into the Health and Retirement Study", in COUNCIL N.R., *Biosocial Surveys*. National Academies Press, Washington, D.C., 78-95.

YAMADA S., GOTOH T., NAKASHIMA Y., KAYABA K., ISHIKAWA S., NAGO N., NAKAMURA Y., ITOH Y., KAJII E. (2001), "Distribution of serum C-reactive protein and its association with atherosclerotic risk factors in a Japanese population", *American journal of Epidemiology*, 153(12): 1183.

ZABA B., WHITESIDE A., BOERMA J.T. (2004), "Demographic and socioeconomic impact of AIDS: Taking stock of the empirical evidence" *AIDS* 18(Supplement 2): S1–S7, doi:10.1097/00002030-200406002-00001.