The effect on cardiovascular risk factors of migration from rural to urban areas in Peru

Juan Jaime Miranda Montero

Submitted in part fulfilment of the requirements for the degree of PhD in Epidemiology

London School of Hygiene and Tropical Medicine

December 2008

Total word count*

82,883 words

Chapters I-XI word count*

69,652 words

Note: word count includes tables and figures, not only text

`

Dedicated to

Claudia and Andrea, my wife and daughter,

with love and grateful for their patience.

To Jaime and Juana, their dedication

and hard-work are my examples.

Statement of own work

The candidate first conceived the idea for this protocol. This was further developed by the student in discussion with his tutor Professor Liam Smeeth. Further exchanges involved Professor David Leon and Professor Shah Ebrahim as members of the Advisory Committee. In Peru, on the fieldwork approach to the study, Professor Robert H. Gilman and Professor Héctor H. García provided practical insights on the conduction of the study during its fieldwork phase. All of them contributed to the final approach of the proposed study.

The candidate did the background research for this protocol, formulated the initial hypothesis, conducted the fieldwork of the study, analysed the data and wrote the present report.

As required by the University of London, a signed statement from the student confirming that the thesis is all his own work is presented in the next page.

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Statement of Own Work

All students are required to complete the following declaration when submitting their thesis. A shortened version of the School's definition of Plagiarism and Cheating is as follows (the full definition is given in the Research Degrees Handbook):

The following definition of plagiarism will be used:

Plagiarism is the act of presenting the ideas or discoveries of another as one's own. To copy sentences, phrases or even striking expressions without acknowledgement in a manner which may deceive the reader as to the source is plagiarism. Where such copying or close paraphrase has occurred the mere mention of the source in a biography will not be deemed sufficient acknowledgement; in each instance, it must be referred specifically to its source. Verbatim quotations must be directly acknowledged, either in inverted commas or by indenting. (University of Kent)

Plagiarism may include collusion with another student, or the unacknowledged use of a fellow student's work with or without their knowledge and consent. Similarly, the direct copying by students of their own original writings qualifies as plagiarism if the fact that the work has been or is to be presented elsewhere is not clearly stated.

Cheating is similar to plagiarism, but more serious. Cheating means submitting another student's work, knowledge or ideas, while pretending that they are your own, for formal assessment or evaluation.

Supervisors should be consulted if there are any doubts about what is permissible.

Declaration by Candidate

I have read and understood the School's definition of plagiarism and cheating given in the Research Degrees Handbook. I declare that this thesis is my own work, and that I have acknowledged all results and quotations from the published or unpublished work of other people.

Signed:..... [redacted] ..... Date: 15 December 2008

Full name: Juan Jaime MIRANDA MONTERO (please print clearly)

Abstract

During the 20 years of political violence in Peru starting in the late 1970's, Ayacucho, an Andean department, was one of the most severely affected areas. Mass-migration to Lima increased largely driven by escaping from violence rather than by economic reasons. This provides a unique opportunity to study the effects of migration on health since selection biases are likely to be reduced.

This study investigates differences in cardiovascular risk factors comparing three groups: i) always lived in Ayacucho (n=289); ii) migrated from Ayacucho to Lima (n=589); and, iii) always lived in Lima (n=199). A cross sectional design was used.

A clear gradient of risk was seen for the majority of factors studied: body mass index (BMI), total and LDL-cholesterol, fasting blood glucose and insulin, CRP and fibrinogen, the rural group having the lowest risk, the urban group the highest. The migrant group had intermediate risk, although generally more similar to the urban than the rural group. Blood pressure did not show a clear gradient of difference between groups. The migrant group had similar systolic blood pressure (SBP) but lower diastolic blood pressure (DBP) than the rural group. The urban group had higher SBP but similar DBP than rural group. In the case of lipid profile, no difference was observed between groups for HDL and triglycerides. Obesity, diabetes, metabolic syndrome and estimated absolute cardiovascular risk were all higher in migrant and urban groups than in the rural sample. Within the migrant group, when classified by time since migration or age at migration, differences were observed in total cholesterol, LDL, fasting glucose and insulin resistance.

The findings of this study suggest the impact of migration on cardiovascular risk is not uniform across risk factors. The study provides new insights into the increased disease risk associated with migration and urbanisation.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank all the friends and colleagues who helped me to assemble my thoughts and supported me in this work. I am particularly indebted to Professor Liam Smeeth for his continuous encouragement, friendship and advice in the last four years. Similarly, I am also very grateful for having Professor Robert H. Gilman and Professor Héctor H. García on my side while in Peru. Their motivation and friendly advice have been invaluable. My gratitude is extended to Professor Dave Leon and Professor Shah Ebrahim, for believing in this idea since its inception.

I also have a large list of friends and colleagues who I wish to thank for their support, patience and friendship in different stages of this fellowship. They know who they are, so, thank you to you all.

Special acknowledgement deserves the amazing team of people that believed in this project and make it real. All of them, who took part in all or in part of this study, deserve to and are listed as part of the research team. Last but not least, my sincere gratitude to all my family for their companion, understanding, and love and caring throughout all this process, including the good and bad moments. I would also like to thank Wellcome Trust for providing me with a Research Training Fellowship as part of their Health Consequences of Population Change programme and a PhD studentship (GR074833MA) that enabled me to start on a research pathway.

Research Team

PhD Supervisor

Professor Liam Smeeth, London School of Hygiene & Tropical Medicine

PhD Advisory Committee

Professor David Leon, London School of Hygiene & Tropical Medicine
Professor Shah Ebrahim, London School of Hygiene & Tropical Medicine
Professor Ricardo Uauy, London School of Hygiene & Tropical Medicine

Academic & Research support in Lima, Peru

Professor Robert H Gilman, Johns Hopkins Bloomberg School of Public Health
Professor Héctor H García, Universidad Peruana Cayetano Heredia
Dr David A J Moore, Imperial College London
Professor Carlos Cáceres, Universidad Peruana Cayetano Heredia
Professor Juan Lema, Universidad Peruana Cayetano Heredia

Research Team in Pampas de San Juan de Miraflores, Lima

Lilia Cabrera, Field research site co-ordinator
Juan Francisco Sanchez Chiroque, Project co-ordinator
Candice Romero Sifuentes, Research assistant
Jorge Bautista, Data management co-ordinator
Janet Vizcarra, Laboratory
Gianmarco Marquez, Administrator

Fieldwork, data collection personnel in Lima

Jessica Benavente Favian; Sonia Cristina Ccoyllo Contreras; Eva Maria Condezo Sajami; Mercedes Escobar Mendoza; Rosario Jimenez Hualpa; Rosa Lopez Bardales; Gladys Carola Mendoza Chacaltana; Maria Luisa Navarrete Garcia; Rosa Nuñez De Prado; Flor De Maria Pizarro Susanibar; Ada Lita Salas Chavez; Pilar Santa Cruz Oña; Gino Vitteri Soto

Fieldwork, data collection personnel in Ayacucho

Margot Bonilla Garcia; Javier Mendoza Huamani; Ahilcie Mendoza Leon; Vilma Deodana Pariona Palomino; Giovanna Quispe Yupanqui

Data entry personnel

Carmen Poemape Cardenas; Marissa Reyes Gallardo; Maria Torres Rivera; Marco Varela Gaona

Laboratory support

Silvia Rodríguez, Universidad Peruana Cayetano Heredia;
Manuela Verástegui, Universidad Peruana Cayetano Heredia;
William Cherre, Medlab; María Pía Baca, Medlab; Cheryl Llagas, Medlab

Terminology

In order to address global health priorities with a specific focus on developing countries, some reports have used the arbitrary classification of dividing all diseases into three major groups: communicable diseases, non communicable diseases (NCD) and injuries[1]. The group of NCD, also referred to as chronic diseases, include cardiovascular diseases (CVD), chronic respiratory diseases, diabetes and cancer.

In this work we will use the term non communicable disease, understanding that CVD is part of this classification. The term chronic diseases will not be used in this document, as it is understood as a synonym of NCD.

[1] See for example:

Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al., editors. *Disease control priorities in developing countries*. 2nd ed. New York: Oxford University Press; 2006.

World Health Organization. *Preventing chronic diseases: A vital investment. WHO Global Report*. Geneva: World Health Organization; 2005.

List of Acronyms

AHA	American Heart Association
ATP	Adult Treatment Panel
BMI	Body mass index
BP	Blood pressure
CHD	Coronary heart disease
CI	Confidence interval(s)
CRP	C-reactive protein
CVD	Cardiovascular disease
DBP	Diastolic blood pressure
GDP	Gross domestic product
HbA_{1c}	Glycosylated haemoglobin
HDL	High-density lipoprotein cholesterol
HIV	Human immunodeficiency virus
HOMA	Homeostatic model assessment
IDF	International Diabetes Federation
IFG	Impaired fasting glycaemia
IQR	Interquartile range
IR	Insulin resistance
LDL	Low-density lipoprotein cholesterol
LMIC	Low- and middle-income countrie(s)
NCD	Non communicable disease(s)
NCEP	National Cholesterol Education Programme
NHLBI	National Heart, Lung, and Blood Institute
OR	Odds ratio(s)
PAHO	Pan American Health Organization
SD	Standard deviation
SE	Socioeconomic
SES	Socioeconomic status
SMD	Standardised mean difference(s)
SBP	Systolic blood pressure
TC/HDL ratio	Total cholesterol / HDL ratio
UK	United Kingdom
USA	United States of America
WHO	World Health Organization
WHR	Waist-to-hip ratio

Conversion factors

In this document, for laboratory tests, conventional units of measurement are reported. The list below provides a selection of conventional units and their conversion factor to the international System of Units (SI)[2].

To convert from a conventional unit to a SI Unit, multiply by the conversion factor listed, e.g. Cholesterol 180 mg/dL x 0.0259 = 4.7 mmol/L. To convert from SI Units to conventional units, divide by the listed conversion factor.

Agent	Conventional Unit	Conversion Factor	SI Unit
Cholesterol	mg/dL	0.0259	mmol/L
C-reactive protein	mg/L	9.524	nmol/L
Fibrinogen	mg/dL	0.0294	µmol/L
Glucose	mg/dL	0.0555	mmol/L
HDL cholesterol	mg/dL	0.0259	mmol/L
Insulin	µIU/mL	6.945	pmol/L
LDL cholesterol	mg/dL	0.0259	mmol/L
Triglycerides	mg/dL	0.0113	mmol/L

[2] Adapted from: Iverson C, Christiansen S, Flanagin A, et al. *AMA Manual of Style: A Guide for Authors and Editors.* 10th ed. New York, NY: Oxford University Press; 2007. Available at: http://www.us.oup.com/us/pdf/9780195176339/table_2.pdf

Chapter I. Introduction

Migration is one example of social and cultural change [1], and Peru has been particularly affected in this regard. Migration can be divided as either internal migration, from rural to urban areas, or international migration. Migration studies have included the evaluation of morbidity and mortality patterns as well as risk factors associated to specific conditions amongst migrant and non-migrant groups [2].

The effects of migration on health has been addressed in a systematic review including hundreds of reports [2]. Depending on the outcome being studied, such as CVD, cancer or mental health, migration has been associated with better, similar or worse health status when migrants groups are compared to their counterparts, i.e. those who are not migrants in the same area [2].

1.1. The impact of non communicable diseases in low- and middle-income countries

Not so long ago, communicable or infectious diseases remained virtually the sole priority for global health policy; however, they do not constitute the major contributor to burden of disease in any region of the world apart from sub-Saharan Africa [3-5]. It is a common view to assume that the developing world suffers mainly infectious diseases [5-7].

In recent years the area of NCD in low- and middle-income countries (LMIC) has increasingly received more attention by being the topic of recent international reports [8, 9], two series of articles in *The Lancet* [10-21] and various commentaries, essays and editorials [10, 15, 22-25] in major biomedical journals. These publications have not only highlighted NCD's overall burden in health but also its economic impact [26, 27]. As part of the growing concern with NCD, and following a similar example from the infectious disease community, expert groups have highlighted the "grand challenges" for research and policy in NCD [28]. Some of such challenges are directly engaged with the study of the impact of poverty and urbanisation on NCD [28].

NCD kill people at economically and socially productive ages and kill them mostly in the developing world: 80% of chronic disease deaths occur in low and middle income countries [8]. Another misconception is that the epidemic of chronic diseases is still to come. As various authors have pointed out, that is no longer true: it is already here [11, 22-24]. The latest global and regional burden of disease analysis for the year 2001 suggest that almost half the disease burden in LMIC is now from NCD [29]. NCD are responsible for more than half of deaths in adults aged 15–59 in all regions except South Asia and Sub-Saharan Africa, where infectious disease conditions, including HIV/AIDS, remain responsible for one-third and two-thirds of deaths, respectively [29]. However, non-communicable diseases are also becoming a significant burden in sub-Saharan Africa [30]. The Global Burden of Disease Study,

conducted in 2001, showed that 20% of deaths in Sub-Saharan Africa were due to non-communicable diseases [31].

According to the World Health Organization (WHO), from a projected total of 58 million deaths from all causes in 2005 globally, it is estimated that NCD will account for 35 million or 60% of all deaths, which is double the number of deaths from all infectious diseases —including HIV/AIDS, tuberculosis and malaria—, maternal and perinatal conditions, and nutritional deficiencies combined [8]. It is estimated that by 2020, coronary heart disease and stroke will be the leading causes of death and disability adjusted life years, and that two thirds of the global health burden will be due to chronic diseases [32].

In 2000 in the Latin American and Caribbean region, the leading cause in mortality were CVDs [23], which accounted for 31% of all deaths, followed by communicable, maternal and perinatal conditions and nutritional deficiencies (24%), other non-communicable causes (14%), cancer (14%), injuries (13%), diabetes mellitus (3%) and mental health (1%) [33].

The greatest impact of this epidemiological transition —that is the change from a burden of disease dominated by mortality from infectious causes to degenerative or chronic causes [34]— will be in LMIC. These countries not only have to deal with their current ongoing burden of infectious diseases and ill-functioning health systems, but also with the growing burden related to NCD [25, 35-38], a situation that has been described as “a race against time” [27].

Decades of research in the developed world have shown that much of the burden associated with NCD are the result of environmental and lifestyle factors including tobacco consumption and decreased physical activity, and may therefore be preventable [32]. Despite this wealth of information available in the developed world, it is also clear that contexts are different —for example the impact of tobacco on mortality differs by geographical region [39]— and that there is an important research gap between developing and developed countries on these issues [40]. Following on this, it is also relevant to note that research findings from developed settings are not necessarily appropriate to other contexts [41, 42], thus local

knowledge is imperative. The latest edition of the widely circulated "*Disease control priorities in developing countries*" [43] suggests that "a basic task of epidemiological research is to assess geographic and secular trends in the distribution of risk factors. Of special relevance is the movement from regional to country levels and the trend within a country [3]." In the same vein, to address the growing concern of NCDs in LMIC, there is an urgent need to better document current rates —incidence and prevalence— and improve surveillance of CVD mortality and morbidity to properly assess burdens and future projections [44, 45]. NCD constitute an important public health challenge to LMIC.

1.2. Migration and health

According to United Nation's projections, by 2015 there will be 21 "megacities" of at least 10 million people —all but 4 in developing countries. In 1975 only 27% of people in the developing world lived in urban areas. In 2000 the proportion was 40%, and projections suggest that by 2030 the developing world will be 56% urban [46]. This growth in urban areas is a phenomenon that has been strongly influenced by migration. As an example of its magnitude, to date, approximately 10% of China's population are rural-to-urban migrants [47].

Although the developed world is already far more urban, at an estimated 75% in 2000, urban areas of developing countries are growing much faster [48], and their populations are larger [49]. Hence, urbanisation —heavily influenced by migration— poses a considerable challenge for public health, especially in developing countries [50-53]. Beyond these facts, it becomes a challenge to understand how city living is linked with a complexity of factors that have an effect on health [53, 54].

In Last's "*Dictionary of Epidemiology*", migrant studies have been defined as "studies taking advantage of migration to one country by those from other countries with different physical and biological environments, cultural background and/or genetic makeup, and different morbidity and mortality experience. Comparisons are made between the mortality or morbidity experience of the migrant groups with that of their current country of residence and/or their country of origin. Sometimes the experiences of a number of different groups who have migrated to the same country have been compared" [55]. This definition does not consider the context and the impact of internal within-country, mostly rural-to-urban, migration.

Barry Bogin, in his "Patterns of Human Growth", argues that migration redistributes the genetic, physiological, morphological, and sociocultural differences found in human populations [56]. In his book, a number of authors are presented as the first to publish studies on the growth patterns of urban migrants, dating from the late 19th to early 20th century. One of them, a paper by Boas titled 'Changes in the bodily form of descendants of immigrants' (published in 1912 and cited by Bogin) [57]

established that changes in growth were due to biological plasticity in the face of the new urban environment. This explanation refuted the idea that natural selection or a genetic mechanism could adequately account for the changes in growth [56].

The picture is complex, since the effect of migration on a particular outcome varies according to who is migrating, when they migrate, where they migrate from, where they migrate to, and what health outcome is measured [2]. Migration is further complicated by the fact that it is not necessarily a random process; the "selection of migrants" and the "healthy migrant effect" —or, in some circumstances, the unhealthy migrant effect— may influence health and disease risk [1, 2].

Such concern with selective migration is not new and this discussion topic dates back to 1938 [58]. Dorothy Thomas reviewed some studies evaluating the conflicting results of rural to urban migration studies which provided "apparently conflicting hypotheses as to the direction of this selection…: (1) cityward migrants are selected from the superior elements of the parent population; (2) cityward migrants are selected from the inferior elements; (3) cityward migrants are selected from the extremes, i.e., both the superior and the inferior elements; and (4) cityward migrants represent a random selection of the parent population" [58].

Thomas' review discusses Sir Austin Bradford Hill's 1925 study of migration and mortality in Essex [59] (referred to and cited by Richard Doll [60, 61] and Thomas [58]), which proceeded on the assumption that migration to the cities is selective of young adults. Bradford Hill observed lower age- and sex-specific mortality rates in cities when compared to rural areas [59-61]. "The observed differential for the age-selected groups favoured urban areas, in general, and especially females in urban areas, thus leading to the inference that at least part of this differential could be attributed to selective migration, and that therefore migrants to the cities represented, on the average, better physical risks than the residual population in rural areas" [58].

The observations made in the earlier in the 20th century [58, 59] are still relevant to most of today's LMIC, since no single explanation was made owing to the complexity of the problem. Migrants to urban slums in today's less developed countries do not necessarily experience the benefits of the urban environment. In

Asia, Africa and Latin America the slums are often on the outskirts of the cities [56]. Not surprisingly, the growth of migrant children living in these slums is not significantly different from that of children living in the impoverished rural areas [56]. Migration due to economic reasons poses additional difficulties in interpreting studies, since those with better health or socio-economic status could be the ones more likely to "afford" to migrate. It could, therefore, be argued that migrant groups are self-selected groups.

Other related factors that complicate this panorama lie in the interpretation —as well as its applicability and generalisability— of research results derived from the available literature. For example, several studies show a lack of reference groups with similar characteristics that facilitate the comparison of the effect to be studied such as, people from the same place of origin who did not migrate. In other instances, difficulties are present with the selection of comparison groups that are not necessarily "similar" to each other, such as those studies that involved migrant groups from different generations.

Razum [62] outlined what would be the requirements of an "ideal" migrant cohort which includes a unique definition of "migrant" that considers duration of stay. Additionally, and ideally, participants would have to be enrolled before they migrate, studies should include the population of origin of immigrants and studies should be based on individual data collected over time to understand the determinants of the relation between migration and health [62].

Due to the complexity outlined in this section, it is, therefore, important to be aware of these issues due to its impact in the design of new research studies, as well as in the interpretation of their findings.

1.3. The effects of migration on cardiovascular disease

Urbanisation is a challenge for today's world and its future [48], and this process has had a profound effect in Latin America [23, 49, 63]. The Latin American region has experienced the greatest urbanisation: more than 60% of its population now live in urban areas [48] —mostly in poor conditions [54]. Obesity and rapid changes in lifestyles have already been expressed as a concern in the region of the Americas [64-66].

In relation to migration and CVD, and its risk factors, the literature is quite prominent and has been approached in a systematic way [2]. This section does not intend to be a comprehensive review of the available literature and briefly mention some of the relevant material published on migration and CVD. For clarification purposes and to avoid mixing evidence from different contexts, this section is divided into two: the effects on CVD of international and internal migration.

In terms of migration and CVD, the NiHonSan study constitutes one of the classical studies of international migrant populations. This study looked at men of Japanese ancestry living in Japan, Hawaii and California describing rates of coronary heart disease and stroke for these groups, and is described below [67]. More recently, Kelleher et al. were able to link historical census data from 1850 onwards, involving European subjects that migrated to the United States of America (USA) and native-born individuals, providing further insights into the coronary heart disease epidemic of the USA in the mid-20th century [68].

1.3.1. International migration

International migration is a complex process, usually associated with deep changes in culture. While migrants may move and keep some protective factors for disease prevention, i.e. low levels of smoking and healthy diets, this is not always necessarily maintained in all migrant generations. As an example, a recent publication from the

Millennium Cohort Study in the UK has shown that after immigration from various countries and settlement in the UK, maternal health behaviours, i.e. smoking during pregnancy and initiation of breast feeding, worsen with length of residency in the UK [69].

When approaching the subject of migration and CVD in the UK it is typical to think about south Asian populations living in this country. Variations in risk factors [70] and mortality [71], particularly due to ischaemic heart disease and stroke, differs markedly by country of origin, being higher amongst non-UK born population. Similarly, in the USA, studies of migrants refer somehow to the larger ethnic groups that live in that country, such as "Latinos" or "Hispanic" groups for example [72, 73].

One of the classical studies of migrant populations published more than 30 years ago, the NiHonSan study, looked at men of Japanese ancestry living in Japan, Hawaii and California. It described that rates of coronary heart disease were low in Japan, high in Hawaii, higher in California, and higher still among white Americans, and an opposite gradient was observed for the case of stroke [67].

On the other hand, lower mortality rates from CVD have been reported in migrants from Latin America, China and South Asia that moved to Canada [74]. An "epidemiological paradox" has also been described amongst Latinos in the USA, who showed lower socio-economic status and also lower all cause mortality [72]. Potential explanations for this phenomenon were attributed to selective migration and the return of those ill to their places of origin, phenomenon also known as "healthy migrant" effect [72, 75]. Additional explanations could be related to bias in the ascertainment of outcome, particularly in the case of illegal migrants that do not want to declare their health status due to fears of being deported.

Altogether, based on this evidence, it has been suggested that the studies of migrant populations could contribute to the understanding of the aetiology of some diseases, particularly the role played by the environment [1, 76]. In addition to the role of environmental factors, it has been proposed that the study of migrants could

contribute to elucidate the role of the genetic background in the "development" of some specific conditions, such as Type 1 diabetes mellitus [77].

1.3.2. Internal migration

In Kenya, migration from rural to urban areas has been shown to be associated with an increase in blood pressure amongst those who migrated [78, 79]. Similar findings, that migrants had higher systolic and diastolic blood pressure than did the people in the rural areas of origin, has also been described in Iran, with the addition that blood pressure levels of migrants and non-migrants in the city were not much different from each other [80]. In South Africa, on another but closely related disease, the prevalence of diabetes has been described as doubling within two decades in populations that adopted a high-calorie, low-exercise lifestyle as a result of migration to urban areas [81].

In Tanzania, however, after 6 months following migration, migrants appear to have lower levels of blood pressure and triglycerides indicating that the direction of change of risk factors is a much more complex phenomena [82]. These findings supported from research from Cameroon which included data on lifetime exposure to an urban environment [83]. In this West African setting, by Sobngwi et al. suggest that both lifetime exposure to urban environment and recent migration are potential risk factors for obesity and diabetes mellitus [83].

Migrants moving from South Western rural China to the urban area of Xichang City showed an increase in serum total cholesterol lipid levels which could lead to elevated coronary heart disease risk [84]. Similarly, the same group has described an increase in blood pressure levels following rural to urban migration [85, 86]. In Guatemala, migration to a city has been reported to increase sedentary and non-healthy eating habits [87], both recognised risk factors for CVD.

1.4. The Peruvian context

In this section some reference is given to Lima (population 8.5 million, 30% of Peru's population, 98% urban [88]) and Ayacucho (population 600,000, 2% of Peru's population, 58% urban [88]), two Peruvian departments located in the coast and in the Andes, respectively. For geographical orientation, please refer to Figure I-1 and Figure I-2 on pages 14 and 15.

Peru had a population of 27,219,264 inhabitants in 2005 [89] and it is estimated to reach 30 million by 2010 [90]. Peru's demographic growth rate for the 2005-2010 period is projected to be 1.4%, which is very similar to the 1.3% projection for the Latin American region [91]. Gross domestic product (GDP) per capita was USD $2,300 in 2004 [92, 93]. As shown in Figure I-3, services and manufacturing are the largest contributors for Peru's GDP in the year 2006 [91], a trend observed in the last 20 years [91, 93].

Total expenditure on health as percentage of gross domestic product is the lowest in the region (4.1% in 2004), even below Haiti, Bolivia and Ecuador. Out-of-pocket expenditure as percentage of private expenditure on health is 79.2% in Peru, a fact that hinders access to health care to the majority of population.

Income is highly concentrated in Peru: while the 20% of the population with the highest income received 47.5% of national income, the 20% of the population with the lowest income received only 6% [92, 93]. The unequal distribution of income, expressed as a ratio between the highest and lowest quintiles, rose from 4.9 to 7.9 between 1997 and 2000 [92]. Access to improved drinking water sources and improved sanitation is 89% and 74% in urban areas but only 65% and 32% in rural areas [94]. Overall, 73% of the country live in urban areas [88, 94]. While historically Lima has been mostly urban, Ayacucho has seen a progressive increase in its proportion of urban inhabitants: 25%, 33%, 37%, 48% and 58% in 1961, 1972, 1981, 1993 and 2007, respectively [88].

Total poverty increased from 48.4% in 2000 to 52.0% in 2004 and was greater in rural areas (73.6%). The percentage of people living in extreme poverty rose from 15.0% to 20.7% during that same period and was 42.5% in rural areas [95]. In 2004, the illiteracy rate was 11.6%, with a wide gap between men and women (5.8% and 17.2%, respectively) as well as in poor areas with a high percentage of *campesino* (peasant) and Quechua-speaking populations, including Ayacucho, where these percentages exceed 6% for men and 25% for women [96].

Unemployment and underemployment rates for 2004 in metropolitan Lima were 10.5% and 42.8%, respectively. A major labour phenomenon in Peru is the large proportion of the economically active population (55.0%) that works in the informal sector of the economy and that therefore has no access to social security, does not receive a steady income, and has no provisions for retirement. Both unemployment and underemployment are greater among women. In 2004, unemployment affected 9.4% of men and 12.0% of women, and underemployment 35.9% of men and 52.5% of women [97]. National statistics for years 2005 to 2007 disaggregated by rural versus urban areas show that nearly 30% of the population in rural areas works in agriculture. On the contrary, in urban areas, small-scale family-run businesses, e.g. street vendors, grocery shops, street food, etc. are the most common type of job (13%), followed by elementary occupations, e.g. cleaners, messengers, watch persons, etc. (10%) [98].

The Peruvian population structure for years 1981, 1993 and 2005 are provided in Figure I-4. This shows a pattern of decrease in proportion of 0-15yrs age-group and increase in >65yrs age-group [92]. The population of Lima, Peru's capital, has increased from 3.5 million in 1972 to 8 million in 2002 [99]. According to Peru's 2000 Demographic and Health Survey [100], urban migrants and their children constitute 11% of the population of Peru [101].

In year 2005, life expectancy at birth was 70 years for males and 74 years for females [94]. This is similar to neighbour Ecuador and Bolivia, but lower than other countries in the region, such as Argentina, Brazil, Chile and Colombia. In the Latin American region, Peru and Bolivia are the countries with greatest stunting, >30% in children

under five years old. Peru, together with Chile, also ranks first in overweight in the same population with >11% children under five being overweight [94].

Figure I-1. Peru in Regional Map of South America

Figure I-2. Map of Peru and location of Study Sites in Lima and Ayacucho

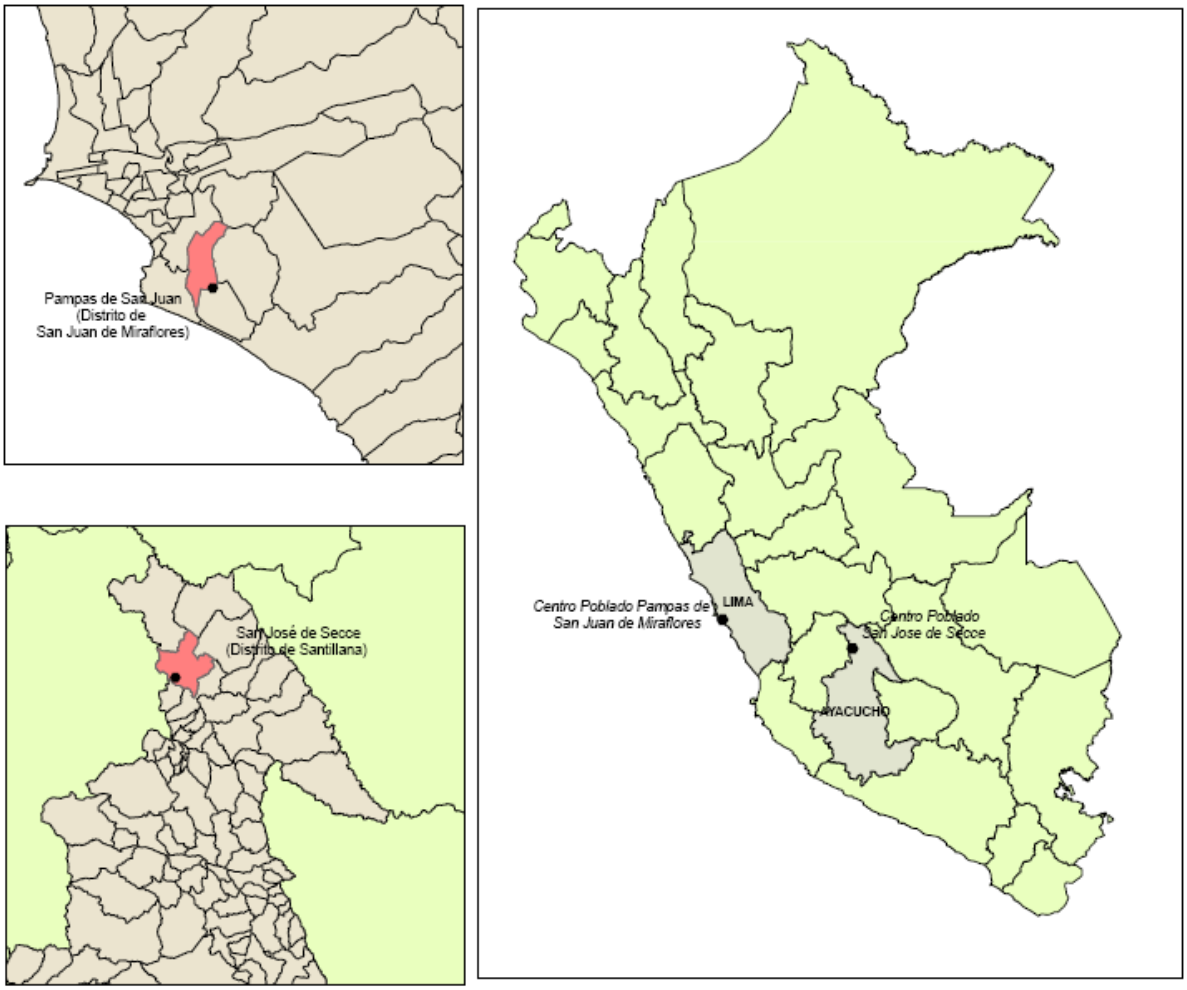
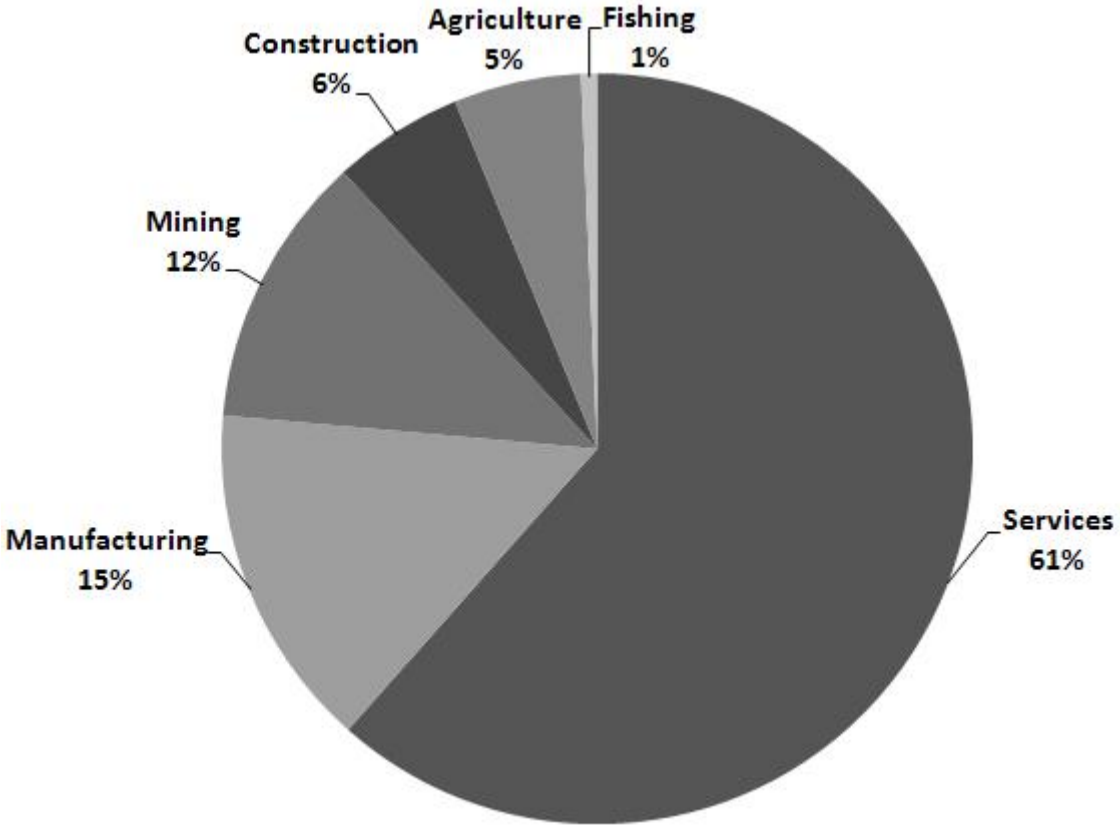


Figure I-3. Peru's gross domestic product by economic sector, 2006



Source: Instituto Nacional de Estadistica e Informatica, Dirección Nacional de Cuentas Nacionales & Banco Central de Reserva del Peru, Gerencia de Estudios Económicos. GDP tables for year 2006 elaborated by Instituto Cuanto (2007).

Figure I-4. Population structure, by age and sex, and by urban and rural areas, Peru, 1993 and 2005

Figure I-4a.

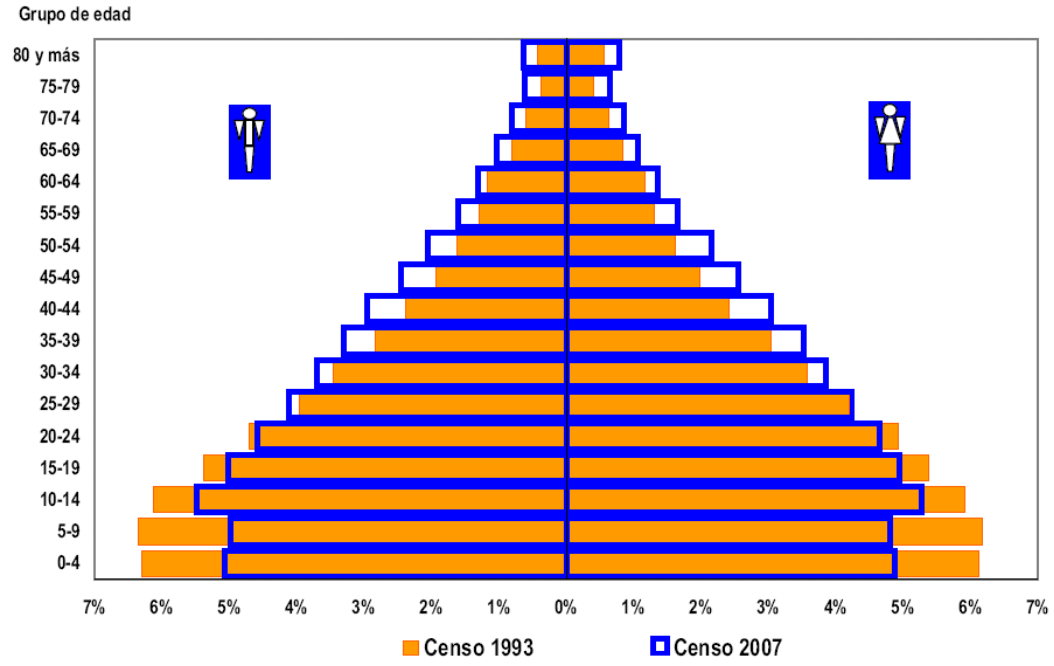
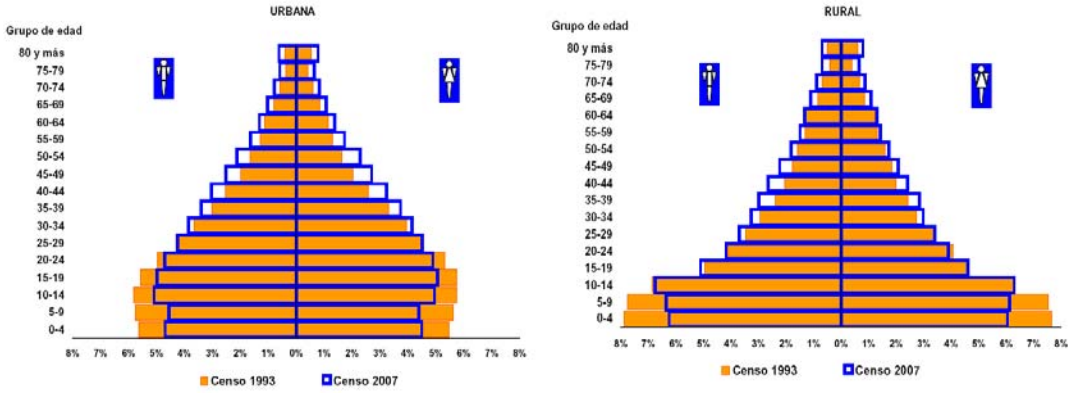


Figure I-4b.



Notes: Both images compare National Population Censuses from 1993 (orange background) and 2007 (blue gridlines). Figure I-4a shows Peru's total disaggregated by age and sex, males on the left-side and females on the right-side of the image. Figure I-4b shows distribution of urban (left-side) and rural population (right-side).

Source: Instituto Nacional de Estadística e Informática, Peru's national censuses 1993 and 2005 [88, page 38].

1.4.1. Cardiovascular diseases in Peru

As a proportion of the total years of life lost due to premature mortality in the population, Peru's years of life lost to non-communicable diseases is 42% [94]. The same indicator for Bolivia, Colombia, Ecuador, Brazil, Chile and Argentina are 34%, 35%, 42%, 50%, 64% and 66%, respectively [94].

At the national level, ischemic heart disease and cerebrovascular disease are ranked as the second and third leading cause of mortality, with a mortality rate of 25.7 and 24.3 per 100,000 population [92]. However, mortality profiles in Peru are very heterogeneous: absolute mortality rate in southern highlands is 2.5 times the mortality in Lima (11.2 vs. 4.5 per 100,000 population). Although a high proportion of mortality is attributed to infectious diseases in the southern highlands compared to Lima, the mortality rate from cerebrovascular disease in the Andes is twice than in Lima (41.3 vs. 20.2 per 100,000 population) [92]. Peru's mortality registration system is very poor, with over 45% of underreporting [102, 103], thus this mortality data should be taken as indicative only.

Age-adjusted mortality rates from diabetes has increased from 10.8 per 100,000 population in 1980-1982 to 18.4 per 100,000 population in 1999-2000 [104]. The prevalence of diabetes is estimated to increase from 5.1% in 2003 to 6.7% in 2025, with a much higher burden in urban areas [105]. Regular sports activities are described to be less frequent in metropolitan Lima than other parts of the country [106].

Overweight and obesity are becoming serious concerns for the Peruvian population [107]. A recent published national survey on NCD risk factors in Peru, the largest ever for the country, shows a prevalence of overweight and obesity of 35.3% and 16.5%, respectively. The prevalence of obesity amongst females is 20%. Prevalence figures at the national level for hypertension, elevated total cholesterol, low HDL cholesterol and diabetes are 9.5%, 19.6%, 1.1% and 2.8%, respectively. However, clear differences are observed between regions [108, 109], and even between Lima [110] and Arequipa [111, 112], Peru's two largest cities. These differences are shown in a detailed description of comparative studies and reported prevalence for CVD

risk factors in Peru in (Table I-1 and Table I-2, pages 20 and 21, respectively). Another population survey in six urban areas from Peru has shown that high total cholesterol, low HDL cholesterol, diabetes, hypertension and overweight and obesity have been found to be strongly associated with lower socio economic status [113, 114].

Also, a number of other smaller studies are contributing towards the understanding of NCD in Peru, but despite the fact that their small sample size limits the ability to draw major conclusions, they provide some interesting findings that warrant further attention. For example, low levels of awareness and control of hypertension were found in a relatively young population in a deprived urban area of Lima [115].

In another setting, women in the capital have been found to have increased body fat mass compared to women in Cuzco, a city in the Andes, despite having similar body mass indexes (BMI), an association reported to be linked to migration [116]. In the same vein, the metabolic syndrome is started to be recognised as a prominent problem for some groups, particularly urban *mestizos* —a category that includes that includes large proportions of Peruvians which is a mixed race between Andean and other non-Andean groups, usually white people [117-119].

Beyond the studies already mentioned for Peru, the literature on CVD and its risk factors is very limited for the country [73, 120-122]. These studies are not discussed in detail because of their small sample size ($n < 200$). However, estimates of disease frequency and risk factor distributions included in these studies are largely in line with the published studies included in the tables here.

Table I-1. Description of large population-based studies on cardiovascular disease and risk factors in Peru

Name of the Study (Year[‡])	Location	Sample size (n)	%Male	Mean age ± SD (y)	Strengths, comments
ENINBSC Survey, Peruvian National Institute of Health (2004-2005) [108]	Urban and rural	4206	49.9	N/A	Nationally representative. Five regions/strata (20% each): Lima, rest of Coast, urban and rural Andes, Jungle. Five age groups, 20 to >60yo.
Non Communicable Risk Factors Survey, Office of Epidemiology, Peruvian Ministry of Health (1998-2000) [114]	Urban	2337	50.1	42.1 ± 8 M 38.0 ± 7.3 F	Six study sites in total: Lima (2 sites), Ica, Arequipa, Huánuco and Ucayali
PREVENCION Study (2004-2006) [111, 112, 123]	Urban	1878	46.2	49.6 ± 17.4 M 48.5 ± 17.0 F	Arequipa, Peru's second largest city
CARMELA Study, InterAmerican Heart Foundation (2008) [110]	Urban	1652	46.6	43.6 ± 1.6	Lima, Peru's capital, first largest city

[‡] Where possible, year of conduction of the study is mentioned. If such data is not available, date of publication is presented.

Table I-2. Estimates of prevalence of various cardiovascular disease risk factors in Peru from large population-based studies

Risk Factors & Diseases[§]	Overweight	Obesity	Hypertension	High Total Cholesterol	Low HDL	Diabetes
ENINBSC Survey [108][**]						
By region						
Lima	40.4	18.8	11.6	20.2	0.7	4.6
Rest of Coast	31.7	20.2	11.2	23.7	1.2	2.5
Urban Andes	32.9	10.8	5.2	16.9	1.4	0.7
Rural Andes	22.1	9.2	7.2	12.1	2.7	0.3
Jungle	34.9	15.2	9.1	11.2	1.8	2.5
By sex						
Male	31.2	12.6	10.9	17.5	1.4	3.2
Female	39.1	20.3	8.3	21.6	0.9	2.4
National	35.3	16.5	9.5	19.6	1.1	2.8

[§] All prevalences are expressed in percentages. Overweight = BMI ≥25Kg/m^2 and ≤29.9Kg/m^2; Obesity = BMI ≥30Kg/m^2; Hypertension = systolic blood pressure ≥140 or diastolic ≥90 mm Hg or use of antihypertensive drugs; High total cholesterol = total serum cholesterol ≥240 mg/dL; Low HDL cholesterol, serum HDL <40 mg/dL; Diabetes = fasting glucose ≥126 mg/dL or self-reported diabetes

[**] In this study high total cholesterol was defined as total serum cholesterol ≥200mg/dL; high LDL cholesterol as serum LDL >130 mg/dL; and, low HDL cholesterol as serum HDL <35 mg/dL. Diabetes was defined as fasting glucose >100mg/dL and self reported diabetes; fasting glucose ≥200mg/dL and negative self reported diabetes or medical treatment for diabetes

Table I-2. (continued)

Risk Factors & Diseases	Overweight	Obesity	Hypertension	High Total Cholesterol	Low HDL	Diabetes
Non Communicable Risk Factors Survey [114]						
Male	44	16	44	27	38	19
Female	40	24	47	21	40	17
PREVENCION Study [111, 112, 123]						
Male		13.8	28.2		32.3	6.3
Female		18.5	25.5		60	5.9
CARMELA Study [110]		22.3	12.6	11.6		4.4
Male		21.1	14.4	10.1		4.3
Female		23.4	10.7	13		4.6

1.4.2. Internal migration in Peru

Dufour and Piperata [63], on the evaluation of rural-to-urban migration in Latin America, have argued that "the classic case of people moving from a distinctly rural setting to a distinctly urban one and staying there for the remainder of their lives is not the norm". As stated previously, migration is not a simple phenomenon. Small-scale migration from rural to urban areas has been present in Peru for many years, largely driven by economic forces, ranging from seasonal migration to neighbouring areas to long term or definite migration [124].

However, the "normal" or "regular" migration process in Peru changed dramatically during the 20 years of political violence that occurred in Peru starting in the late 1970's. These periods of violence and internal migration, notorious for the occurrence of systematic murders [125], has been well documented in the social science literature [126-128]. Although the exact figures are not known, Peru's violence in the 1980–1990's period yielded a shameful and painful balance of thousands of deaths —in the order of 70,000 deaths— with 79% of them occurring in rural areas [125]. Also importantly were the large amounts of displaced people due to the violence of those years, calculated in 120,000 displaced families [127]. This process of displacement showed a pattern of migration that did not resemble the "traditional" short-term short-distance rural-to-urban migration as usually described [63]. People moved long distances, usually to stay in larger cities without establishing initially in nearer towns [128].

Ayacucho, an Andean department, was one of the most severely affected areas during this period of violence [129]. More than 50% of all deaths attributed to the violence between Shining Path guerrilla Peruvian militia and reported to Peru's Truth and Reconciliation Commission [125] occurred in Ayacucho. Such was the extent of these brutalities that some of Ayacucho communities were totally destroyed during the fight between Shinning Path and the Peruvian armed forces.

According to Peru's 1993 national census, Ayacucho was the only region with a -3.3% negative balance in total population when compared with 1981 [128]. National statistics show a doubling in the number of people leaving Ayacucho from 38,539

during the period 1972-1981 to 94,708 during 1981-1993. The latter period coincides with the political turmoil and violence that affected the area. It has already been stated that the population of Lima increased from 3.5 million in 1972 to 8.5 million in 2002. For the period 1988-1993, 50.7% of the total emigrants from Ayacucho moved to Lima, making Ayacucho the leading source of migrants to Lima [99, 130] (See Figure I-5. Rural to urban migrants entering Lima in 1973).

Thus the mass-migration seen in Peru, and particularly in Ayacucho, from the 1980s onwards was largely driven by the need to escape from politically motivated violence rather than only a migration for economic reasons. In some rural communities, a very high proportion of people migrated to the cities: the migrants were not simply a small self-selected atypical group.

In the Andean region most people work as subsistence farmers and their lifestyle is shaped by the basic conditions: hard physical work, much walking, diets of unrefined foods and few televisions. Once in Lima, unemployment is common; what jobs exist consist largely of menial work in service industries and factories. Public transport is commonly used to travel long distances, so people walk less, diet moves towards refined energy-dense foods, and leisure time is dominated by television [66].

Figure I-5. Rural to urban migrants entering Lima in 1973

Image taken from an exhibition about Ayacucho in the recent decades (Huamanga, Ayacucho, July 2007). Image source: La Prensa newspaper archives, photograph taken by Mr. Talledo, 1973.

Chapter II. Methodology

Following on the previous introductory section, it becomes relevant to evaluate the impact of rural to urban migration in terms of NCD in the Peruvian context. More specifically, this study addresses the impact of migration on cardiovascular risk factors, would enable us to have a better approximation to the problem in the Peruvian setting. This chapter describes the methodology used in the present study.

2.1. Hypothesis, objectives and research question

2.1.1. Hypothesis

The risk of CVD increases following migration from rural to urban areas in Peru.

2.1.2. General objective

To assess cardiovascular risk factors among people who migrated from rural to urban areas during the 1980-1990 period in Peru.

2.1.3. Specific objective

To describe differences in cardiovascular risk profiles amongst migrants from Ayacucho, a rural area, to Lima, an urban area, and those who did not migrate.

2.1.4. Research questions

The present study was conducted to address the following research questions.

2.1.4.1. Overall research question

i) Is there a difference in specific CVD risk factors in the rural-to-urban migrant group compared to those who did not migrate?

2.1.4.2. Specific research questions

Does the pattern of CVD risk factors in the migrant population vary by:

ii) length of residence in urban environment?

iii) lifetime exposure to urban environment?

iv) age at first migration?

Additionally,

v) what are the specific CVD risk burdens on each of the study populations?

2.2. Study design

2.2.1. Type of study

Cross-sectional survey of three population-based groups:

1) Rural, people who have always have lived in a rural environment;

2) Rural-to-urban migrants, people who migrated from rural to urban areas; and,

3) Urban, people who have always lived in an urban environment.

2.2.2. Study population and setting

The Peruvian departments of Lima (Peru's capital) and Ayacucho (for a map please refer to Figure I-2, page 15) were selected for the conduction of this study due to the migration patterns and rural-urban contrasts existing between these two geographical areas. Ayacucho is a remote mountainous region characterised by high levels of migration to Lima, as presented before in the introductory chapter. Most of the immigrants who have come from Ayacucho have settled in defined areas in Lima's shantytowns. In addition, Ayacucho lifestyle and environment could hardly be more different to the one in Lima. People living in rural communities in Ayacucho and Lima's shantytowns could be deemed representative of a large majority of Peruvian population living in similar conditions.

The village of San Jose de Secce, located in the Santillana district, Huanta province in Ayacucho was selected as the rural study site (See Figure II-1 and Figure II-2, pages 32-33). The area called "Las Pampas de San Juan de Miraflores" in the district of San Juan de Miraflores in Lima, where migrants tend to settle, was selected as the urban area for the study. Both urban and rural-urban migrant participants were selected from the Pampas de San Juan de Miraflores area, a periurban shantytown in the south of Lima (see Figure II-3 and Figure II-4, pages 34-35).

Peruvian population consists largely of *Mestizos* (‘‘mixed race’’) and shows a high degree of admixture, being predominantly Andean Amerindian (i.e., autochthonous Quechua and Aymara populations), with small contributions from Spanish Whites and minimal contributions from West African populations [131-137]. Thus is difficult to differentiate the groups by ethnicity. In fact, Peruvian statistics usually do not collect of record ethnicity, since more than 80% of the population qualifies as *mestizos*. Indigenous groups are usually identified as such if they have any other native Peruvian language as a mother tongue but Spanish. In this study, rural and rural-to-urban migrant groups have strong links to traditional Andean indigenous populations and could be regarded as indigenous population. In the same vein, the urban group is most likely to have some degree of connection to indigenous populations by being *Mestizos* and/or first or second-generation descendants from relatives from Andean ancestry.

Figure II-1. San Jose de Secce, Santillana: Landscape

Rural study site: village of San Jose de Secce, landscape. Average journey time to Huanta, the nearest town, is about 3 to 4 hours using small minibuses (J. Jaime Miranda, 2007).

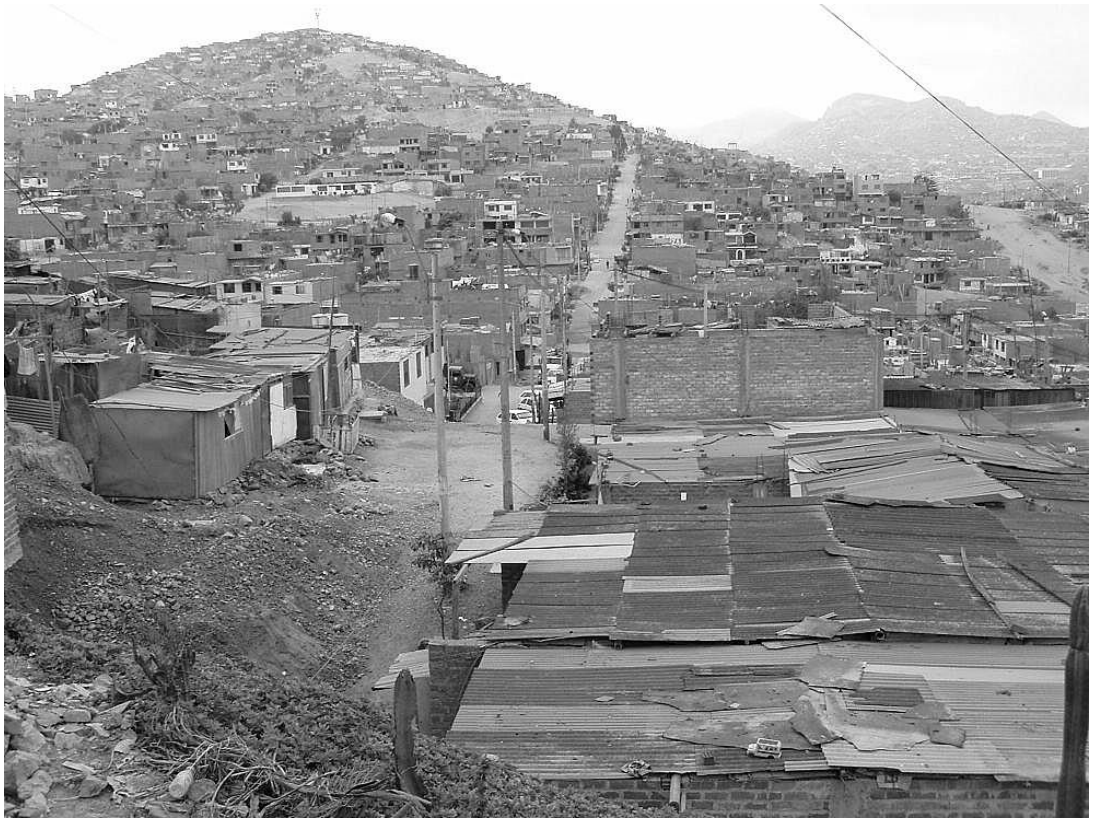
Figure II-2. San Jose de Secce. Santillana: Farmers

Rural study site: inhabitants of San Jose de Secce working in the land. Most of the population work as small-scale subsistence farmers. (Salud Sin Límites, 2002)

Figure II-3. San Juan de Miraflores, Lima: Cityscape 1

Urban study site: view towards Las Pampas de San Juan de Miraflores in Lima. Average journey time from the area towards downtown Lima, using public transport similar to the bus shown in the picture, is 2.5 hours. (J. Jaime Miranda, 2006)

Figure II-4. San Juan de Miraflores, Lima: Cityscape 2



Urban study site: view from within Las Pampas de San Juan de Miraflores. Most houses are use corrugated iron and cardboard as roof materials. Electricity and tap water is in place. (J. Jaime Miranda, 2006)

2.2.3. Sampling strategy

A single-stage random sampling method was used in all groups. Potential participants were randomly identified from censuses in both the urban and rural settings.

In the case of San Jose de Secce in Ayacucho, a census was conducted in mid 2007 to identify all adult population permanently living in the area that would serve as sampling frame for the rural group.

The sampling frame for the urban group was derived from the local census carried out by Asociación Benéfica PRISMA, conducted in year 2000, in the area called "Las Pampas de San Juan de Miraflores" in the district of San Juan de Miraflores in Lima. All those who reported to have been born in Lima in the 2000 census and currently living permanently in the recorded address were considered eligible for the study.

In the case of the rural-to-urban migrant group, the same 2000 census was updated in 2006 to identify all those who referred to have been born in the department of Ayacucho and were currently living in "Las Pampas de San Juan de Miraflores" in Lima.

For all study groups, individuals from both sexes aged 30 years-old and over, permanently living in their residence were considered eligible to take part on this study. Pregnant women were excluded from the study.

The study target was to recruit a total of 1000 people, i.e. 200 people in each of the rural and urban groups and 600 people who migrated from rural to urban areas. Participant's selection was matched by age-groups and sex to ensure sufficient number of people in each stratum. Sapsford [138] suggests that stratification (referring to matching) is recommended in survey research since such procedure "improves the estimates of population means, provided the strata are fairly distinct from each other and that each stratum is reasonably homogeneous, … and, its major use in normal practice is to ensure that important groups are adequately covered" [138, pp. 69].

Sample size targets and achieved sample size are presented in Table IV-2, Table IV-3 and Table IV-4 and discussed in Chapter 3.

2.2.4. Exposure

The primary exposure was migration from a rural to an urban environment, defined by study group, i.e. rural, rural-to-urban migrant and urban groups.

In order to address the study's specific research questions, the migrant group was subsequently divided by length of residency in an urban area, lifetime exposure to an urban area, and age at first migration. The exact definitions used are presented in Table II-6 and Table II-7, in pages 49 and 50, respectively.

2.2.5. Outcomes

The primary outcomes of the study were blood pressure, prevalence of hypertension, BMI, WHR and prevalence of obesity, fasting glucose and diabetes, total cholesterol and lipoprotein profile.

Secondary outcomes were behavioural risk factors (alcohol consumption, smoking status), inflammation markers (CRP and fibrinogen), insulin resistance and metabolic syndrome and aggregate measures of CVD risk.

All the definitions for each outcome of the study were presented in Chapter II, in separate tables (Table II-9 to Table II-13). The results chapters are organised in relation to the analyses carried out for specific outcome groups. The last results' chapter, Prediction of CVD risk, has been set out to address the last specific research question of the proposed study, that is what are the specific CVD risk burdens on each of the study populations?

2.2.6. Sample size and power

The study target was to recruit a total of 1000 people, i.e. 200 people in each of the rural and urban groups and 600 people who migrated from rural to urban areas.

Power calculations were made with the `sampsi` command in Stata using conservative estimates of the prevalence of major risk factors in the areas of Huaraz (urban, Andes) and Ingeniería (urban, Lima) from preliminary work in Peru [109]. At the time of study planning, in 2005, these two urban settings were the best approximation available for the calculation of required sample size and power. Due to the limited research available in the country, there is a lack of studies on rural-to-urban migration and CVD in Peru.

Comparing the Lima with the Andes group, with 200 people in each group, the study had 80% power or greater at the 5% significance level to detect a difference in the prevalence of: hypertension (>140/90mmHg) 33% versus 19.5%, power 0.84; hypercholesterolaemia (>6.3mmol/L) 22.7% versus 10.6%., power 0.88; and diabetes 7.6% versus 1.3%, power 0.81.

It is likely that the study will have more power than shown by the calculations since it will be able to look at rural groups from the Andes, whereas the calculations made previously derive from information based on urban groups from one city in the Andes.

2.3. Data requirements

2.3.1. Continuous variables

Age, systolic blood pressure, diastolic blood pressure, height, leg length, weight, body mass index, abdominal and hip circumferences, waist-to-hip ratio, skinfold measurements, fasting glucose, fasting insulin, total cholesterol, LDL cholesterol, HDL cholesterol, triglycerides, haemoglobin, C-reactive protein, fibrinogen.

2.3.2. Categorical variables

Sex, childhood and current socioeconomic status, migration (age at first migration, age of arrival to lima, years of settlement, reasons for migration, place of origin, previous migration records), smoking and alcohol intake, history of diabetes and medication, history of hypertension and medication, employment type (current) and self-reported health.

2.3.3. Definitions used

The following tables show the information collected through questionnaires, clinical measurements and laboratory assays and the definitions used to create the variables used in the analysis. In order to guide the reader, the following subheadings have been used: socioeconomic variables, definition of exposure variables, migration variables, clinical assessment variables, and definition of outcome variables.

For a detailed presentation of the processing of blood samples, please refer to Appendix G, Field and laboratory processes for blood testing and storage of samples. Data collection methods are presented in detail in Chapter IV, Execution of main study.

Table II-1. Definitions of general demographic variables

Variable	Type	Categories	Definition	Comments
Age	Continuous		Age at the time of interview	Based on date of birth or, if not known, self-reported age
Age group	Categorical	30-34 years old 35-39 years old 40-44 years old 45-49 years old 50-54 years old 55-59 years old ≥60 years old	Age grouped into 5-year categories, 7 in total	
Sex	Categorical	Male Female	Gender category	
Language	Categorical	Spanish Quechua Both	Language of interview	

Table II-2. Definitions of current socioeconomic position variables

Variable	Type	Categories	Definition	Comments
Educational level, individual	Categorical	None Primary incomplete Primary complete Secondary incomplete Secondary complete or more	Educational level attained at the time of interview	Proxy of current socioeconomic position [139-142]
Household income	Categorical	≤$50 dollars $51-150 dollars $151-250 dollars $251-350 dollars $351-450 dollars ≥$450 dollars	Income generated by all the members living in the household	An additional proxy measurement for socioeconomic position [139-142]. Family income was measured through a questionnaire using the same ranges as the ones used in Peru's national census [89]. The census and this survey consider range values in US dollars.

Table II-2. (continued)

Variable	Type	Categories	Definition	Comments
Number of people per room	Categorical	<2 people per room 2-3 people per room 3-4 people per room 4 or more people per room	Number of people living in the household divided by total number of rooms, excluding kitchen and bathroom	This indicates household overcrowding, another proxy indicator of socioeconomic position. Recommended thresholds for overcrowding are two people per room [140]. However, for the Latin American context, such figure has been set out as three or more [143].

Table II-2. (continued)

Variable	Type	Categories	Definition	Comments
Assets index	Categorical	Lowest tertile Middle Highest tertile	Possessions weighted asset index in tertiles	Index constructed based on current ownership of household assets, weighted according to the relative proportion of ownership following methods described elsewhere [144, 145]. For example, 15% of the study's sample owned a bicycle, so a score of 85 (i.e. 100 minus 15) was applied to those individuals who own a bicycle and zero to those who did not. The twelve items included were current ownership of gas cooker, radio, colour television, refrigerator, computer, telephone, mobile phone, cable, internet, bicycle, motorcycle, and car. (Cronbach's alpha = 0.7486, indicating high reliability). The final index, a continuous variable generated from the sum of individual item's scores (median 171, IQR 95 – 257), was then divided in tertiles.

Table II-2. (continued)

Variable	Type	Categories	Definition	Comments
Individual's current socioeconomic deprivation	Categorical	No Yes	Deprivation of at least two socioeconomic indicators	Deprivation indexes have been widely used for the appropriate measurement of poverty [143-153]. Please refer to Chapter V, Section 5.4 for full details on elaboration of this index.

Table II-3. Definitions of childhood socioeconomic position variables

Variable	Type	Categories	Definition	Comments
Educational level, parents	Categorical	None Primary incomplete Primary complete Secondary incomplete Secondary complete or more	Educational level attained by the participant's father, mother or tutor when the subject was 12 years old	Measured separately for paternal and maternal educational level as a marker of childhood socioeconomic status [139-142]
Highest parental education level	Categorical	None Some primary Primary completed or more	Aggregation of both, paternal and maternal educational level attainment	Priority was given to any highest education level in any of both parents. For example, if a household had a mother without education and a father with some primary, this would have been coded as some primary.

Table II-4. Definitions of migration variables

Most of these questions were only applicable to the migrant group, and this information was used to create more detailed description of migration patterns, to be used later as alternative exposure groups (see Chapter V).

Variable	Type	Categories	Definition	Comments
Place of birth	Categorical	Rural Urban	Self ascertainment of place of birth	
Age at first migration	Continuous		Self ascertainment of age (years) when first migrated for a period of 6 months or more.	
Age at arrival in Lima	Continuous		Self ascertainment of age (years) when arrived and established in any part of Lima, Peru's capital	This definition considers the arrival to any location in Lima and not only to the urban study site

Table II-4. (continued)

Variable	Type	Categories	Definition	Comments
Education level at first migration	Categorical	None Primary incomplete Primary complete Secondary incomplete Secondary complete or more	Self ascertainment of education level attained at the moment of first migration	
Years lived in urban area	Continuous		Self-ascertainment of number of years living in an urban area	
Lifetime exposure to urban area	Continuous		Number of years lived in an urban area divided over age, expressed as percentage.	

Table II-5. Definition of general exposure variables

Variable	Type	Categories	Definition	Comments
Population group	Categorical	Rural Migrant Urban	Rural: any individual aged 30years or more who permanently lives in the San Jose de Secce (rural study site) and was randomly selected from the updated census in the area Migrant: any individual aged 30years or more who reported to have been born in Ayacucho, permanently lives in Las Pampas de San Juan de Miraflores (urban study site) and was randomly selected from the updated census in the area Urban: any individual aged 30years or more who reported to have been born in Lima, permanently lives in Las Pampas de San Juan de Miraflores (urban study site) and was randomly selected from the updated census in the area	This group categorisation was used to answer the study's overall research question: "is there a difference in specific CVD risk factors in the rural-to-urban migrant group compared to those who did not migrate?"

Table II-6. Sub-classification of exposure variables in migrant group by length of exposure to urban area

Variable	Type	Categories	Definition	Comments
Length of residency in urban area	Categorical	Migrant <20 years in urban area Migrant 20-29 years in urban area Migrant 30-39 years in urban area Migrant ≥40 years in urban area	Migrant by length of residence in urban area, in absolute number of years. The migrant group was subsequently divided into five 10-year groups, using the variable "Years lived in urban area"	This group categorisation was used to answer study's specific research question: "Does the pattern of CVD risk factors comparing migrants with non-migrants vary by length of residence in urban environment?"
Lifetime exposure to urban area	Categorical	Quartile 1, lowest Quartile 2 Quartile 3 Quartile 4, highest	Lifetime exposure to urban area as a proportion of current age (number of years lived in an urban area divided over age) in quartiles.	As above, this group categorisation was used to answer study's specific research question: "Does the pattern of CVD risk factors comparing migrants with non-migrants vary by length of residence in urban environment?"

Table II-7. Sub-classification of exposure variables in migrant group by age at first migration

Variable	Type	Categories	Definition	Comments
Age at first migration	Categorical	Migrant aged ≤ 12 years old when first migrated Migrant aged > 12 years old when first migrated	Migrant by age at first migration: The migrant group was subsequently divided into two groups, using the variable "Age at first migration".	This group categorisation was used to answer study's specific research question: "Does the pattern of CVD risk factors comparing migrants with non-migrants vary by age at first migration?" Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults.

Table II-8. Definitions and measurement techniques used for clinical assessments

All measurements were made by the same trained personnel in both rural and urban areas.

Variable	Type	Definition / Measurement technique
Height	Continuous	Total height measured without shoes to the nearest 0.1 cm using stadiometer
Sitting height	Continuous	Sitting height measured to the nearest 0.1 cm using standard stools
Trunk length	Continuous	Calculated from sitting height minus stool's height
Leg length	Continuous	Calculated from total height minus trunk length
Leg/trunk ratio	Continuous	Calculated as leg length divided over trunk length
Weight	Continuous	Weight measured with the individual wearing light clothes to the nearest 0.05 Kg using SECA 940 electronic scale.

Table II-8. (continued)

Variable	Type	Definition / Measurement technique
Skinfolds, sites	Continuous	Measured in triplicate in each measurement site (biceps, triceps, subscapular and suprailiac) to the nearest 0.2 mm using a Holtain Tanner/Whitehouse Skinfold Calliper (http://www.anthropometer.com/tw.php). The average of three measurements was calculated for each measured site.
Waist circumference	Continuous	Measured in triplicate at the midpoint between the lower rib and the iliac crest. Measurements were made in the horizontal plane, while the participants were standing, using a tape measure to measure to the nearest 1 cm. The average of three measurements was calculated and used in the analysis.
Hip circumference	Continuous	Measured in triplicate at the point yielding the maximum circumference over the buttocks. Measurements were made in the horizontal plane, while the participants were standing, using a tape measure to measure to the nearest 1 cm. The average of three measurements was calculated and used in the analysis.
Blood pressure	Continuous	Systolic and diastolic blood pressure was measured using appropriate cuffs for arm circumference in the sitting position using the right arm, supported at chest level. Three measurements were made at least 5 min apart using an oscillometric device (Omron M5-i, Omron, Japan), validated for use in adult population [154].

Table II-9. Definition of outcome variables: behavioural risk factors

Variable	Type	Categories	Definition / Measurement technique
ALCOHOL			
Frequency of alcohol consumption	Categorical	Low High	Question asked: Frequency of consumption of any alcohol-containing drink in the last year. Low: Never drinks or drank one or less times per month High: Two or more times per month
Volume of alcohol consumption	Categorical	Low High	Question asked: Frequency of consumption of 6 or more alcohol-containing drinks in the same occasion in the last year Low: Never drinks or drank 6 or more alcohol-containing drinks in the same occasion less than once per month High: Drank 6 or more alcohol-containing drinks in the same occasion once or more than once per month
Frequency of hangover	Categorical	Low High	Question asked: How many times have you have had hangover in the last month Low: Never or less than once in the last month High: Once or more in the last month
Heavy drinkers	Categorical	No Yes	No: those who do not fulfil criteria for "yes" Yes: High volume of alcohol consumption or high frequency of hangover

Table II-9. (continued)

Variable	Type	Categories	Definition / Measurement technique
TOBACCO			
Current smoker	Categorical	No Yes	No: Has never tried cigarettes or has smoked less than 100 cigarettes in his/her lifetime (never smoker); or, smoked more than 100 cigarettes in lifetime and last cigarette was more than 6 months ago (former smoker). Yes: Has smoked more than 100 cigarettes in lifetime and last cigarette was less than 6 months ago.
Cigarette consumption in the last month	Continuous		Number of cigarettes consumed in the last month amongst current smokers.

Table II-10. Definition of outcome variables: anthropometric risk factors

Variable	Type	Categories	Definition / Measurement technique
Skinfolds, total	Continuous		Sum of all four (biceps, triceps, subscapular and suprailiac) averaged skinfolds measurements, in mm
Waist to hip ratio	Continuous		Calculated as waist circumference (cm) / hip circumference (cm)
Body mass index	Continuous		Calculated as weight (Kg) / height2 (m2)
Overweight	Categorical	No Yes	No: BMI < 25 Kg/m^2 Yes: BMI ≥ 25 Kg/m^2 and BMI < 30 Kg/m^2
Obesity	Categorical	No Yes	No: BMI < 30 Kg/m^2 Yes: BMI ≥ 30 Kg/m^2

Table II-11. Definition of outcome variables: blood pressure, lipids and inflammation risk factors

Variable	Type	Categories	Definition / Measurement technique
BLOOD PRESSURE			
SBP	Continuous		Mean of the last two SBP measurements
DBP	Continuous		Mean of the last two DBP measurements
Hypertension	Categorical	No Yes	Yes: SBP ≥ 140 mm Hg or DBP ≥ 90 mm Hg, or self report of physician diagnosis and currently receiving antihypertensive medication [155, 156]
LIPID PROFILE			
Total cholesterol	Continuous		Measured in mg/dL, in fasting conditions in serum. Technique: CHOD-PAP (Modular P-E / Roche- Cobas, Germany)
Triglycerides	Continuous		Measured in mg/dL, in fasting conditions in serum. Technique: GOD-PAD (Modular P-E / Roche- Cobas, Germany)
HDL	Continuous		Measured in mg/dL, in fasting conditions in serum. Technique: enzymatic-colorimetric methods (Modular P-E / Roche- Cobas, Germany)

Table II-11. (continued)

Variable	Type	Categories	Definition / Measurement technique
LDL	Continuous		In individuals with triglycerides below 400 mg/dL, LDL was calculated using the Friedewald equation [157, 158], in mg/dL: LDL = total cholesterol – HDL – (0.2 × triglycerides) In individuals with triglycerides greater than 400 mg/dL, LDL was measured in mg/dL, in fasting conditions in serum, using the following technique: enzymatic-colorimetric methods (Modular P-E / Roche- Cobas, Germany)
TC/HDL ratio	Continuous		Total cholesterol divided over HDL
INFLAMMATION			
CRP	Continuous		Measured in mg/L, in fasting conditions in serum. Technique: immunoturbidimetry (Modular P-E / Roche- Cobas, Germany)
Fibrinogen	Continuous		Measured in mg/dL, in fasting conditions in plasma. Technique: Coagulometry (CA-500, Dade-Behring, Germany)

Table II-12. Definition of outcome variables: glucose, insulin and metabolic-related risk factors

Variable	Type	Categories	Definition / Measurement technique
Glucose	Continuous		Measured in mg/dL, in fasting conditions in whole venous blood. Technique: GOD-PAP (Modular P-E / Roche- Cobas, Germany)
Insulin	Continuous		Measured in μIU/mL, in fasting conditions in serum. Technique: electrochemiluminescence (Modular P-E / Roche- Cobas, Germany)
Glycated haemoglobin	Continuous		Measured in %, in fasting conditions in whole blood with EDTA. Technique: high performance liquid chromatography (D10- BIORAD, Germany)
Diabetes	Categorical	No Yes	Yes: Fasting plasma glucose ≥ 126 mg/dL (or ≥ 7 mmol/L) [159] or self report of physician diagnosis and currently receiving antidiabetic medication.
Impaired fasting glycaemia	Categorical	No Yes	Yes: Fasting plasma glucose ≥ 110 mg/dL (or ≥ 6.1 mmol/L) and < 126 mg/dL (or < 7 mmol/L) [159].

Table II-12. (continued)

Variable	Type	Categories	Definition / Measurement technique
Insulin resistance	Continuous		Insulin resistance (IR) Calculated using HOMA calculator (Oxford Centre for Diabetes, Endocrinology & Metabolism, Diabetes Trials Unit, http://www.dtu.ox.ac.uk/) [160] and excluding those with diabetes. For HOMA-%S higher values indicate higher insulin sensitivity, or less insulin resistance, and are potentially less harmful to health. By contrast higher levels of fasting insulin are related to greater insulin resistance and are potentially harmful to cardiovascular health [161].

Table II-12. (continued)

Variable	Type	Categories	Definition / Measurement technique
Metabolic syndrome, WHO	Categorical	No Yes	Yes: WHO 1999 definition [159] Glucose intolerance, IGT or diabetes and/or insulin resistance (defined as fasting plasma glucose ≥ 110 mg/dL (6.1 mmol/L)) together with two or more of the following: 1. Blood pressure ≥ 140/90 mmHg 2. Raised triglycerides ≥ 1.7 mmol/L (150 mg/ dl) and/or HDL: Men: < 0.9 mmol/L (35 mg/ dl), Women: < 1.0 mmol/L (39 mg/ dl) 3. Obesity: Men: waist–hip ratio > 0.90, Women: waist–hip ratio > 0.85, and/or BMI > 30 Kg/m^2 4. Microalbuminuria*: Urinary albumin excretion rate ≥ 20 μg/min or albumin:creatinine ratio ≥ 30 mg/g * This study did not collected information on microalbuminuria.

Table II-12. (continued)

Variable	Type	Categories	Definition / Measurement technique
Metabolic syndrome, IDF 2005	Categorical	No Yes	Yes: IDF 2005 definition [162, 163] uses central obesity as a must criteria, and uses a lower plasma glucose threshold than in the WHO definition. Central obesity: Waist circumference (ethnicity specific)*, plus any two: 1. Raised triglycerides: ≥ 150 mg/dL (1.7 mmol/L) or Specific treatment for this lipid abnormality 2. Reduced HDL: <40 mg/dL (1.03 mmol/L) in men, <50 mg/dL (1.29 mmol/L) in women, or Specific treatment for this lipid abnormality 3. Raised blood pressure: Systolic ≥130 mm Hg, Diastolic ≥ 85 mm Hg or Treatment of previously diagnosed hypertension 4. Raised fasting plasma glucose: Fasting plasma glucose ≥ 100 mg/dL (5.6 mmol/L) or Previously diagnosed type 2 diabetes. * Waist circumference is ethnicity specific. For this study, South Asian recommendations (Male ≥ 90 cm, Female ≥ 80 cm) were used as suggested in the IDF definition.

Table II-12. (continued)

Variable	Type	Categories	Definition / Measurement technique
Metabolic syndrome, AHA/NHLBI 2005	Categorical	No Yes	Yes: AHA/NHLBI 2005 [164], a revised definition from the NCEP ATP III guidelines [165] that considers the individual's treatment status for specific conditions, and uses a lower plasma glucose threshold. Any three of the following: 1. Elevated waist circumference: ≥ 102 cm (≥ 40 inches) in men, ≥ 88 cm (≥ 35 inches) in women 2. Elevated triglycerides: ≥ 150 mg/dL (1.7 mmol/L), or Drug treatment for elevated triglycerides 3. Reduced HDL: < 40 mg/dL (1.03 mmol/L) in men, < 50 mg/dL (1.3 mmol/L) in women, or Drug treatment for reduced HDL 4. Elevated BP: ≥ 130 mmHg SBP or ≥ 85 mmHg DBP, or drug treatment for hypertension 5. Elevated fasting glucose: ≥ 100 mg/dL or drug treatment for elevated glucose

Table II-13. Definition of outcome variables: high-risk of cardiovascular disease

Variable	Type	Categories	Definition / Measurement technique
WHO/ISH high risk of CVD	Categorical	No Yes	Yes: high risk of a fatal or non-fatal cardiovascular event defined as 10-year risk > 20% based on WHO/IHS risk prediction charts for low- and middle-income countries [166, 167]. Risks were derived and aggregated using the chart AMR-D (Americas region, mortality strata D: high child mortality and high adult mortality), indicated for Peru [166].
Lancet's Chronic Disease Group high risk of CVD	Categorical	No Yes	Yes: high risk of fatal ischaemic heart disease or cardiovascular event defined as 10-year risk ≥ 15% and derived for Mexican population [18].
Framingham high risk of CHD	Categorical	No Yes	Yes: high risk of coronary heart disease defined as a 10-year risk ≥ 20% based on the Framingham equation [168]

2.4. Statistical analysis

The exposure variable, migration was generated based on three categories —rural, migrant (meaning people who had migrated from rural to urban areas) and urban— as indicated in the previous section. The rural group was used as the reference category for the main analysis. The migration variable was subsequently sub-divided by length of residence in urban environment, lifetime exposure to urban area as a proportion of current age, and age at first migration, using the lowest categories in each of them as baseline for comparisons. Only data from all individuals who completed the study were used in the analysis. Participants who completed the study were defined as those with completed questionnaires, clinical measurements and laboratory analyses (see section 5.1.1, Definitions used at each stage of study).

All data derived from this study follows recent STROBE guidelines for reporting observational studies [169]. Statistical analyses were carried out using Stata software version 10 (StataCorp LP, College Station, Texas, USA). To minimise errors in transcriptions, estimation results were converted to output datasets with one observation for each of a set of estimated statistical parameters using the `parmby` command within Stata [170].

2.4.1. Descriptive analysis

For general description of data, frequency analyses are presented as number (percentages), mean (± standard deviations (SD)) or median (interquartile range) when appropriate. Because of the study's matched-design by age-group and sex, no difference between calculations with and without adjustment for age and sex were expected in univariate analyses of categorical data, e.g. prevalence rates. This assumption was verified in all calculations, using direct standardisation against whole sample studied, and thus, such adjustment was not pursued for the reporting of categorical data.

Continuous non-normally distributed variables (triglycerides, CRP, fibrinogen, fasting glucose, HbA_{1c}, fasting insulin and HOMA insulin resistance) were log

transformed. Such logarithm transformation led to normal or near normal distributions. Age- and sex-adjusted arithmetic means (± SD) or geometric means (ratios) [171, 172] are presented. In the case of age, since the study-design only included participants from 30 years-old or more, a mid/centre age point was used such that age 45 years-old was used as the baseline in all age-adjustments.

2.4.2. Multivariable analysis

Multivariable logistic regression and linear regression were used for categorical and continuous outcomes respectively. Adjustment for treatment effects in specific continuous outcomes, e.g. antihypertensive therapy on blood pressure outcomes, was done using censored normal regression [173].

In logistic regressions results, odds ratios (OR) compare against the baseline exposure group of interest. For the linear regression dummy variables were created for the main exposure variable and other confounders when appropriate. Interpretations of categorical exposures for a continuous outcome were based on the β coefficients. β coefficients represent the average change of the outcome of interest, maintaining the units of measurement, in each category of exposure compared to the baseline group for that exposure.

Adjustment for potential confounding was done in a step-wise approach. A conceptual discussion of potential confounding related to the study is presented in this chapter (see section 2.7.3). Later on, the distribution and aggregation of measured socioeconomic indicators and its aggregation for their use as confounders is discussed in section 5.4.

R^2 for linear regression is also provided in each result table that reports outputs derived from multivariable analyses. R^2 is the proportion of variance in the outcome variable explained by the predictors. Of note, R^2 is an overall measure of the strength of association, and does not reflect the extent to which any particular independent variable is associated with the dependent variable [174].

Adjustment for treatment effects could potentially be used in the case of glucose as an outcome censoring those who are taking antidiabetic medication, or similarly, in

lipid traits censoring those on statins. In the case of lipids, none of the participants of this study reported to be on any lipid-lowering medication thus there was no need to make any adjustment based on lipid treatment. However, censored normal regressions, one of the approaches recommended for such treatment effects adjustments [173], has only been described for blood pressure. Such models make two key assumptions. First, it assumes that the underlying —not affected by medication— blood pressure is as least as high as the observed or measured blood pressure. Second, it assumes that the distribution of the underlying blood pressure above any given threshold in treated subjects is the same as the corresponding blood pressure distribution of those untreated, implying a non-informative censoring [173]. The second assumption can be criticised. It is because of such assumption that this approach for censoring was not used in the case of glucose and antidiabetic medication. Basically, glucose presented a skewed distribution and the proportion of those on medication was very low. Thus, censored normal regression was only be used in the case of blood pressure as previously described [173].

2.4.3. Standardised mean differences

To answer the study's main and specific research questions, which evaluates if there is a difference in specific CVD risk factors in the rural-to-urban migrant group compared to those who did not migrated, odds-ratios (OR) and standardised mean differences (SMD) were calculated. In categorical outcomes, OR and 95% confidence intervals (CI) were calculated using logistic regression. In the case of continuous outcomes, SMD were chosen because of its advantage to interpreting results of continuous data measured with different scales or units, thereby facilitating comparisons of difference sizes for individual measures. The Cochrane Collaboration has defined SMD as "the difference in means between two groups, divided by the pooled standard deviation of the measurements" and suggests that "the value of a SMD thus depends on both the size of the effect (the difference between means) and the standard deviation of the outcomes (the inherent variability among participants)" [175]. In this study, a slight variation in the normal SMD calculation was made to be able to include the comparison of more than two groups.

SMD were calculated through a regression to a within-group SD scale which took into account the variation in all three rural, migrant and urban groups. As such, the denominator is not a pooled SD of measurement of two groups only but it is very similar to the SD of the whole population studied. SMD regressions were carried out in fully adjusted models taking into account age, sex, individual's socioeconomic indicators and parental education.

In this sense, all comparisons would be expressed as differences in units of SD of normally-distributed variables or as differences in units of SD in the log scale of transformed variables. Due to the lack of units, these SMD allows for comparison of magnitude of differences across various risk factors. In terms of the interpretation of SMD, the Cochrane Collaboration indicates that "rules of thumb exist for interpreting SMD (or 'effect sizes')… 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 a large effect [175, 176]".

2.5. Institutional support and funding

The study was carried out in collaboration between academic institutions from Peru and the UK, specifically the School of Medicine and the School of Public Health, Universidad Peruana Cayetano Heredia in Lima, Peru, and the Department of Epidemiology and Public Health, London School of Hygiene and Tropical Medicine, London, UK. The coordination of fieldwork in Peru was made in conjunction with Asociación Benéfica PRISMA and access to its office in Pampas de San Juan de Miraflores was provided.

The study was funded by the Wellcome Trust through a Masters Research Training Fellowship and PhD Studentship for Dr J. J. Miranda Montero (Grant number GR074833MA).

2.6. Ethical considerations

Ethical approval was obtained from ethics committees at Universidad Peruana Cayetano Heredia in Peru and London School of Hygiene and Tropical Medicine in the UK.

The purpose of the study was explained to each of the study participants and informed consent was obtained, following international standards for ethical research in developing countries [177, 178]. The patient's information leaflet and informed consent forms revised and approved by the Ethics Committees in both, English and Spanish versions, are provided as Appendix A and Appendix B, respectively.

2.7. Sources of bias and confounding

Error is a false or mistaken result obtained in a study or experiment. The three objectives of a sound design are to minimise selection and information bias, to control confounding, and to attempt to rule out chance [179, 180]. This section will focus on the discussion of the nature, extent and actions to be taken to minimise systematic error or bias and address confounding, issues that could affect the design of the present study, a migrant study. It is considered that the role of chance —random error— has been dealt with in the previous section related to sample size. The discussion of selection bias in relation to the population that participated in this study is discussed in Chapter V.

2.7.1. Selection bias

Selection bias is a major issue of concern in the design and interpretation of migrant studies [58, 62]. Selection bias would occur in the event that participants of a study substantially differ from the general population or group that they belong to. It has been previously mentioned that migration is not necessarily a random process, existing the possibility that migrants could be a self-selected group —socially or economically different from their peers who do not migrate— and not necessarily a representative group. In the same vein, the evaluation of migrant groups in host environments could be prone to misleading observations. For example, it could be possible to have situations where a better health status is observed amongst migrant groups, which could be further explained by the migration of healthy individuals, by the return of those ill to their places of origin, or by a mixture of these two situations.

Due to the unique circumstances of the Peruvian context, where a forced migration process occurred, it would be expected that a wide diversity and majority of people from the rural part of Ayacucho had strong pressures to migrate, and not only the better-off —biologically and socioeconomically— sectors of this population. Additionally, rural and urban control groups were defined a priory to match the rural area of origin of most migrants as well as their urban destination. In this sense,

although unavoidable, this study has benefited from a less extreme type of selection bias amongst the migrant populations in comparison to other studies of the same nature.

2.7.2. Information bias

The estimate of the strength of the association between exposure and outcome may be biased as a result of inaccurate information being collected on either the exposure or the outcome or both [180]. The probability of misclassification may be the same in all study groups (non-differential misclassification) or may vary between groups (differential misclassification) [55].

2.7.2.1. Non-differential misclassification

The exposure used in the present study is a well-defined one, migration from rural to urban areas, and therefore, major problems with the ascertainment of exposure were not expected a priori. The selection of migrants from a defined place of origin (Department of Ayacucho, Peru) and currently settled in a defined area in Lima was useful to select similar comparison groups who did not migrated in both, rural and urban areas.

On the contrary, the assessment of the different outcomes of interest might pose some challenges, and errors in its measurement could produce varied effects depending of the direction of the misclassification. To the benefit of this proposal, most of the outcomes of interest were obtained through measurements using trained personnel, standardised equipments and/or biochemical assays.

2.7.2.2. Differential misclassification

This type of bias could occur when the probability of outcome status being misclassified depends upon exposure status or vice versa [180]. Recall bias and observer or interviewer bias are the commonest sources of bias of this type. Minimisation of recall and observer bias was addressed by using simple standardised questionnaires previously piloted and a well-defined protocol.

2.7.3. Confounding

Confounding, by definition, "occurs when a relationship between the effects of two or more causal factors as observed in a set of data such that is not logically possible to separate the contribution that any single causal factor has made to an effect" [55]. In other words, a confounding variable is one that is both associated with the exposure and independently associated with the outcome or disease, and does not lie on the causal pathway between exposure and outcome of interest [180]. This section discusses potential confounder variables in the association of interest for this study.

2.7.3.1. Socioeconomic status

It is largely recognized that lower socioeconomic position is associated with poor health [181]. Whatever social classification system is adopted, social disadvantage is consistently associated with poorer health [182]. In developed countries it has been shown that the lower the socioeconomic position the higher the cardiovascular risk [183, 184]. On the contrary, in some developing countries, obesity and other cardiovascular risk factors have been described to be more prevalent amongst upper social classes, although this pattern is changing in some societies [185-187].

The scenario is more complicated in the case of studies of migrant populations. The propensity to migrate is linked to health status and socioeconomic position. It has been described in the UK that areas with high mortality were also areas with higher net out-migration [149, 188]. A series of factors are involved, and the main potential explanation could be that those individuals who are able to migrate could be part of a better socioeconomic strata, therefore they could "afford" —via self-confidence, social networks and/or economical means— to do it. In the event of established violence, as described in the present study, it could well be that the better-off migrants-to-be considered —and probably took— at an earlier stage the option to migrate. Similarly, people in rural areas may have perceived additional gains or benefits of migrating considering the societal status of violence a few decades ago. Socioeconomic status is, therefore, associated with the exposure migration.

From this analysis of the available evidence applied to the Peruvian context, it is being considered, a priori, that socioeconomic status would be a confounder variable. For that, in this study we have considered measurements of education and income as proxies for socioeconomic status [139-142]. A life-course approach has been adopted in the measurement of socioeconomic position in this study [139-141]. Father and mother's education level were used as a proxy for childhood socioeconomic position while individual's education level attained, overcrowding, assets possession and family income were used as markers of adulthood socioeconomic position.

2.7.3.2. Mental health

The process of urbanisation in developing countries has been described to affect mental health [189]. Recently, a large international case-control study has made the case for the association between psychosocial risk factors, mainly stress, and CVD, specifically myocardial infarction [190]. Controversy still remains in this area, where little evidence has been described for the role of psychosocial factors in the causation of physical disease, in contrast to material deprivation [191].

It is recognised that mental health, understood as self rated health or levels of adaptation —also referred as acculturation in specialised fields— in the host environment, could play a role in the decision to migrate as well as being associated to the outcomes being studied. In the specific case of migrant populations, those who have a better self-perceived health or those who are able to better cope with social change, could be more prone to engage in a process of migration, as well as to maintain healthier life styles. In this study, mental health has not been considered confounder but as a variable that lies on the causal pathway. Once migrant groups settle in a new unknown environment, with limited social networks and social support, increased chances of engaging in "unhealthy" behaviours —such as alcohol drinking and smoking— can happen, that in turn can be reflected a negative cardiovascular risk profiles.

2.7.3.3. Acculturation

Acculturation has been considered as the adoption of health behaviours from the new dominant culture and loss of health behaviours from the original culture [192, 193].

As demonstrated elsewhere [194, 195], a simple acculturation scale can be a very informative variable in the potential associations between migration and cardiovascular risk. Acculturation in the US, measured by language spoken at home, place of birth, and years living in the US, has been associated with a higher prevalence of hypertension and other CVD risk factors [196, 197]. In the UK, acculturation has been explored in relation to likelihood of mothers engaging on unhealthy behaviours, including smoking and alcohol consumption [69]. Acculturation as a construct, however, has been criticised in that it is a concept that looks at the individual and does not take into account the social complexity of relationships with environments and society. Viruell-Fuentes contends that "socio-cultural explanations for this apparent epidemiological paradox propose that culture-driven health behaviours and social networks protect the health of the first generation and that, as immigrants acculturate, they lose these health-protecting factors. However, the prominence granted to acculturation within these explanations diverts attention from structural and contextual factors, such as social and economic inequalities, that could affect the health of immigrants and their descendants" [198].

One of the major limitations of acculturation is that it has always been explored in the context of international migration and no instrument has been tested for internal within-country migration or rural-to-urban migration. In addition to that, in this study such variable could only be addressed on exploring changes within migrants if differences were to be observed. As that was not one of the objectives of this research, this variable, although collected, was decided not to be included in the analysis of this study.

2.7.3.4. Altitude and haemoglobin

While both migrant and urban groups studied currently live at sea-level, the rural group is located in the Andes at an altitude around 2,500 meters above sea-level. Few studies have considered the potential association between high altitude and CVD outcomes. One of them, the Yi Migrant Study, treating altitude as categorical variable, looked specifically at the impact of migration on blood pressure [86]. Other examples appear in physiology or high altitude studies where prevalences of high blood pressure [199, 200] and metabolic syndrome [117] have been reported in high

altitude settings. None of these high-altitude studies however, have explored the presence or absence of an association between altitude and the outcome studied beyond the reporting of prevalences.

It is well-known that red blood cell and haemoglobin levels are higher in individuals living at high altitudes as a compensatory mechanism for the low environmental oxygen [199]. The references cited could support the case could that high-altitude has a role on blood pressure, either by treating altitude as a categorical variable (yes/no to living in altitude) or using haemoglobin level (continuous variable) as a surrogate marker for altitude. However, the scenario is much more complicated as to find potential explanatory roles for high-altitude in the wider array of CVD to be explored in this study. For example, it is very improbable that altitude itself or haemoglobin levels can provide credible explanations for differences in body mass index or inflammation markers in this study. Under these circumstances, it was decided not to include controlling for high altitude as a confounder variable to ensure comparability of results across risk factors studied.

Chapter III. Pilot Study

This section describes the conduct of a pilot study in a small sample of participants in both the rural and urban site and lessons learnt in this process.

3.1. Objectives

The main objective of the pilot study was to assess the scientific and logistical feasibility of the proposed study. Scientific, to explore whether or not the differences hypothesized were relevant to be studied or a change in selected outcomes was necessary. Logistical, to explore if the main study could be well conducted in two different settings by the investigator.

A secondary objective was to pilot the questionnaire to be used in the study.

3.2. Execution of pilot study

In April and May 2006 a pilot study was carried out in Lima (urban setting) and in Huanta, a small city in Ayacucho (rural setting). The city of Huanta was chosen because of its convenient location in terms of transport and because its hospital is the head of the micro-network of health services in that area. Thus, in addition to the pilot itself, this visit enabled the investigator to initiate co-ordinations with the health sector and request the necessary permissions.

[For clarification purposes, the final rural site chosen for the main study —San Jose de Secce in Santillana— is a smaller village that belongs to the administrative boundaries of Huanta where the pilot study was conducted. This village is located 3-4 hours inland from Huanta.]

3.3. Results from pilot study

This pilot included 38 adult people in total: 20 in Huanta and 18 in Lima (11 migrants and 7 urban residents). The results, in terms of general demographics, anthropometry and laboratory results are shown in Table III-1.

Results of this pilot show a lower degree of years of education in the rural and rural-urban groups as well as in all female groups. In terms of behavioural risk factors, the rural and rural-urban groups show higher reporting of never smoking status. It is difficult to analyse the alcohol patterns based on this smaller numbers. In terms of BMI, a tendency of increased BMI in both sexes from all groups is observed, being the urban group the one with higher average BMI. A similar pattern is seen in waist-to-hip ratio for males but not females. A higher number of male participants were found to be hypertensive.

As in BMI, a similar increasing pattern of fasting glucose, total cholesterol and fasting insulin is observed in the three groups when looking at the rural to urban spectrum.

Being a small sample is difficult to draw any major conclusion.

Table III-1. Demographic, anthropometric and laboratory results from pilot study, 2006

	Rural n = 20		Migrants n = 11		Urban n = 7	
	Male	Female	Male	Female	Male	Female
	n = 8	n = 12	n = 6	n = 5	n = 2	n = 5
Demographic and socioeconomic variables						
Age mean	43	40	50	45	42	42
Age range	32-62	26-61	26-72	39-51	39-44	35-52
Years education average	7	6	12	7	13	12
Years education range	1-16	0-11	4-18	0-14	11-14	8-16
Behavioural risk factors						
Ever smoked tobacco daily?						
Never	4	12	6	5	–	5
Ex-smoker (6mo ago)	3	–	–	–	1	–
Current	1	–	–	–	1	–
Regular drinking alcohol more than 10 days per month						
Never	3	7	5	5	–	5
No longer (6mo ago)	5	5	1	–	–	–
Current	–	–	–	–	3	–
Pattern of drinking						
During the week	–	–	–	–	–	–
Weekends	3	1	–	–	2	–
Last 3 months	2	3	1	2	2	2

Table III-1. (continued)

	Rural n = 20		Migrants n = 11		Urban n = 7	
	Male	Female	Male	Female	Male	Female
	n = 8	n = 12	n = 6	n = 5	n = 2	n = 5
Anthropometric and clinical variables						
Height (m) average	1.57	1.50	1.56	1.55	1.70	1.56
Weight (Kg) average	60.3	59.7	61.4	72.2	78.2	70.9
BMI (Kg/m^2) average	24.5	26.4	25.1	30.1	27	29.3
Obese (BMI≥ 30 Kg/m^2) (n)	1	3	–	2	–	1
Waist circumference (cm)	84.7	163.6	90.1	95.3	97.1	93.1
Hip circumference (cm)	90.4	97.6	90.9	105.5	97.4	107
Waist/Hip ratio mean	0.94	1.68	0.99	0.9	0.99	0.87
Blood Pressure						
Systolic (mmHg) mean	134	113	127	119	127	123
Diastolic (mmHg) mean	85	68.8	79.3	76.8	83.8	81.7
Hypertension (SBP >140 or DBP >90) (n)	4	1	1	–	1	2
Laboratory tests						
Fasting Glucose (mg/dL) mean	85.8	78.9	87.7	90	91.5	90.4
Total Cholesterol (mg/dL) mean	176.9	173.5	185.8	219.6	221	198
HDL Cholesterol (mg/dL) mean	42.6	46.4	39.3	44.9	37.6	46.0
Triglycerides (mg/dL) mean	150	131.9	146	176.2	245.5	148.2
LDL Cholesterol (mg/dL) mean	104.3	100.7	117.3	139.4	134.3	122.3
HbA_{1c} (%) mean	5.6	5.4	5.4	5.64	5.5	5.5
PCR (mg/L) mean	2.5	2.1	2.9	2.64	3.5	2.7
Fasting Insulin (μIU/mL) mean	6.6	8.7	11.6	12.6	9.5	23.7
Fibrinogen (mg/dL) mean	221.5	323.9	314.8	385.8	255	312.4

3.4. Lessons learnt from pilot study

A few lessons, important for the conduction of the main study later on, were drawn from the pilot experience. These lessons can be aggregated as lessons for improvement of the study proposal and lesson for the logistical organisation and conduction of the study.

All the lessons were taken into account and action was taken upon them during the planning of the main study. Therefore an explanation on how the study was re-shaped is also provided.

3.4.1. Lessons related to study proposal

- First, the most important one, the results of this very small study points towards the proposed hypothesis of this study, that cardiovascular risk factors increase following rural to urban migration.

- Second, the data suggest that a gradient also exists in education level between groups. As discussed in the "Sources of bias and confounding" section in the previous chapter, socioeconomic status could be a factor that is not similar across groups. In this sense, the questionnaires were modified to include education level at the moment of migration as well as parental educational status as proxy indicators of childhood socioeconomic position. In addition to this, an extra question aiming to explore, in the migrant group, whether or not such migration process had an impact on education level attained, e.g. if migrants had to stop their studies, continued their studies or started new studies.

- Third, the questionnaire piloted included a reduced version of a nutritional survey as used on the WHO STEPS approach [201]. The use of such instrument proved difficult as misconceptions are prevalent about what constitutes a salad, to name an example. To avoid this more detailed food/nutrition instruments would have been the best option to choose. However, since diet was not the main outcome of the study it was decided to drop this survey component.

- Fourth, the pilot study did not record any information of response rates, and this was an important limiting factor in interpreting the pilot study's results. Based on this experience it was decided that for the main study a short rejection form would be elaborated and implemented to be completed by those who were randomly selected but did not agreed to participate in the study. This would capture some information about survey non-responders to address potential response bias.

3.4.2. Lessons related to study execution

- The pilot study was conducted by the main investigator aided by two research staff. It became clear that a larger research team was necessary in order to achieve the study targets. Also, a Quechua speaking team was also required to conduct the study in the rural site. Thus, as explained later on in section "4.1.6 Logistical organisation of research team", a larger research team with specific tasks was assembled.

- Invitations to take part on the study based on postal questionnaires were not a feasible option ever since such service is unavailable in the study areas. Telephone communication was also discarded because such data was not available for the majority of participants. Thus, in order to reach potential participants the only viable strategy was through household visits.

- In terms of laboratory procedures, a challenging task was to ensure that the transport of blood samples from the rural site to a laboratory was made within reasonable time, mostly to avoid glucose degradation. The first option explored was to use local hospital laboratory services; however such change would have introduced a potential measurement bias by using different laboratory techniques and quality control standards. A consultation was carried out with the laboratory in charge of the blood tests together with the evaluation of frequency of transportation from Ayacucho into Lima. It was agreed that all samples from the rural study site would be sent daily into Lima and would receive priority in the processing. It was also based on this experience that the same degree of carefulness was also implemented in the urban study site. Such policy was

introduced and, as a result, only one glucose sample was lost in a participant from the rural group.

- It was necessary to conduct the pilot study in the rural site avoiding the rainy season (January – May) to ensure adequate transportation of study team and blood samples. Plans were made to ensure the main study in rural areas avoided the rainy season.

Chapter IV. Execution of main study

This section describes the data collection methods used for the main study as well as its organisation and execution. This chapter closes with a presentation of achieved study sample size compared to the initial target.

4.1. Data collection methods

Information required for the study was collected through questionnaires and direct measurements using trained clinical and non-clinical fieldworkers to ensure standardisation in measurements, use of questionnaires and blood sampling.

4.1.1. Questionnaires

Questionnaires were constructed after checking relevant work for the study. Instruments that address objective and subjective social status [100, 113, 182, 202-205], behavioural risk factors [206], reproductive health [207] and mental health [208] were reviewed. An important material used was the WHO STEPwise approach to Surveillance (STEPS) [201].

For the fieldwork, the questionnaires were prepared and piloted in Spanish. Particular advantage will be taken from the Demographic Health Survey's Household Assets questionnaire [100] and the Center for Disease Control and Prevention's Behavioral Risk Factor Surveillance System Survey Questionnaire [206], both of which have versions available in Spanish.

The final version of the full questionnaire elaborated and used in the study is presented in Appendix C. In addition to the full survey, a rejection form was also elaborated for non-responders, shown in Appendix D.

4.1.2. Direct measurements

Detailed explanation of clinical measurements, e.g. weight, height and blood pressure measurements has already been provided in "Table II-8. Definitions and measurement techniques used for clinical assessments" in page 51.

A specific standard operation procedures' protocol for anthropometric, skinfold and blood pressure measurements was elaborated for the research study team (see Appendix E).

All the study team was trained on the measurement of waist and hip circumference as well as in the use of skinfolds. The best three performers were selected as responsible for these measurements following standardised protocols for the evaluation of accuracy and precision in anthropometric measurements [209, 210].

In relation to the measurement of sitting height, 95.5% of all of these measurements were made using two specified equipments. In the urban site, a single standard 44 cm stool that fitted with standing stadiometer was used to measure sitting height (580/589, 98.5% in migrant group and 197/199, 99% in urban group). In the rural site a single standard 38.8 cm stool from the health clinic was used in 167/201 (83.1%) of total observations. The remaining observations were used different stools, ranging from 31 to 55.1 cm, and these stool lengths were also recorded to be able to calculate leg length.

4.1.3. Laboratory assessments

All laboratory assessments were performed on venous samples taken in the morning after a minimum of 8 hours fast. Venous blood was drawn into 5 different collection tubes, 3mL each, using vacutainers.

Blood testing was carried out by a certified laboratory (Medlab, www.medlab.com.pe) using their qualified personnel to do all sampling procedures and to be in charge of transport of samples to the laboratory's facilities. Medlab's fulfilment of laboratory quality control analysis has been made independently by US-based company Bio-Rad (www.biorad.com). The investigators had access to previous Medlab's quality control evaluations and reports. In brief, quality control procedures involved the supply of samples with exact and known levels, and the Medlab's measurements on these samples were compared against a reference laboratory. Most laboratory results for Medlab were within a good quality reference range, that is, ± 1 standard deviation and < 1 for the coefficient of variation, a reference range provided by the independent assessment company. (Quality control reports for glucose, HDL, total cholesterol and insulin tests are presented as examples in Appendix F, Laboratory performance overview).

Medlab's personnel followed a pre-established quality control protocols for health and safety and for management of blood samples for this study (see Appendix G, Field and laboratory processes for blood testing and storage of samples). All study personnel and laboratory personnel were aware of these procedures. This protocol outlined all laboratory tests to be carried on each subject's blood sample, including the minimum amount of venous blood sampling necessary to secure that enough blood, serum and plasma were available for the assays. Additional carefulness was placed in processing blood samples for the measurement of glucose in plasma by using a separate tube for this specific assay (see Appendix G, page 441).

In Lima, Medlab's laboratory personnel visited the research clinic and took directly the blood samples. One of Medlab's personnel travelled with the research team to the rural site and ensured that the same quality control protocol was followed in the rural site. All blood samples, from all participants in both study sites, were sent to the laboratory headquarters within 24 hours of sampling, and tests were conducted as soon as they arrived to the laboratory facilities.

Serum glucose, insulin, CPR, total cholesterol, HDL, and triglyceride levels were measured using a Cobas® Modular Platform automated analyser and reagents supplied by Roche Diagnostics [211]. The same procedure was used to measure LDL in people with triglycerides greater than 400 mg/dL (n=29, 2.9% of all participants studied). In those with triglycerides below 400 mg/dL, LDL was calculated using the Friedewald equation [157, 158]:

$$\text{In mg/dL: LDL} = \text{total cholesterol} - \text{HDL} - (0.2 \times \text{triglycerides})$$

$$\text{In mmol/L: LDL} = \text{total cholesterol} - \text{HDL} - (0.45 \times \text{triglycerides})$$

HbA_{1c} was measured on fasting whole blood and analysed by High Performance Liquid Chromatography using automated D-10 Bio-Rad haemoglobin analysers [212]. Plasma fibrinogen was measured by coagulometry technique using a Dade-Behring CA-500 analyser.

Laboratory result outputs were available electronically on an intranet site with access to the main investigator through a secured password. An excel spreadsheet with all

participants laboratory tests and two printed copies of laboratory results were provided to the research team on a regular basis. A copy of the results was later given to the participants and the other copy was archived.

4.1.4. Study organisation and execution

4.1.5. Census update

The main objective of this activity was to identify the sampling frame for the proposed research in the urban area in Lima. This issue is of particular importance since Lima hosts nearly 9 million people, a third of total Peruvian population. It will be impossible to use the whole city of Lima as the primary sampling unit for the sampling process. The specifics of this project, looking at migrant population, also require a focused approach.

Based in previous research undertaken in one of Lima's shantytowns it was decided to conduct the study in part of the district of San Juan de Miraflores. Asociación Benéfica PRISMA carried out a census of the area called Pampas de San Juan in 1997 with further update in 2000. The settlement in this area started in 1980's, which coincides with the periodicity of the mass migration observed in Peru detailed in the background section.

Based on information from these censuses in Las Pampas de San Juan de Miraflores, a total 43,891 inhabitants were counted, organised in 37 settlements. Information on place of birth was available for 99.7% of them, and a total of 4,738 people of all ages responded to have been born in Ayacucho.

An updated census was co-ordinated and carried out by the investigator in January and February 2006. Concretely, the specific aim of an updated census was to identify adult people for the rural-urban group, defined as those who were born in and migrated from Ayacucho into Lima, and who were permanently living in the area of study.

After updating records for place of birth, further migration and other factors, a total of 1,264 people aged 25 or more were found. Information about reasons for

migration, e.g. socioeconomic (job or studies) or forced migration (afraid or terrorism) were also obtained. This updated census was used as the primary platform for random selection of individuals for the rural-urban group.

In the case of San Jose de Secce in Ayacucho, a small-scale census was conducted in mid 2007 to identify all adult population permanently living in the area. In this census in the rural area, 398 adults were identified and this was used as the sampling frame for the rural group.

4.1.6. Logistical organisation of research team

Apart from the investigator, a medium-size fieldwork team, including 10 full-time people and some extra part time at specific times in the study, was assembled to complete the tasks set out in this project in both urban and rural study sites. Key personnel included two doctors, two nurses, one data manager and various fieldworkers. The later group comprised individuals that live in the same community of the urban study site who were trained as community health workers in the past and had experience in other research studies.

Once training and piloting was completed in the research office responsibilities were assigned to various members of the team, such as fieldwork enrolment, clinical appointment and data monitoring.

The main investigator of this study monitored and supervised all activities in both sites on a daily basis.

4.1.6.1. Fieldwork enrolment

Briefly, the fieldwork team had the responsibilities to locate in the community through household visits the selected individual from the census, invite them to the study, read consent forms and make an initial clinical appointment and apply questionnaires to participants. Fieldwork personnel would also have the task to revisit up to three times the randomly selected individual. If the selected person was not found after a third visit, he/she was dropped from the study as a not contacted

person. Based on the design of the study —age and sex matched— a replacement subject from the same age- and sex group was randomly selected.

If at any point a randomly selected individual was found but did not agree to participate in the study, they were considered as contacted but refused to take part in the study —a type of non-responders— and were asked to fill a short rejection form with basic demographic and socioeconomic information (form shown in Appendix D, presented in next chapter). If the non-responder did not agree to complete such information, the team fieldworker had the task to try to fill as much information as possible from observation or through neighbours' information.

In some circumstances the participant will initially agree to take part on the study and gave their informed consent. This was considered an enrolled participant. However, for various reasons, mostly time constraints, study fieldworkers were not able to regain contact with the enrolled participant thus affecting the completion of the study. For this, fieldworkers were also asked to revisit the selected person up to three times after the person was considered an enrolled participant. All these definitions are discussed in the next Chapter, in the section related to response rate.

For practical and working reasons for some participants, it was easier for them to attend the research office during the weekend. Mechanisms were put in place to have an effective working team during those weekends and to be able to work with an increased flow of participants. Some of these tasks included use of motorised transport to collect participants and an increased number of fieldworkers in research office to ensure completion of surveys.

4.1.6.2. Laboratory and clinical measurements

The clinical team had the responsibility to complete all clinical measurements and laboratory blood samples following standardised protocols. Guidelines for measurements and definitions used have already been presented (see section 2.3, Data requirements).

At the beginning of the study it was observed that some people feared the amount of blood drawn for laboratory tests. Thus, fieldwork personnel were explained and

trained to provide explanations about blood testing. These included reasons to use different tubes and practical exercises showing the real volume of blood that filled each tube. A practical local explanation was that each tube contains the volume "equivalent of a tea spoon". Such information was then transmitted to people in the community, and additional emphasis in explaining reasons for blood testing was placed at the research clinic at the moment of the laboratory tests itself.

A free consultation clinic was arranged to return laboratory results to each participant and provide free interpretation of their tests, including data from blood pressure and BMI. If a subject was identified with abnormal levels of known risk factors, usually high blood pressure, high cholesterol or high glucose levels, they were referred to the local health post. This was done by the investigator and two medically trained personnel who were part of the study. If a participant did not collect his/her laboratory test results at the research clinic, it was then transferred to the fieldworkers to return those tests at each of the participant's households.

4.1.6.3. Data management

A lesson from early stages of fieldwork was to ensure a continuous monitoring system. This was observed once fieldworkers were set out to visit the community to enrol participants. Some difficulties observed during fieldwork activities included the need to revisit households of selected individuals, the difficulty to locate some participants, and the need to make quick decisions once a subject was classified as a non-responder. A list of randomly selected replacements for non-responders (age and sex matched individuals) were prepared in advance and provided to fieldworkers.

A more challenging issue was to keep track of all households' visits and revisits. Initially all household visits were recorded in paper and managed by the fieldworker. However, this system was difficult to update and risked the loss of information. It was decided that a data monitoring team was necessary to keep all household visits in an electronic dataset. Such tasks were initially delegated to two clinical fieldworkers and later a data management worker joined the study.

The data monitoring team had the responsibility to monitor daily progress in various parts of the research (enrolment, questionnaires, and laboratory) and ensure that data

was fully recorded. A more strict monitoring system was placed for the supervision of the enrolled individuals to ensure completion of the study (see for example Figure IV-1, page 98).

A standard double-entry method was used for inputting all data collected through questionnaires, a process which was also supervised by the data monitoring team. Laboratory tests results were also double checked for outliers, comparing the data provided in the electronic spreadsheet against the printed data and the intranet information site.

4.1.6.4. Study execution in rural study site

Once the mechanisms were tested in a number or urban-resident participants, planning for the study in the rural site took place. For the rural site, local Quechua speaking people were hired to take on the same fieldworker's tasks as their counterparts in Lima. Two members of the team were fluent in Quechua, being themselves migrants or descendents from migrants, and they were part of the team in the rural study site providing training to local fieldworkers. Co-ordinations were made with the local health centre in San Jose de Secce, rural study site, which agreed to provide space for the clinical measurements.

As shown in the table in the following page, most interviews in the rural site were conducted in Quechua, the local language, whereas in the urban site most interviews were carried out in Spanish.

Table IV-1. Language used at interviews

	Spanish		Quechua		Both		Total
	n	%	n	%	n	%	n
Rural	20	10	156	77.6	25	12.4	201
Migrant	575	97.6	2	0.3	12	2	589
Urban	197	99	0	0	2	1	199
Total	792	80.1	158	16	39	3.9	989

4.1.7. Period of study execution and recruitment rate

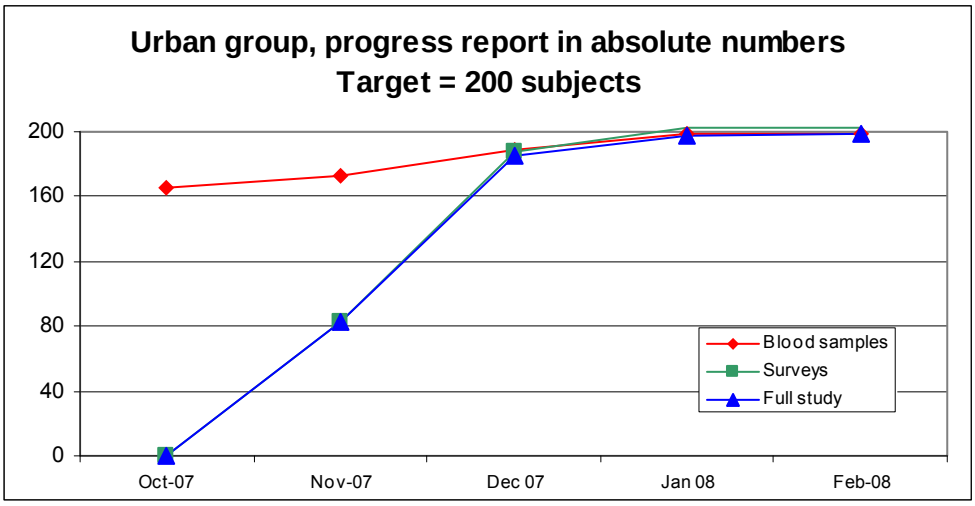
From January to June 2007, a preparatory phase was conducted for training of fieldwork personnel, logistical arrangements and co-ordinations for the different component of the study.

Once all arrangements were in place, the main study took place from July 2007 until January 2008. It started initially in Lima, aiming for a total recruitment target of 800 participants (see next section on study's targets). Figure IV-1 and Figure IV-2 show the progress of the study in the urban site, in both the urban and migrant groups.

Once the team was more confident with the study processes, a parallel group conducted the study in the rural site, with a target of enrolling 200 participants (see next section on study's targets). This task was accomplished in two months; from October to December 2007 (a detailed graph is not available as the task was shorter in time).

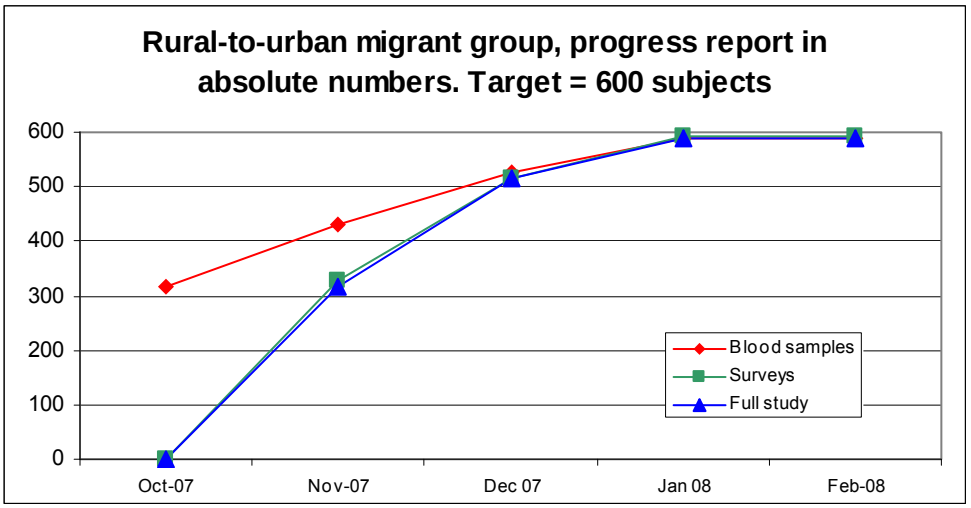
The continuous ongoing monitoring process in the urban site was set-up in order to have continuous and up-to-date information on study's progress. The goal was to achieve 100% target completion defined as participants with completed questionnaires and laboratory testing. Nearly 90% of the total study was accomplished by December 2007.

Figure IV-1. Monthly monitoring of fieldwork progress, urban group



Note: Data is shown from the point when the monitoring system was made into electronic records. This enabled an ongoing supervision and accomplishment of targets.

Figure IV-2. Monthly monitoring of fieldwork progress, rural-to-urban group



Note: Data is shown from the point when the monitoring system was made into electronic records. This enabled an ongoing supervision and accomplishment of targets.

4.2. Target versus achieved sample size

This is a brief section to present the target versus achieved sample size. As mentioned in Chapter 2, the study target was to recruit a total of 1000 people, i.e. 200 people in each of the rural and urban groups and 600 people who migrated from rural to urban areas. These participants were age and sex matched.

Table IV-2, Table IV-3 and Table IV-4 show the target and achieved sample sizes by sex and age-group in each of the study groups: rural, migrant and urban.

The final total sample size achieved was 98.9% of the target sample size (989/1000). Of these, 52.8% (522/989) were females.

Table IV-2. Distribution of rural study group by age and sex

Rural Group, target n = 200

Projected				Achieved			
Age group	Male	Female	Total	Age group	Male	Female	Total
30-34	15	15	30	30-34	15	16	31
35-39	15	15	30	35-39	16	14	30
40-44	15	15	30	40-44	15	16	31
45-49	15	15	30	45-49	13	12	25
50-54	15	15	30	50-54	8	14	22
55-59	15	15	30	55-59	9	15	24
60+	15	15	30	60+	19	19	38
Total	105	105	210	Total	95	106	201

Table IV-3. Distribution of rural-to-urban migrant study group by age and sex

Rural-to-urban Migrant Group, target n = 600

Projected				Achieved			
Age group	Male	Female	Total	Age group	Male	Female	Total
30-34	43	43	86	30-34	31	50	81
35-39	43	43	86	35-39	36	47	83
40-44	43	43	86	40-44	48	41	89
45-49	43	43	86	45-49	41	43	84
50-54	43	43	86	50-54	50	42	92
55-59	43	43	86	55-59	32	43	75
60+	43	43	86	60+	42	43	85
Total	301	301	602	Total	280	309	589

Table IV-4. Distribution of urban study group by age and sex

Urban Group, target n = 200

Projected				Achieved			
Age group	Male	Female	Total	Age group	Male	Female	Total
30-34	15	15	30	30-34	13	14	27
35-39	15	15	30	35-39	14	16	30
40-44	15	15	30	40-44	13	14	27
45-49	15	15	30	45-49	11	15	26
50-54	15	15	30	50-54	14	16	30
55-59	15	15	30	55-59	14	15	29
60+	15	15	30	60+	13	17	30
Total	105	105	210	Total	92	107	199

Chapter V. Response rates, response bias and selection bias

This is the first of the results chapters. In this section a description of the participants at each stage of the study is presented. It goes onto describe the characteristics of non-responders compared with responders. Being a study of migrants, particular attention is devoted to reasons for and impact of migration in the population studied. This is done in order to address the issue of selection bias amongst migrants.

This section also presents the unadjusted socioeconomic indicators measured in this study as a preamble to subsequent results chapters, where more detailed statistical analyses will be presented. The basic descriptive presentation of outcomes of interest is presented in each result chapter, as these need to be calculated adjusted for age and sex.

5.1. Response rate

The last section of the previous Chapter showed that 98.9% of target sample size was achieved, thus a great achievement in enrolment of participants into the study.

Following recent recommendations for good practice of reporting in observational studies, this section reports numbers of individuals at each stage of study and sample attrition [169].

Figures IV-1, IV-2 and IV-3 present the total number of individuals in the primary sampling unit, the number of individuals randomly selected to take part in the study, the number of individuals attempted to contact or potentially eligible as well as those who were contacted, enrolled and completed the study.

5.1.1. Definitions used at each stage of study

The definitions for each category used in each stage of the study were as follow:

- Randomly selected: a randomly selected individual chosen from the updated census specific for each study site.
- Attempted to contact: as above excluding the dead, those who no longer live permanently in the area, pregnant women, any condition that impairs communication such as alcoholism or schizophrenia, or any other reason related to the fulfilment of the study group (previously defined as rural, migrant, urban).
- Contacted: individuals who had a contact with the research team.
- Enrolled: individuals who agreed to take part in the study and signed a consent form. A maximum of three home visits were made to arrange dates for completion of questionnaires and attendance to research clinic.
- Completed study: participants with completed questionnaires, clinical measurements and laboratory analyses. In this case completed questionnaires refer to those people who provided some information to all sections of the main

questionnaire, i.e. socioeconomic, migration and risk factors survey. Completed clinical measurements and completed laboratory analysis were considered as those individuals who, after an appointment made, went through the clinical measurements and blood sampling processes. Only data from individuals who completed the study, that it all three stages, were used in the analysis. It is expected that missing data will exists in certain specific questions, measurements or laboratory tests for various reasons, but this did not disqualify the individual's information to be used on the data analysis process.

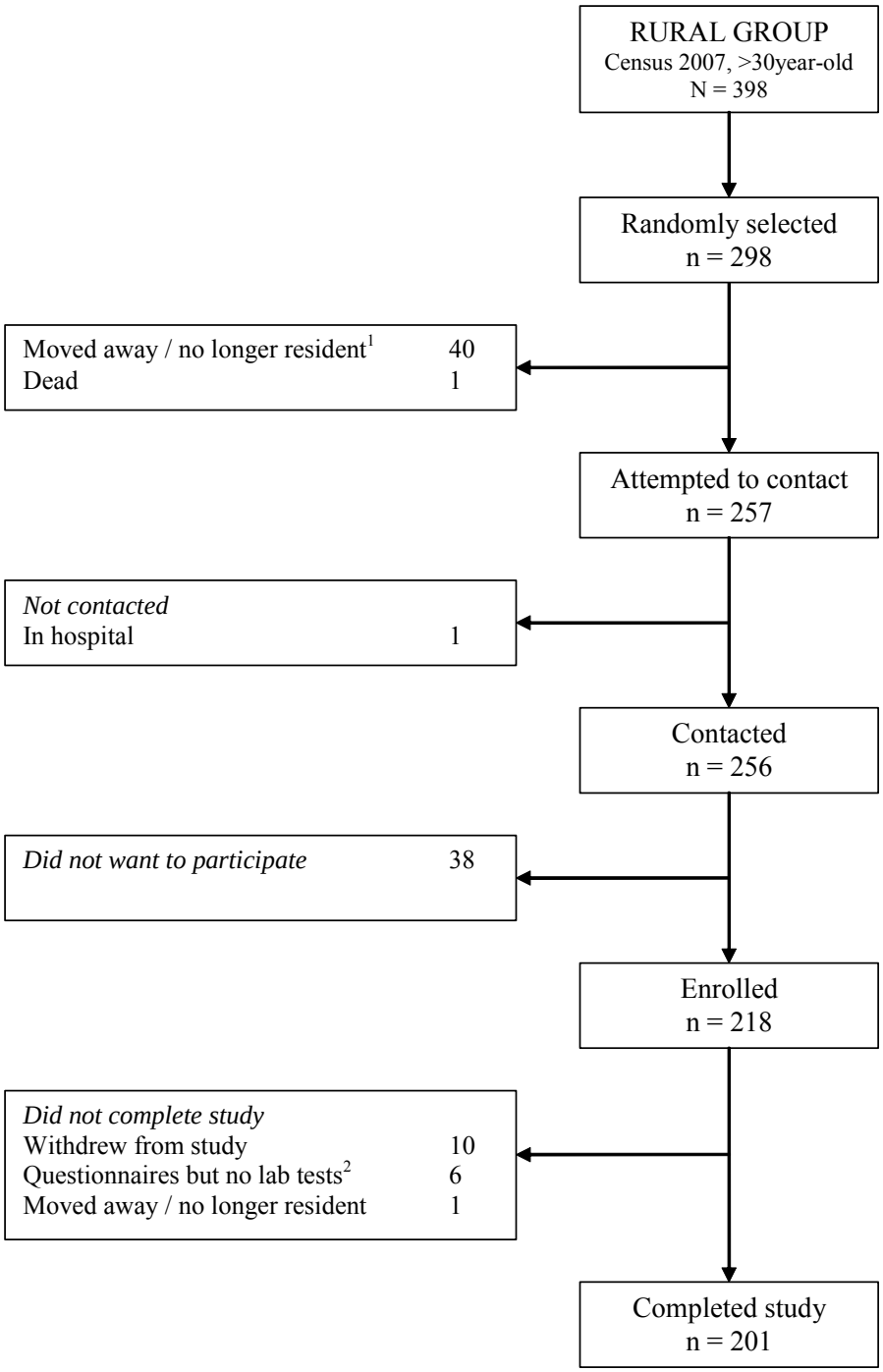
- Non-responder: a non-responder is anyone for whom an attempt was made to contact but did not participate in the study. Non-responders could be identified at two stages of the study. First, all of those that were attempted to contact but were not contacted. Second, all of those contacted but refused to take part in the study. An additional third category could be defined as all of those who agreed to take part and were enrolled in the study but did not complete the study.

5.1.2. Definitions of response rate

Response rate is calculated out of the number of eligible respondents successfully included in the study, as a percentage of the total eligible study population [213, pp 264]. Based on the definitions outlined above, two different response rates were considered. First, response rate at enrolment (Table V-1, page 111), calculated as individuals enrolled in the study as a proportion of individuals attempted to be contacted. Second, response rate at completion of the study, considering number of individuals who completed the study —all of them included in the analysis— as the numerator.

The rationale for presenting these two responses rates are that all those enrolled were eligible and could have potentially completed the study after their agreement to take part on it. For various reasons, presented in the flowcharts, they were unable to do so.

Figure V-1. Study participants' flowchart, rural group



[1] Moved away / no longer resident was defined as those people who no longer live in the given address and moved to another area or are continuously living outside the area of study (e.g. house maids and security guards working and living full-time on employer's houses/properties). This definition applies for all study groups.
[2] Questionnaires completed, but no laboratory tests were done because required sample size was reached and limited funds.

Figure V-2. Study participants' flowchart, rural-to-urban migrant group

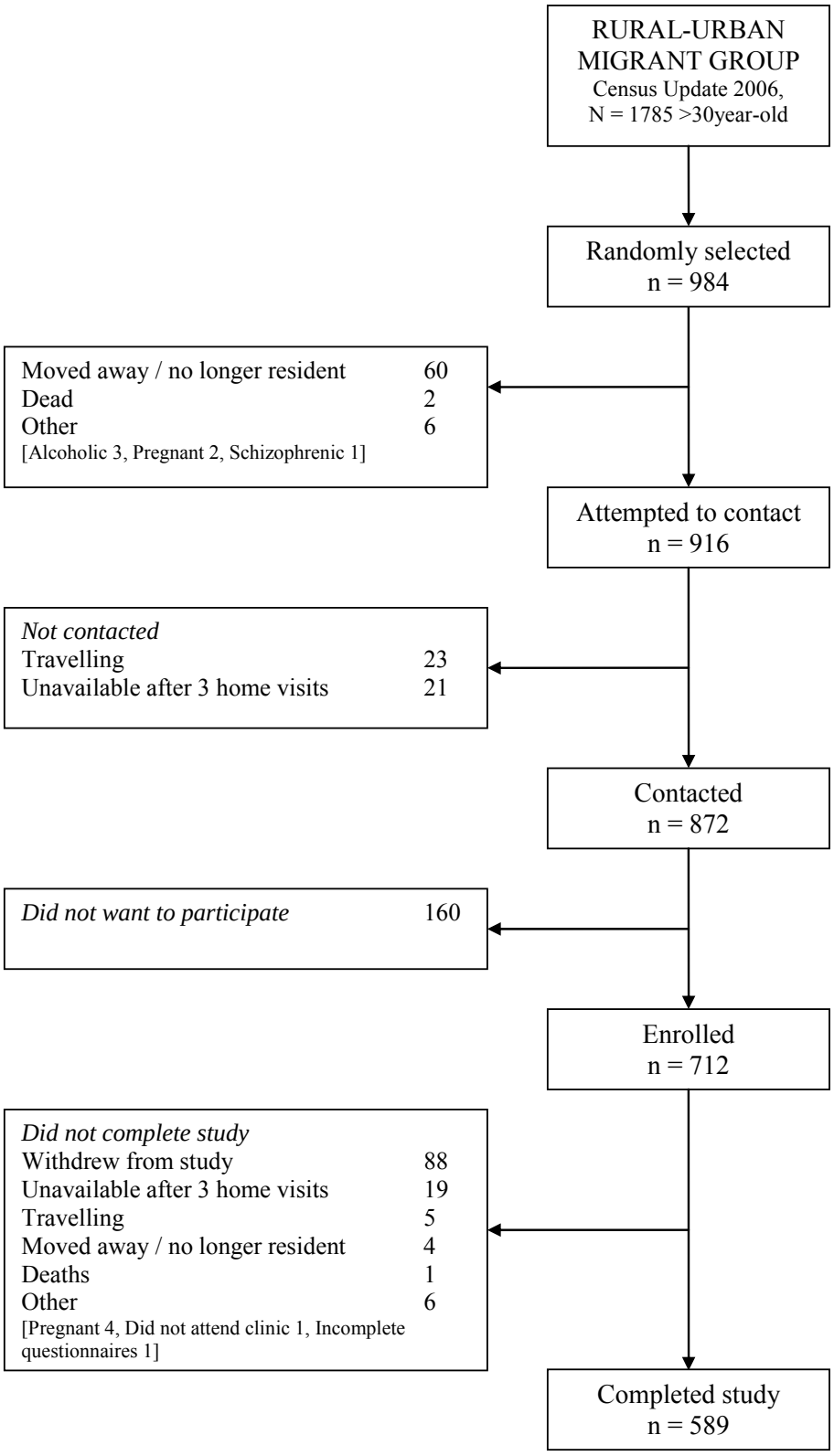


Figure V-3. Study participants' flowchart, urban group

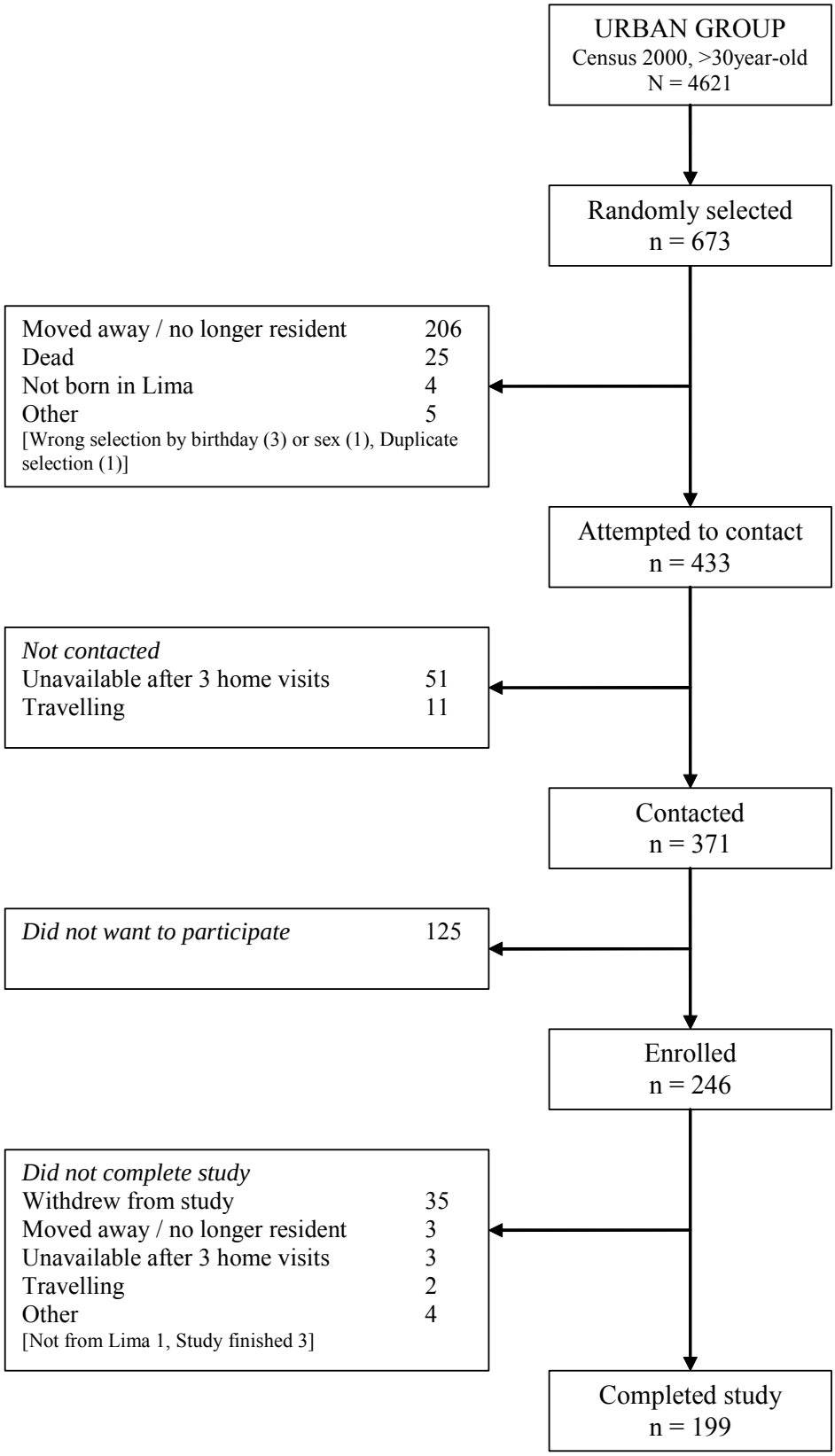


Table V-1. Response rates in study groups

	Response rate at enrolment*		Response rate at completion of study**		Difference
	n	%	n	%	%
Rural	218 / 257	84.8%	201 / 257	78.2%	6.6%
Migrant	712 / 916	77.7%	589 / 916	64.3%	13.4%
Urban	246 / 433	56.8%	199 / 433	46.0%	10.9%
Total	1176 / 1606	73.2%	989 / 1606	61.6%	11.6%

Notes:
* Response rate at enrolment = enrolled study / attempted to contact
** Response rate at completion of study = completed study / attempted to contact

5.1.3. Remarks in relation to response rate

The information from previous pages shows a detailed breakdown of study participants in specific flowcharts for each study group and a summary table. Overall response rate at enrolment was 73.2% and overall response rate at completion of the study was 61.6%.

There were a total of 187 individuals who were enrolled in the study but did not complete it, derived from a total of 1176 enrolled minus 989 who completed the study (16%). The most common reasons for not completing the study were that 71.1% (133/187) withdrew from the study after initial consent and 11.8% (22/187) were unavailable after 3 home visits. Some of the explanations provided for withdrawing from the study (133/187 people) was that some participants did not agree to the blood sampling component of the study which required the extraction of venous blood in 5 tubes (quoting some of them "you are taking too much blood from me").

Response rates were higher in the rural group and lower in the urban group. Attempts to record demographic and socioeconomic information on the non-responders were made and are presented in the next section of this Chapter.

The response rates observed reflect the difficult fieldwork conditions in a poor periurban setting in a developing country. The whole fieldwork was based on household visits, instead of telephone or postal communications as used in other studies. The household strategy was opted as the most effective strategy to increase response rates. The other fieldwork options, contact via telephone or post, were considered not feasible due the lack of telephone or the availability of telephone number on which to contact people, and mail services in this area are of poor quality. If enrolment in the study were to be made by telephone or mail communications, lower response rates would have been obtained.

As expressed before, some participants consented to take part in the study and potentially could have completed it. Various limitations occurred in the period between time of consent given to arrangement or re-schedule of appointments for

examination and interviews. This affected the response rate at completion of the study, being lower that the response rate based on enrolled participants.

Under ideal conditions, if enough fieldwork resources were available permitting a larger fieldwork base, all those who consented to take part on the study could have completed the study in a shorter time frame. This observation is based on the 6% difference between response rates at enrolment versus response rate at completion of the study in the rural area, where fieldwork was quicker (target sample size was 200). On the contrary, in the urban site, the targets were higher (total target sample size in the urban area was 800) and its difference between response rates, at enrolment versus study completion, are double than in the rural ones.

It is worth noticing that the response rates observed in the present study working in poor urban and rural communities are not so different from those of larger studies conducted in developed countries. To name a few, the UK National Women's Health Study had a response rate of 49% for Stage 1 (a total of 26,050 questionnaires were returned in this stage) and 73% for the more targeted Stage 2 [207]. The British Regional Heart Study 1975–2004 response rate was 78% [214-216]. The British Women's Heart and Health Study had a 59.8% response rate (a total of 4286 women of the 7173 invited) [217, 218]. In the US, the Atherosclerosis Risk in Communities (ARIC) Study' response rate ranged from 46% to 67% in the communities studied [219]. Therefore, the response rates observed in this study were within or above the range of response rates of internationally recognised well-conducted observational studies.

5.2. Response bias: survey non-response

Elliot, from the UK's Office of Population Censuses and Surveys, clearly stated that "Non-response is a pervasive fact of life in any social survey research… The problems of non-response become particularly acute when the survey is required to estimate population totals since in this case some explicit assumption must be made about the nature of non-response… However, even in situations where the focus of the survey is the estimation of means, distributions and relationships among the measured variates, ignoring the effects of non-response amounts to making an often unacknowledged assumption that non-respondents form a random sub-sample of the full sample" [220]. These concerns are relevant to the present study.

The literature on population survey research recommends post-stratification weighting for non-response as an strategy to overcome response bias in the calculation of population estimates [220, 221]. In order to apply such correction methods effectively, information on all type of non-responders —not-contacted and refusals— is required. In most cases, information from non-responders is obtained from those who were contacted but refused to participate as long as this group is larger than those non-responders who were not contacted. This information, ideally based on known strata such as age group or sex distribution, should be derived from the whole "true" population or census. Also, a valid alternative, in the case of special population surveys, would be to use estimates of the population distribution made from some other (larger) source or study [220]. Such migrant-specific information was not available from Peru's census, and no larger migration studies have even been conducted in Peru, thus limiting potential efforts to calculate the data taking into account non-response rates. In addition to this, this study does not intend to provide national estimates for all Peruvians or all of Peru's migrant population but only for specific groups as defined in this study. Herein, weighting for non-response was not included in the estimations made in this study.

This section intends to address the issue of response bias potentially introduced by non-responders. The issue of selection bias by migration status is discussed in the

next section. Following lessons learnt from the pilot project, it was decided that a rejection form (shown in Appendix D) would only be used amongst those who were located and refused to participate. As such, this section considers only the non-responders up to the point of enrolment.

5.2.1. Summary of non-responders

A definition for non-responder was described in the previous section. Briefly, a non-responder is anyone who was attempted to contact but did not take part in the study. As such this would include all of those attempted to contact but were not contacted and all of those contacted but refused to participate in the study. This can be summarised as all those who were attempted to be contacted but were not enrolled in the study. (Also, as discussed in the previous section, this could be expanded to include as non-responders those who were enrolled but did not complete the study, already discussed in the previous section). In numbers, non-responders were calculated as total attempted to contact minus total enrolled.

There were a total of 430 non-responders in the study (26.8% of 1606 attempted to contact). 75% (323/430) of non-responders were because of their refusal to take part in the study. The remaining 25% of non-responders were because they were not contactable (107/430 in total, including 72 unavailable after three home visits, 34 on long-term travelling, 1 in hospital). In the urban study site there were a bigger proportion of individuals who were not contacted due to travelling or unavailability after three home visits. A detailed presentation of non-responders is presented in Table V-2.

Table V-2. Summary of non-responders at each stage of the study

	Rural	Migrant	Urban	Total
Attempted to contact	257	916	433	1606
Enrolled in the study	218	712	246	1176
Completed the study	201	589	199	989
Total non-responders*	39	204	187	430
% non-responders	39 / 257 15.2%	204 / 916 22.3%	187 / 433 43.2%	430 / 1606 26.8%
Non-responders: Not contacted				
Travelling	0	23	11	34
Unavailable after 3 home visits	0	21	51	72
In hospital	1	0	0	1
Proportion of not-contacted	1 / 39 (2.6%)	44 / 204 (21.6%)	62 / 187 (33.2%)	107 / 430 (24.9%)
Non-responders: Refused**				
Refused to take part in the study	38	160	125	323
Proportion of refused	38 / 39 (97.4%)	160 / 204 (78.4%)	125 / 187 (66.8%)	323 / 430 (75.1%)

* Non-responders = attempted to contact minus enrolled

** A short rejection form was used in this group, non-responders who refused to take part in the study.

5.2.2. Rejection form

A short rejection form survey was applied to those who refused to take part in the study. This form gathered information on age, sex, reasons for refusing to take part in the study, educational level attained, number of individuals living in the household, smoking status, previous diagnosis of hypertension and diabetes, current treatment for hypertension, age at migration and reasons for migration (the actual form used is provided in Appendix D).

This form was obtained in 282 people of the total 323 who refused to take part in the study (87.3%). Male non-responders were more likely to complete a rejection form (167/282, 59.2%). Among the different exposure groups, the proportions of non-responders providing a rejection form were: rural 31/282 (11%) migrants 121/282 (42.9%) and urban 130/282 (46.1%). Efforts were made to gather as much information as possible in these forms.

Information analysed and presented in this section was based on the 282 who provided a rejection form. A breakdown, by age and sex categories and study group are presented in Figure V-4 and Figure V-5. A high proportion of refusals were observed in males in the urban group. In terms of age, most refusals were observed in the oldest age-group (>60 years old) in all study groups. The potential bias that these age and sex differences could have exerted in the main study are controlled, as the final population studied included similar proportions of sex and age strata.

Table V-3 in page 121 show the different reasons provided for refusal to take part in the study. This was an open-ended question and responses were aggregated into four main categories. The most common reason for refusal was accessibility to social security insurance or access to medical check-ups in the public health system if necessary, thus not needing to take part in the study to benefit from the free evaluation provided. The next category for refusal was unwillingness to take part in the study, followed by logistical circumstances and finally health status.

Figure V-4. Distribution of refusals by gender in each study group among the 282/323 non-responders who completed a rejection form

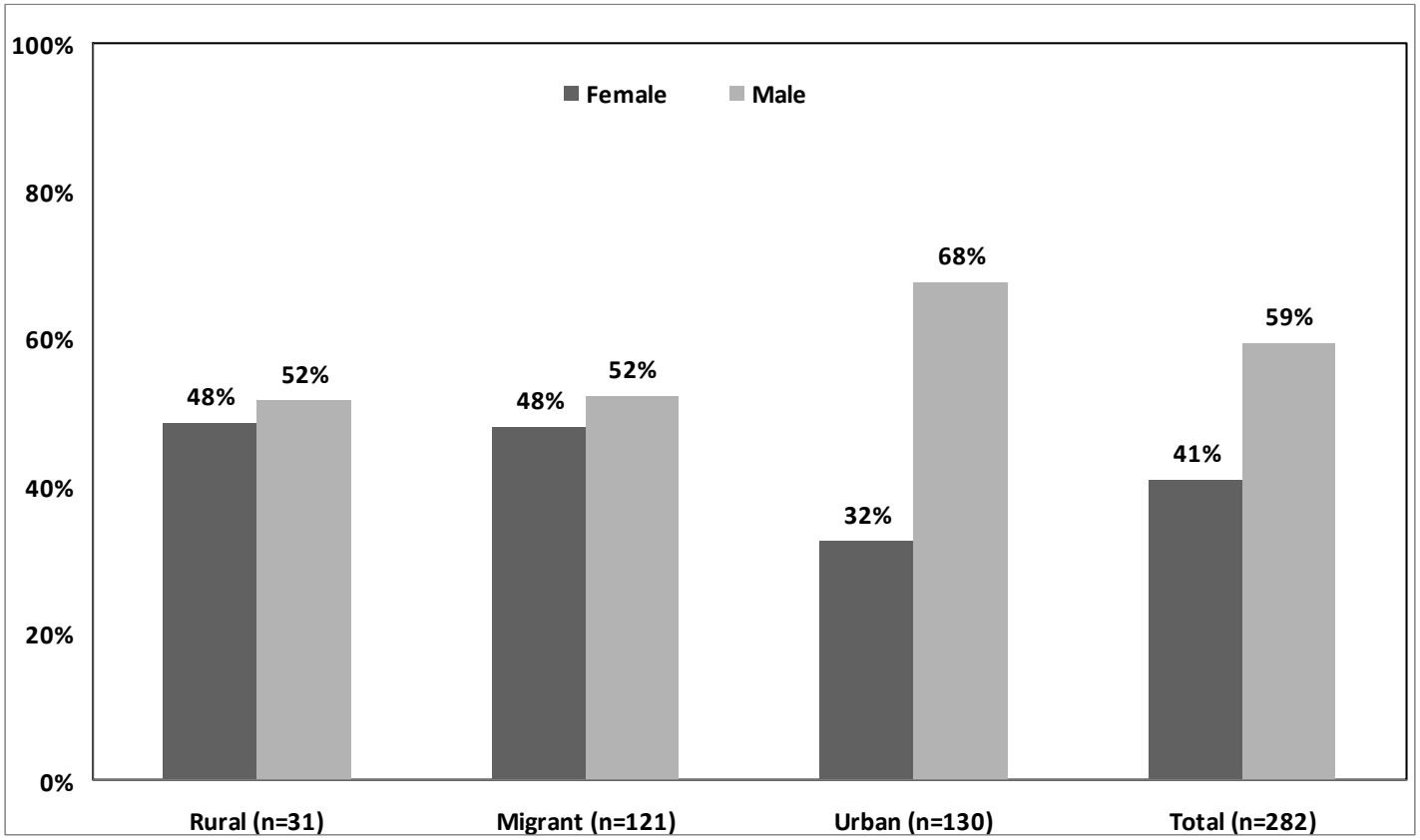


Figure V-5. Distribution of refusals by age in each study group among the 282/323 non-responders who completed a rejection form

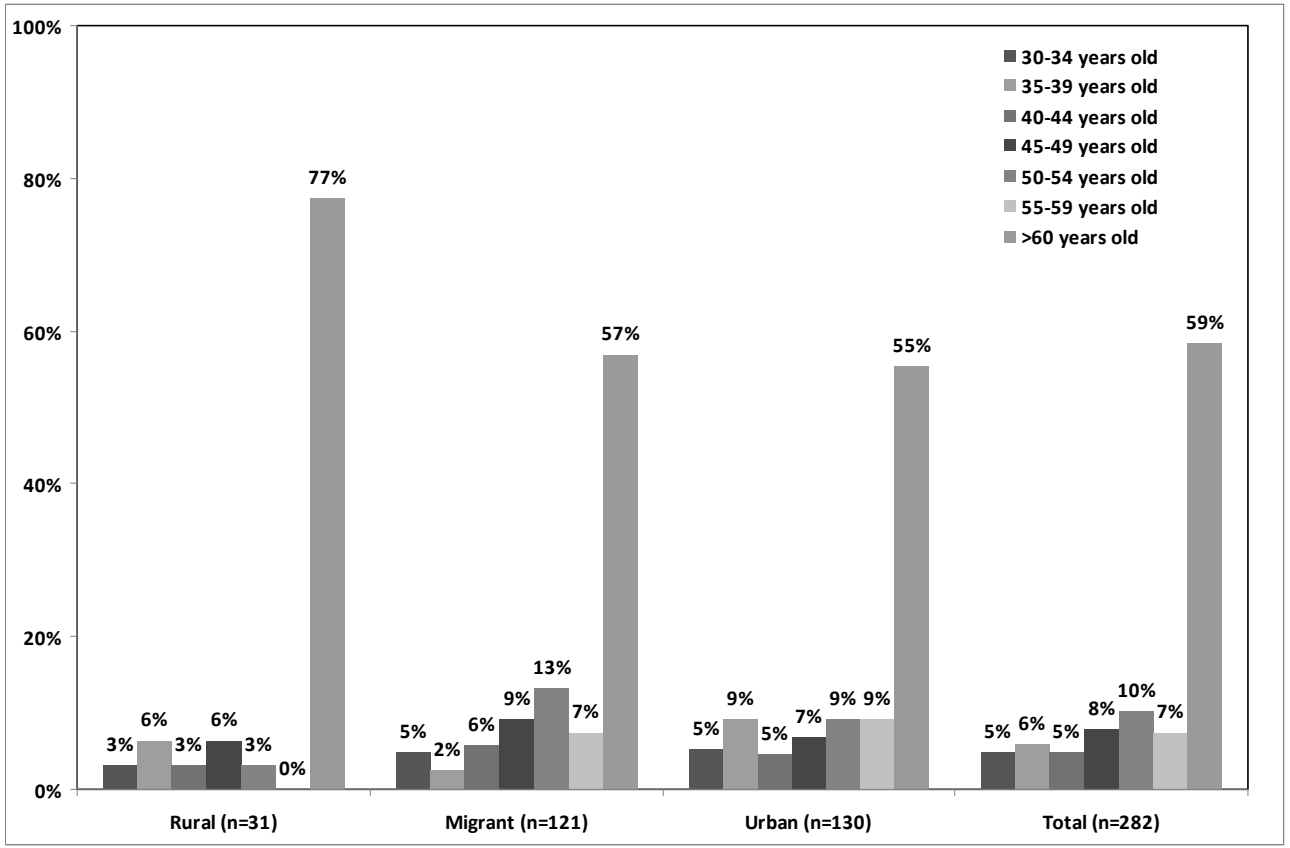


Table V-3. Reasons for refusing participation in the study

	Rural	Migrant	Urban	Total	Proportion*
	n	n	n	n	%
Access to health care					
Total	0	46	46	92	32.6%
Have social security insurance	0	26	33	59	20.9%
Have access to medical check-ups if wanted/needed	0	20	13	33	11.7%
Unwillingness to participate					
Total	0	43	24	67	23.8%
Distrust (e.g. disclosure of personal details, signature)	0	13	12	25	8.9%
"Do not want to participate" statement	0	11	7	18	6.4%
Did not want blood samples to be taken	0	12	4	16	5.7%
Negativism from relative of selected participant	0	7	1	8	2.8%

Notes: Reasons for refusal shown in this table are not mutually exclusive and were aggregated using four categories. The same individual could have answer more than one option.

***** This proportion is calculated using the total of responses as numerator and the total of non-responders (n = 282) as denominator

Table V-3. (continued)

	Rural	Migrant	Urban	Total	Proportion
	n	n	n	n	%
Logistical					
Total	1	17	32	50	17.7%
Time constraints due to working	0	6	12	18	6.4%
Time constraints, unspecified	0	5	13	18	6.4%
Travelling	0	4	1	5	1.8%
Other**	1	2	6	9	3.2%
Health status					
Total	0	16	14	30	10.6%
Recent laboratory test done, no results provided	0	9	5	14	5.0%
Sickness, non-CVD	0	2	5	7	2.5%
Recent laboratory test done, negative for CVD risk factors	0	1	3	4	1.4%
Sickness, unspecified	0	3	0	3	1.1%
Sickness, CVD***	0	1	1	2	0.7%

** Other category included the following reasons: 4 religious beliefs or ideologies, 3 fieldwork mistakes with appointments, 1 pregnant, 1 do not live permanently in the area

*** Sickness CVD included 1 case with diagnosed diabetes and 1 case with diagnosed hypercholesterolaemia. In both cases treatment status was not specified

5.2.3. Comparison between responders and non-responders

The short rejection form, albeit not available for all non-responders, allowed the evaluation of comparable information between those who refused and those who completed the study.

Table V-4 shows the basic characteristics of non-responders. For comparison purposes these have been aggregated into proxy indicators of socioeconomic status, cardiovascular risk factors and indicator of migration process. In the same vein, results are presented disaggregated by study group.

No major differences were observed between rural non-responders when compared to their counterparts who completed the study, but numbers of non-responders were small in the rural group.

However, amongst the urban group, non-responders differed from those who completed the study in education level. More urban non-responders had completed secondary level education (70.3% vs. 56.6% in urban responders). No differences in self reported diagnosis of diabetes or hypertension were seen between response groups.

In relation to migration indicators, non-responders migrant's median age at first migration was similar compared to responders. Both, individual socioeconomic reasons —studies or working reasons— and terrorism were listed amongst the two main reasons for migration in both responders and non-responders.

Table V-4. Characteristics of responders versus non-responders

	Non-responders			Completed study (responders)		
	Rural	**Migrant**	**Urban**	**Rural**	**Migrant**	**Urban**
Socioeconomic						
Number of people living in the same household	n = 6	n = 66	n = 72	n = 200	n = 589	n = 199
Median (IQR)	3.5 (3 - 4)	6 (4 - 7)	5 (3.5 – 6.5)	5 (4 – 7)	5 (4 – 7)	5 (4 – 7)
Education level attained (n, %)	n = 6	n = 67	n = 74	n = 201	n = 588	n = 198
None	0	9 (13.4%)	2 (2.7%)	68 (33.8%)	59 (10%)	2 (1%)
Primary incomplete	2 (33.3%)	18 (26.9%)	3 (4.1%)	64 (31.8%)	124 (21.1%)	11 (5.6%)
Primary complete	3 (50%)	7 (10.5%)	4 (5.4%)	30 (14.9%)	99 (16.8%)	23 (11.6%)
Secondary incomplete	1 (16.7%)	12 (17.9%)	13 (17.6%)	16 (8%)	126 (21.4%)	50 (25.3%)
Secondary complete or more	0	21 (31.3%)	52 (70.3%)	23 (11.4%)	180 (30.6%)	112 (56.6%)

Table V-4. (continued)

	Non-responders			Completed study (responders)		
	Rural	**Migrant**	**Urban**	**Rural**	**Migrant**	**Urban**
Cardiovascular Risk Factors						
Current smoker* (n, %)	1/6 (16.7%)	8/66 (12.1%)	12/75 (16%)	11/201 (5.5%)	59/589 (10%)	40/199 (20.1%)
Diabetes diagnosis, self-report (n, %)**	0/6	0/67	5/75 (6.7%)	0/201	14/589 (2.4%)	9/199 (4.5%)
Hypertension diagnosis, self-report (n, %)**	0/7	7/67 (10.5%)	9/74 (12.2%)	12/201 (6%)	59/589 (10%)	28/199 (14.1%)

Notes:
* Current smoking status in the non-responders was evaluated as a Yes/No question. In the case of the ones who completed the study, current smoker was defined as someone who smoked more than 100 cigarettes in lifetime and last cigarette was less than 6 months ago.
** Diabetes and hypertension correspond to self-report only to enable a similar comparison across groups. These figures will differ from prevalences to be reported later because they do not include diagnosis based on blood pressure or glycaemia measurements.

Table V-4. (continued)

	Non-responders			Completed study (responders)		
	Rural	**Migrant**	**Urban**	**Rural**	**Migrant**	**Urban**
Migration history						
Age when left place of birth		n = 62			n = 572	
Years, Average (±SD)		18.7 (±13.7)			14.7 (±9)	
Years, Median (IQR)		15 (10 – 22)			14 (10 – 17)	
Main reason for migration (n. %)		n = 67			n = 590*	
Socioeconomic, individual		53 (79.1%)			384 (65.1%)	
Terrorism		9 (13.4%)			114 (19.3%)	
Socioeconomic, family		0			5 (0.8%)	
Family/partner		3 (4.5%)			84 (14.2%)	
Other factors		2 (3.0%)			3 (0.3%)	

Notes:
* Main reasons for migration were drawn from information gathered initially in the census. A full description and discussion of those reasons is presented later in Section 0.

5.2.4. Remarks in relation to response bias

Despite the difficulties observed during fieldwork in gathering information about non-responders, the short-rejection form used provided useful data to contextualise the profile of non-responders. Also, it is important to differentiate non-response within groups studied from comparisons between groups.

Some key observations were:

i) 25% (107/430) of non-responders were because of failure to establish contact, while 75% (323/430) of non-responders were due to refusals to participate in the study;

ii) Within-group comparisons showed that, within the urban group, non-responders reported higher levels of education attained (~70% reported secondary complete or more);

iii) In comparisons between groups, non-response rates were highest in the urban group. The two commonest reasons for refusal were having social security insurance or having access to medical check-ups if wanted/needed.

In order to have access to social security health insurance in Peru, either you must be fully employed and this will be deducted from your monthly income, or, if self-employed, the individual has to arrange monthly contributions. Also, to report having access to medical check-ups is necessary also indicates that economic constraints for primary care consultations are not a barrier, considering that health care in Peru's public health sector operates under a strong user-fees policies. Based on these facts, together with the higher education levels of urban non-responders, it could be argued that the urban non-respondents were of a better socioeconomic status in relation to those urban people who did completed the study. A similar profile of higher non-response rates amongst urban, more educated and wealthier men has also been observed in HIV testing in various countries in Africa [222].

The next question to address is how this type of response bias, i.e. by socioeconomic position, might affect the study's results? And if it does, in which direction will it

affect them? In the example cited before, of higher non-response for HIV testing amongst urban better educated male, such non-response was found to have a small, non-significant effects on national HIV seroprevalence estimates obtained from national household surveys [222].

In the case of CVD, for developed societies it has been shown that the lower the socioeconomic position the higher the cardiovascular risk [183, 184]. In contrast, in some developing countries, obesity and other cardiovascular risk factors have been described to be more prevalent amongst upper social classes, although this pattern is changing in some societies [185-187]. At present, the common factor for both rural and urban study sites in this study is poverty [223, 224]. Goldstein et al. have described, using household material possessions as a proxy for socioeconomic status, that prevalence of CVD risk factors is higher amongst lower socioeconomic status groups in a study of six urban areas of Peru [114]. However, Table V-4 shows an inconsistent pattern of CVD risk, particularly when comparing responders and non-responders in the urban group, and therefore extrapolations of the direction of bias cannot be made.

The fact that, in this study, urban non-responders appear to have higher socioeconomic status than urban responders does not imply that they belong to the richest sectors of Lima. As presented before, the urban study site is a shantytown within Lima. In 2003, Peru had a Gini coefficient of 52 (100 equals perfect inequality) [225], indicating a high level of economic inequality. Such inequalities are manifested in the spread of low-quality housing and informal settlements, as in the case of Las Pampas de San Juan de Miraflores, the urban study site. While there will be some degree of socio-economic stratification within these urban settlements hosting urban and migrant individuals, it is also the case that these groups are less well off than the general or 'native' population of Lima [54].

To what extent non-response will exert an impact on CVD risk in this study is less clear given the paucity of evidence from Peru. Predicting the direction of bias in terms of CVD risk remains difficult. A sensible approach is to differentiate non-responders within study groups, and then, separately, how these within-group biases might affect the comparison of interest between groups. Based on the results

presented, it could be that in all three groups, in within-group comparisons, a small tendency for responders to be slightly poorer than non-responders was observed. Therefore, when aggregated, the overall bias in between-group comparisons might be quite small.

It was thus, very useful that a priori, at the design stage as presented in section 2.7.3.1, it was decided that socioeconomic status was a potential source of confounding and relevant detailed information was collected. At the analyses stage, careful consideration will be placed to explore the role of socioeconomic factors in the associations of interest. Additionally, the selection of three well defined study groups —rural, migrants and urban— will enable the exploration of socioeconomic gradients in CVD risk within groups.

5.3. Selection bias: migration

It has been stated that one of the main challenges of migration status is the potential of selection bias being introduced by the very same migrant population. Relevant questions may arise, such as "are migrants a self-selected group?" Such concerns are addressed and discussed in this section, in relation to the population studied.

5.3.1. Ascertainment of exposure by migration status

As stated in Chapter 2, in the section 2.7.2.1 related to non-differential misclassification as a potential source of bias in this study, it was not expected a major degree of misclassification of exposure. The exposure migration has been previously defined as any individual aged 30 years-old or more who reported to have been born in Ayacucho, permanently lives in Las Pampas de San Juan de Miraflores (urban study site) and was randomly selected from the updated census in the area.

The need to confirm the ascertainment of exposure lies in the fact that the migrant population studied has to be comparable to the rural population in the sense that they share the same place of origin as indicative of exposure to the same rural environment. The rural group selected was truly a rural environment and it was selected because of that criteria. As inclusion criteria it was set out that both rural and migrant population should have been born in the department of Ayacucho, Peru. Ideally, all migrants evaluated should have been born in that same village but such endeavour —without wider population censuses or historical records— was not feasible. By design, this was not a before-after migration study. Thus, the potential risk that some migrants may have been born in a non-rural part of Ayacucho, including its capital or smaller towns, was a possibility that would have severely affected the selection of migrants included in the study.

Additional ways to confirm that the final migrant population is similar to the rural one can be achieved by i) individual's categorisation of their place of origin as rural or urban; ii) mother tongue; and, iii) the pattern of migration from Ayacucho into Lima is well established. The following sub-sections address these three scenarios.

5.3.1.1. Confirmation of migrant status by place of origin

All participants in all study groups were asked to describe their place of birth, and type of location, i.e. village, town or city. Table V-5 and Figure V-6, in the next two pages, show the type of place of birth in participants grouped by exposure groups.

Table V-5 shows that nearly 80% in both rural and migrant group reported being born in a rural area. In addition to this, as shown in Figure V-6, the most common description of the type of place for these groups was town and village, which coincide with the rural classification stated before. In the same vein, 81% of urban participants reported being born in an urban area, and nearly 90% described the place where they were born as a city or as the Capital (Lima).

It is normal not to have the same perceptions about their place of origin across all individuals. For example, in some cases urban participants responded to have been born in a rural place. This can be explained by the fact that large cities also host large new settlements or shanty towns where urban poverty is concentrated [48, 54]. As shown in previous figures about the urban site chosen for this study, despite being located in the boundaries of Lima, Peru's capital and largest city, some participants are not wrong in qualifying their place of origin in Lima as a rural one.

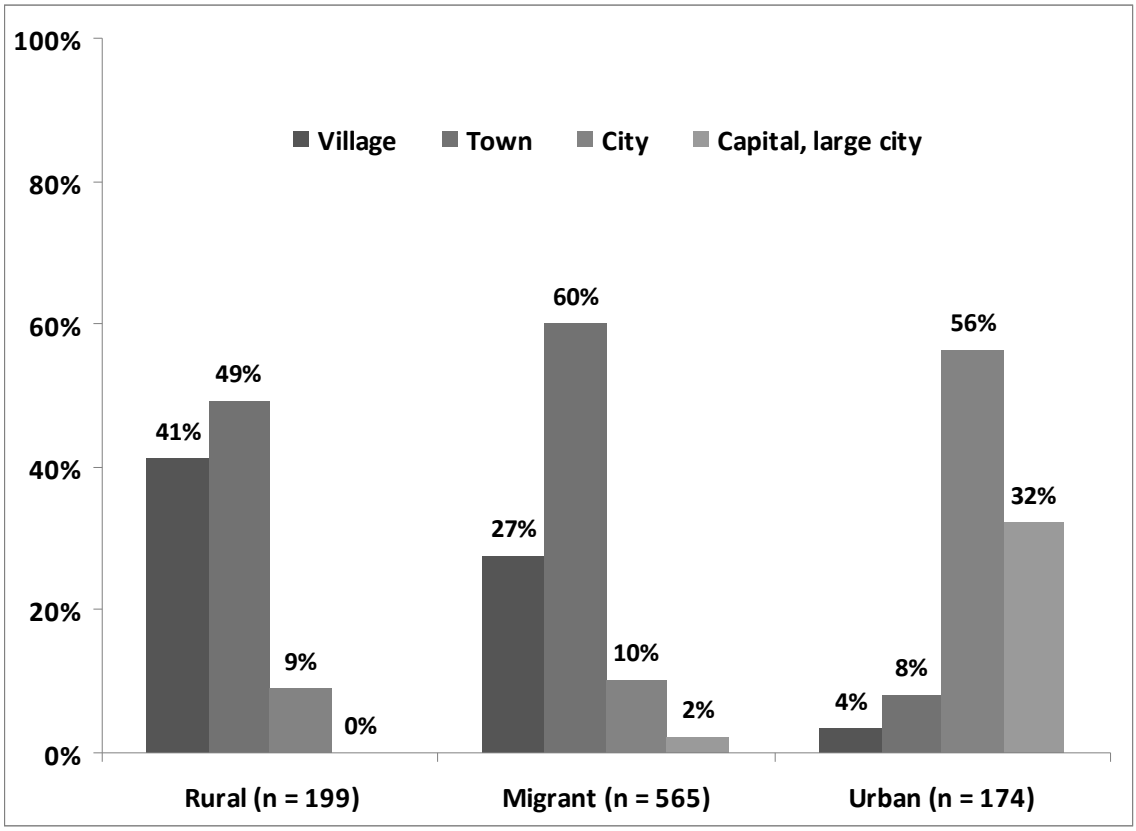
Thus, this information confirms that although difficult to achieve perfection, the current study did very well in the ascertainment of the study groups by place of origin, thus minimising a potential source of selection bias by misclassification of exposure.

Table V-5. Self reported place of birth, urban and rural, by study group

Exposure Groups	Born in an Urban place		Born in a Rural place		Total
	n	%	n	%	n
Rural	44	21.9	157	78.1	201
Migrant	130	22.1	458	77.9	588*
Urban	163	82.3	35	17.7	198*
Total	337	34.1	650	65.9	987

Notes:
* One missing value in this group.

Figure V-6. Self reported place of birth by study group



* Information only available for 938/989 evaluated in total.

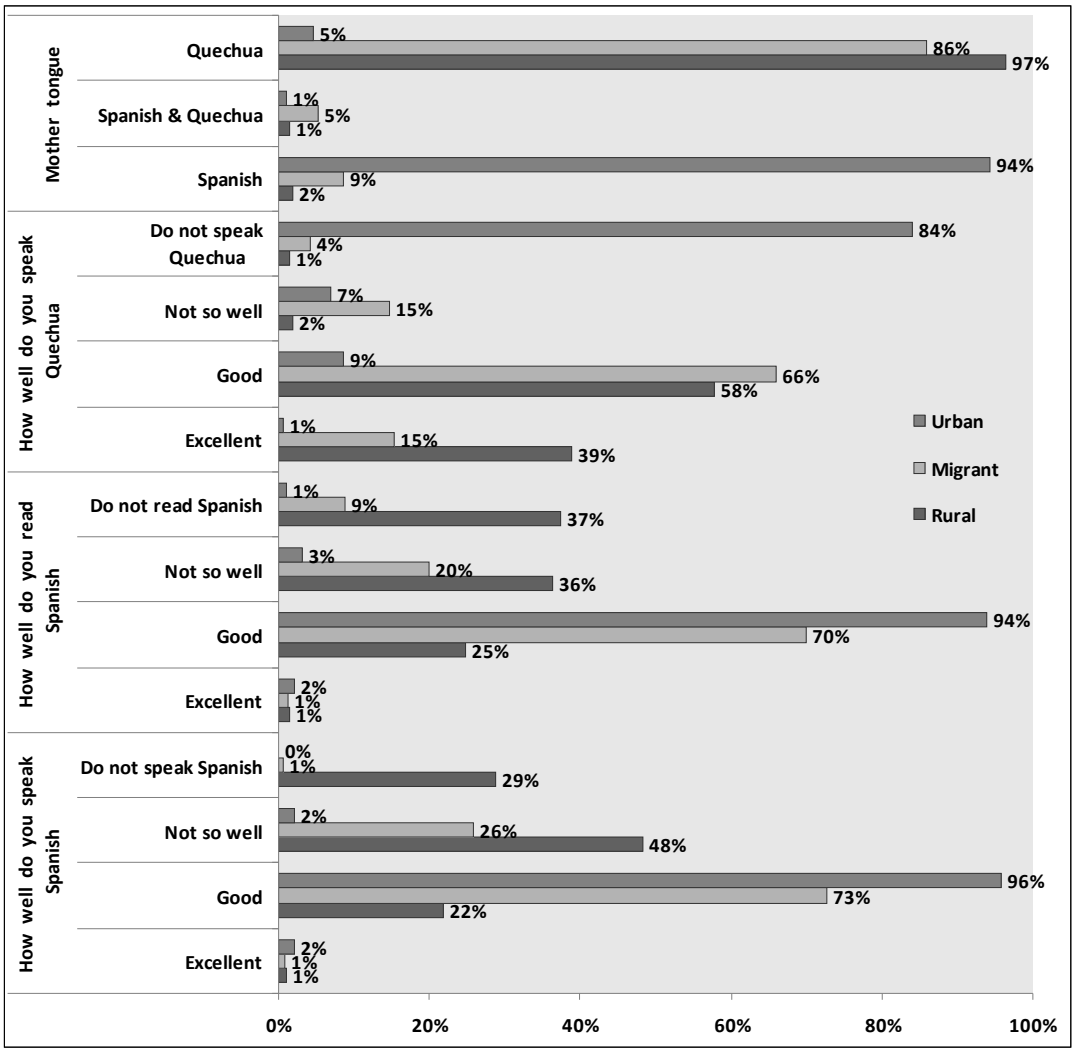
5.3.1.2. Confirmation of migrant status by mother tongue

An additional source of information to evaluate how comparable are migrants to their rural counterparts is mother tongue and self-reported ability of language proficiency. Ayacucho, the place of origin of both migrant and rural population in this study, is largely an indigenous area where most people speak Quechua, an ancient Peruvian language. In Lima, the language most used is Spanish.

All study's participants were asked to name the first language they learnt to speak, and "mother tongue" was obtained from this question. Also, based on their self-perceptions —and not in person-to-person assessments—, they were asked to report how well they speak Quechua, how well they speak Spanish and how well they read Spanish. The question "How well do you read Quechua?" was not asked because Quechua is a language mostly spoken without books or texts to read. The use of Quechua is largely for verbal communications. Responses to these four questions ranged from 971 to 986 out of 989 participants in total (98-99%).

Figure V-7 shows mother tongue and language proficiency in both Quechua and Spanish by study group. Both rural and migrant groups reported Quechua as their mother tongue (over 85%). Approximately 30% of migrants in Lima read not so well or do not read Spanish and similarly, 27% do not speak so well Spanish.

Figure V-7. Mother tongue and language proficiency by study group

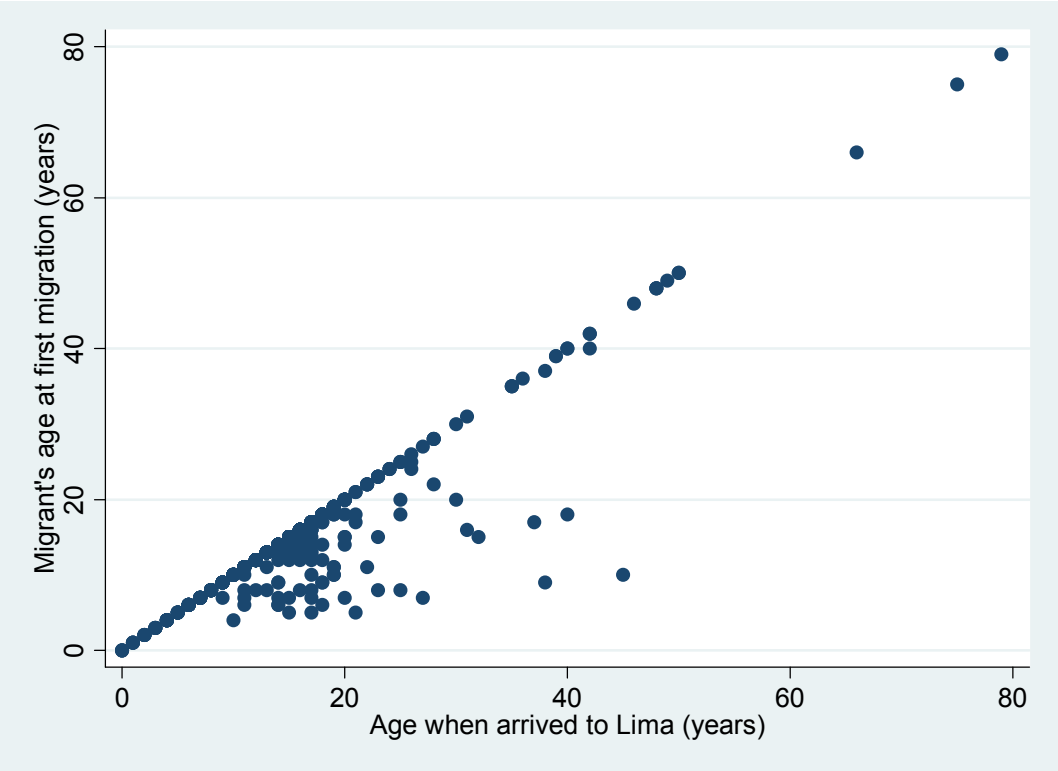


5.3.1.3. Confirmation of Ayacucho-Lima migration pattern

The introduction, Chapter I, section 1.4.2, described how migration from Ayacucho into Lima was a common phenomenon in recent decades, particularly fostered by the years during terrorism and its violence occurred mostly in very rural villages and, at a later period and lesser scale, in urban Ayacucho and the rest of Peru. Also, it is important to confirm such assumption since it could well be the case that most of migrants move to Lima but only after spending considerable time in another rural or urban area. This could affect the ratio of lifetime exposure to a rural environment versus an urban one.

Average (±SD) age at first migration in migrant group that took part in this study was 14.4 (± 8.5 years). Average (±SD) age when arrived into Lima was 15.5 (± 8.8 years). As shown in Figure V-8 these two variables showed a high correlation. This indicates that the pattern of migration in the population studied is in line with the information presented and discussed previously. This is that migrants who were born in Ayacucho tend to move largely directly into Lima.

Figure V-8. Scatter plot of age at first migration versus age when arrived into Lima in migrant group



Correlation coefficient = 0.92, p = <0.0001

5.3.2. Ascertainment of exposure by length of residence in urban environment and age at first migration

Migration in general is not a uniform process —that is not all people migrate at the same time. Therefore, at a given age, people could vary in terms of their length of exposure to urban or rural environments [51, 62]. The primary definition of migration exposure in this study uses a categorical definition of migrants based on place of birth and current place of residency. It is possible, however, to sub-divide migrants according to their length of residence in urban environment —either in absolute numbers of living in an urban area or as a proportion of total age—, and, separately according to the age at first migration. These classifications have been defined a priori as alternative exposure groups for the present study, but it deserves exploration before its analysis.

Detailed information about lifetime exposure to rural and urban environments is important to disentangle its contributions in relation with any outcome of interest. This approximation has been used in other relevant migrant's study in Cameroon [83]. Also, the differentiation of migrants by age at first migration can provide insights on the impact of migration at earlier/younger ages compared to migration later into adulthood.

In this study, migrant participants were asked to report the age when they first left their place of origin as well as the total number of years lived in a rural area and in an urban area. Such self-reporting is prone to recall bias. Participants may not recall the exact age when first migrated, particularly if they migrated as infants. Being migration out-of their rural area of origin a very significant change in their lives, it was expected that the majority will remember or have information about this event. Similarly, the recall of number of years lived in an urban and a rural environment is affected by the own individual's perception of rural and urban places, a subjective description based on each individual's experiences. Despite these potential biases, such information is still a fairly simple data to obtain that permits the evaluation of length of exposure to rural and urban environments and age at first migration.

Information on both number of years lived in an urban area and lifetime exposure to an urban area was available in 95% (559/589) of migrants. Age at first migration was available in 99% (585/589) of the total migrant group.

Figure V-9, page 143, shows a plot between ages at first migration versus number of years living in an urban area. No clear patterns are noted between these two variables, except that a high number of migrants belong to the 20-years-old or less for age-of-first-migration group. This indicates that both, length of residence in an urban area and age at first migration can separately inform to the understanding of migration patterns. Thus the decision made a priori of using them as separate exposures in the analysis stage of this study remains valid. Break down of new exposure groups by length of residence in urban area, lifetime exposure to urban area and age at first migration, following definitions set out in Table II-6 and Table II-7 (pages 49 and 50), are presented in Table V-6, Table V-7 and Table V-8, pages 140 to 142.

In addition to the information presented in this section, amongst the migrant group, lifetime exposure to urban area was calculated as the umber of years lived in an urban area divided over age, expressed as percentage. Prior to this, data was double checked to detect inconsistencies, such as number of years living in an urban area plus number of years living in a rural area did not exceed total age, allowing for ±1 year due to rounding up of months. Information on lifetime exposure on urban areas was available in 559/589 (95%) of migrants, median (IQR) was 69.4% (59.4 – 77.8%), thus indicating that a high proportion of migrant's life has been in an urban environment. Plots of lifetime exposure to urban area against years lived in an urban area and versus age are presented in Figure V-10 and Figure V-11, respectively.

Table V-6. Distribution of migrants by length of residence in urban area

Length of residence in urban area (in years)	n (%)	Mean (SD)*	Median (IQR)*
Migrant <20 years in urban area**	53 (9.5%)	15 (4.1)	16 (4)
Migrant 20-29 years in urban area	203 (36.3%)	25.1 (2.6)	25 (4)
Migrant 30-39 years in urban area	169 (30.2%)	34.4 (2.9)	34 (5)
Migrant ≥40 years in urban area	134 (24%)	46.1 (6.3)	44 (6)
Total***	559 (100%)	32 (10.5)	31 (14)

Notes:
* Summary statistics, mean (SD) and median (IQR), are provided for this variable in relation to each group. All units for descriptive statistics correspond to absolute number of years.
** In order to ensure sufficient numbers in each strata, the first two groups, "<10 years in urban area" (n = 6) and "10-19 years in urban area" (n = 47), were merged into a single stratum "<20 years in urban area".
*** Sub-classification based on 559/589 (95%) observations in the migrant group.

Table V-7. Distribution of migrants by lifetime exposure to urban area

Lifetime exposure to urban area*	n (%)	Range (min – max)	Mean (SD)**	Median (IQR)**
Quartile 1, lowest	141 (25.2%)	0 – 59.5	47.8 (11.6)	51.8 (14.6)
Quartile 2	139 (24.9%)	59.5 – 69.4	64.8 (2.9)	65 (4.9)
Quartile 3	142 (25.4%)	69.6 – 77.8	73.7 (2.4)	73.5 (4)
Quartile 4, highest	137 (24.5%)	78 – 100	85 (5.7)	83.6 (7.7)
Total***	559 (100%)	0 – 100	67.7 (15.1)	69.4 (18.3)

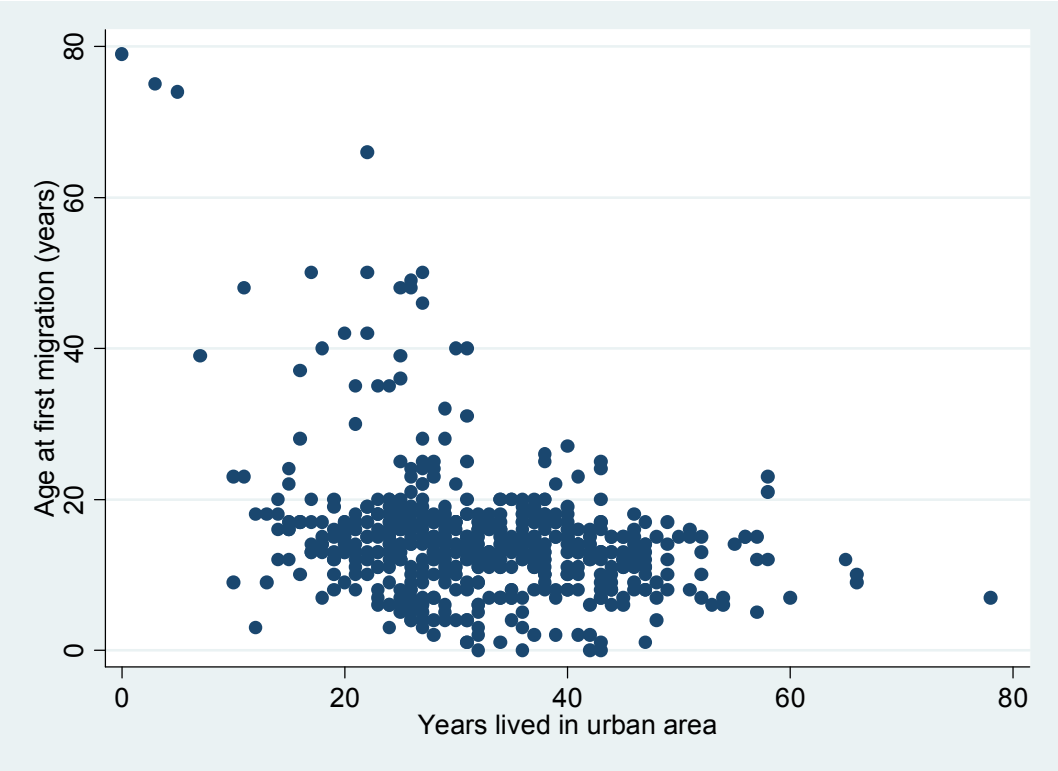
Notes:
* Quartiles were created based on proportion (percentage) of lifetime exposure to urban environment, defined as number of years lived in an urban area divided over age.
** Range and summary statistics, mean (SD) and median (IQR), are provided for this variable in relation to each quartile. All units for descriptive statistics correspond to percentages of total age.
*** Sub-classification based on 559/589 (95%) observations in the migrant group.

Table V-8. Distribution of migrants by age at first migration

Group	n (%)	Mean (SD)*	Median (IQR)*
≤ 12 years old when first migrated	225 (38.4%)	8.2 (3.2)	8 (5)
> 12 years old when first migrated	360 (61.6%)	18.7 (9)	16 (4)
Total**	585 (100%)	14.7 (9)	14 (7)

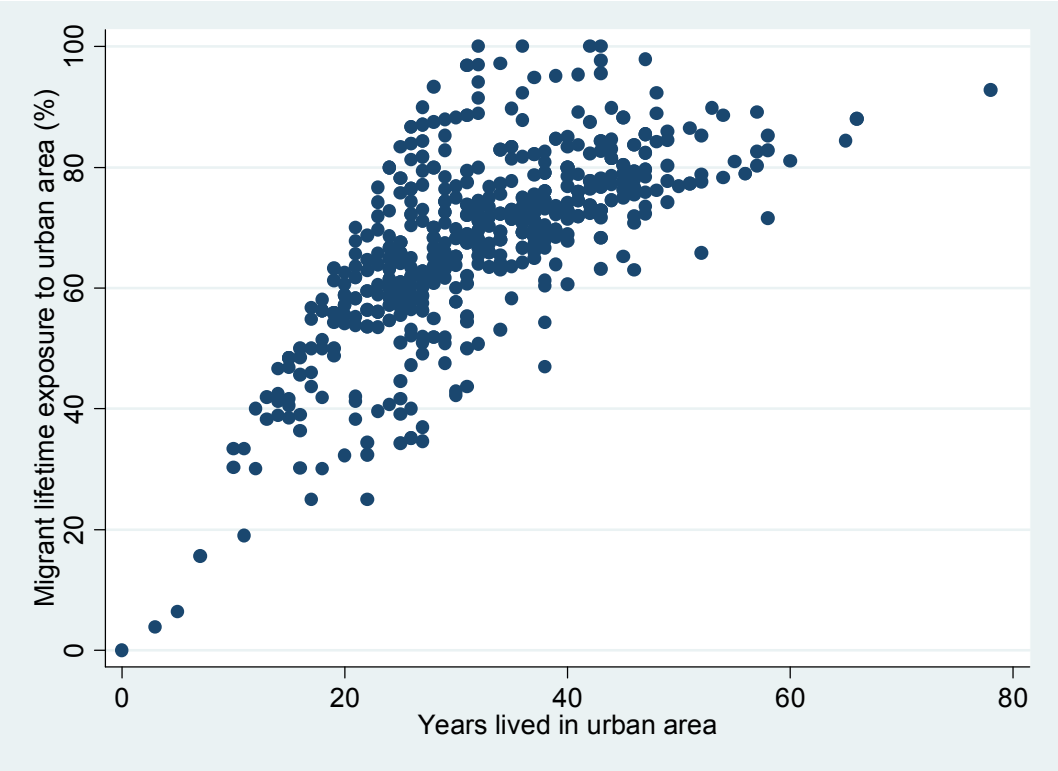
Notes:
* Summary statistics, mean (SD) and median (IQR), are provided for this variable in relation to each group. All units for descriptive statistics correspond to absolute number of years
** Sub-classification based on 585/589 (99.3%) observations in the migrant group

Figure V-9. Scatter plot of age at first migration versus number of years living in an urban area in migrant group



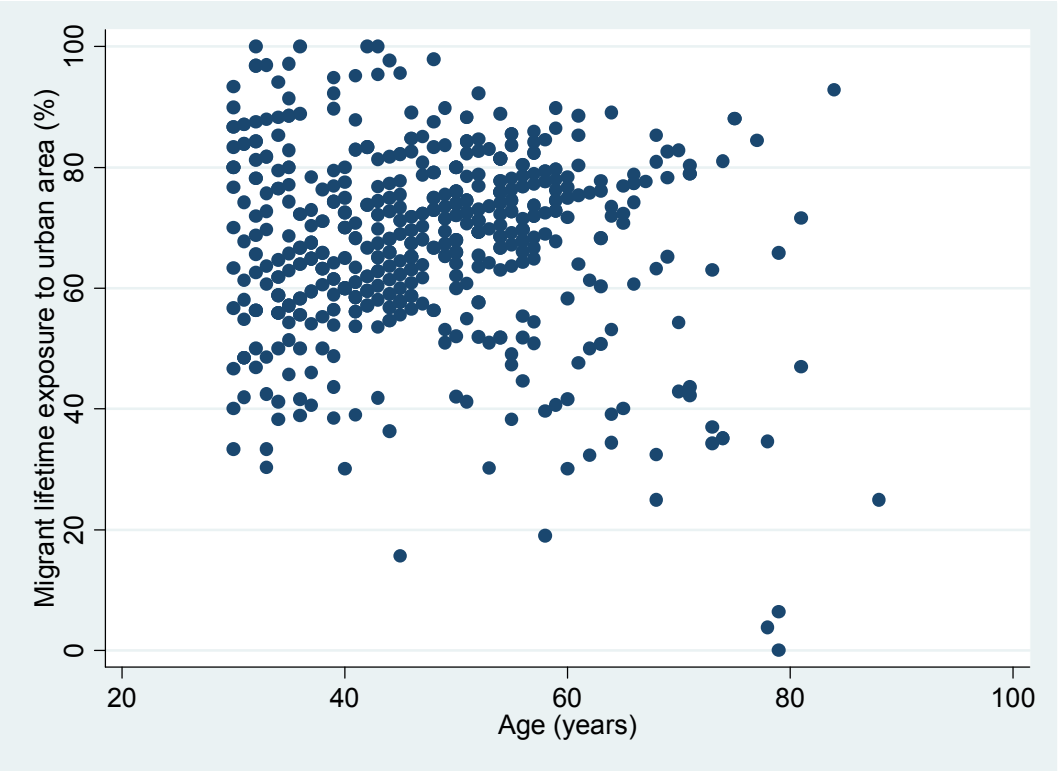
Correlation coefficient = -0.3, p = <0.0001

Figure V-10. Scatter plot of lifetime exposure to urban area versus number of years living in an urban area in migrant group



Note: Four people had 100% lifetime exposure to urban area because they migrated when they were 1 year-old or less. Similarly, one individual had 0% of lifetime exposure to urban area because of moving into the capital within the last year of the study, as an adult.

Figure V-11. Scatter plot of lifetime exposure to urban area versus age in migrant group



Note: Four people had 100% lifetime exposure to urban area because they migrated when they were 1 year-old or less. Similarly, one individual had 0% of lifetime exposure to urban area because of moving into the capital within the last year of the study, as an adult.

5.3.3. Reasons for migration in the population studied

Observations from the pilot study conducted before seem to indicate that migrants were more likely to be of higher socioeconomic status than rural non-migrants, based on level of education attained. This prompted a more detailed assessment of the socioeconomic status of migrants according to reasons for migration. This was done in order to have a better description of reasons for migration in the migrant group, and to provide further understanding about potential selection biases that may still be present with the chosen migrant group.

Violence and conflict are difficult life experiences. The Peruvian experience with years of Shining Path guerrilla activity and terrorism makes it a very sensitive subject to deal with, especially for the migrant group studied in this project. This factor is postulated to be an important contributor for rural to urban migration amongst people from Ayacucho. However, it is very difficult to assess this condition due to the emotional consequences attached to this period. Similarly, some people may prefer not to refer to this issue due to security concerns (since some people were targeted and murdered in other cities for their links with terrorist movements). Additionally, terrorism may not have been the only (main) reason for someone to decide to leave their place of origin.

Taking the opportunity of revisiting all people who reported having been born in Ayacucho and registered to have been living in Pampas de San Juan de Miraflores in Lima at the time of the 2000 Census, it was decided to evaluate the patterns of migration of these people which would provide further insights into the migration patterns of the random sample that will later take part in the main study. Specifically they were asked about all the movements for longer than 6 months starting from their place of origin, their age and the reason that prompted such decision. Participants were free to refer more than one reason for every change of residency through a multiple choice question and an open question. These responses were later aggregated into five main categories: terrorism, individual socioeconomic reasons (e.g. studies or working opportunities of the subject), familial socioeconomic reasons (e.g. parent's work), personal or family-related reasons (e.g. marriage, sickness of a relative) and other reasons. These five categories were analysed separately, first

listing all reasons given (not mutually excluding, more than one reason per subject was considered valid, do not add 100%) and then selecting the main reason given (mutually excluding, only one reason per subject, adds 100%). For the aggregation of responses by main reason given, a priority criterion was set out in the following order: terrorism, socioeconomic individual, socioeconomic family, family reasons, and other reasons.

A priori, three migration-patterns at different migration points were postulated in the migrant group studied. These were evaluated separately because the profile of driving forces —pushing or pulling factors— related to the migration process could differ at different migration points. The three scenarios considered were: i) all migrations, from any place of origin to any destination; ii) first ever out-migration, from place of birth to any destination anywhere and not necessarily Lima; iii) migration into lima, coming from any destination but arriving and settling into Lima. If patterns of migration were to be very unspecific and scattered, each scenario could provide different results or generate a different cadre of migrants. If, as postulated in this study, most migrants from Ayacucho tend to settle into Lima, no major differences would be observed in any of the three scenarios placed above.

Table V-9, page 149, summarise all reasons and the single main reason for migration provided by interviewed participants in the different scenarios outlined for a differential profile in reasons to migrate —those related to any migration, to out-migration from place of birth, and migration into Lima. These figures present data from the migrant sample studied.

Some observations are worth noticing. The profile of reasons for migration in migrants who were born in Ayacucho seems to be a consistent one without major variations neither in the driving forces associated with first migration, arrival into Lima or any migration nor with the aggregation of all or main responses. Interestingly, individual socioeconomic reasons, i.e. migration to pursue work or study opportunities, were referred as the most common reason for migration. This was followed by reasons related to terrorism. This factor, terrorism, was listed as a reason for migration in 20-25% of responses. Thus, albeit not listed as the main reason for migration and acknowledging some degree of underreporting, it can be

concluded that terrorism constituted an important driving factor for migration in the population studied.

Table V-9. Driving reasons for migrating at different migration points in migrant group

	All migrations	First out-migration	Migration into Lima
	n = 589	n = 589	n = 589
All reasons given*			
SE individual	441 (74.7%)	413 (70.0%)	411 (69.7%)
Terrorism	147 (24.9%)	114 (19.3%)	135 (22.9%)
Family/partner	149 (25.3%)	94 (15.9%)	94 (15.9%)
SE family	10 (1.7%)	5 (0.8%)	5 (0.8%)
Other factors	6 (1.0%)	5 (0.8%)	3 (0.3%)
Main reason given*			
SE individual	374 (63.4%)	384 (65.1%)	377 (63.9%)
Terrorism	147 (24.9%)	114 (19.3%)	135 (22.9%)
Family/partner	64 (10.8%)	84 (14.2%)	72 (12.2%)
SE family	4 (0.7%)	5 (0.8%)	5 (0.8%)
Other factors	1 (0.2%)	3 (0.3%)	1 (0.2%)

Notes: SE = socioeconomic

* People interviewed could provide more than one reason for migration ("All reasons given", do not add 100%). Responses were later reclassified into a single "Main reason given" following a pre-established order of priority.

5.3.4. Impact of migration in educational achievements

Are migrants better-off than rural counterparts in terms of educational achievement? This question poses a potential bias, selection bias, inherently to migrant's studies that deserves attention in this study. Both the pilot study and the previous section in this chapter points towards differential characteristics in education level in the migrant population studied compared to the rural group. This section briefly addresses if migration per se had an impact on educational level attained as adulthood. This evaluation was necessary considering that the majority of surveyed participants referred either educational or working reasons as major driving factors to migrate. Such driving factors could introduce a source of selection bias in the population studied, by wrongly assuming that all migrants are "educational" or "working" migrants, which are thus different from their non-migrants counterparts.

In the main study, a specific question on educational status at the point of first migration was asked to the migrant group. This information, available in 484/589 migrants, was disaggregated by adult educational level attained, and is shown in Table V-10. This table shows that about half of individuals (ranging from 38% to 57%) maintained in adulthood the same educational level they had when migrated for the first time. This observation indicates that either they migrated as adults, once education was completed, or that they had to stop their education. The other proportion is spread in the next levels of educations, thus pointing towards progression in educational attainment after migration. Closer examination by age at migration suggests that those who migrated aged 18 years-old or less were more likely to progress in their educational level achievement (less than 40% remained in the lowest level of education at migration, see Table V-11, page 153). The opposite is observed for older migrants, where more than 70% remained in the same education level attained at the point of migration (Table V-12, page 154).

This information is confirmed by a separate observation. In the main study, one specific question asked whether or not such migration had an impact on the individual's educational pathway. Forty-five percent of migrants (261/573 who

answered the question) referred that migration had an impact on their educational pathways because they could continue their studies. 24% responded that migration did not have any impact on their education, while 18% reported having to stop their studies because of the migration process.

In terms of educational achievement, the information presented in this section suggests that, roughly, half of migrants would "advance" in their socioeconomic position —measured by educational attainment at adulthood— following migration, while the other half maintains the same socioeconomic position. However it is difficult to establish firm statements on migrant's educational achievement without having a direct comparison with their rural counterparts to explore whether or not such educational paths are similar or not if they did not migrate.

Table V-10. Education level at first migration disaggregated by education level attained in adulthood in migrant group

		Education level attained in adulthood					
		Level 1	Level 2	Level 3	Level 4	Level 5	Total
		n = 31	n = 108	n = 89	n = 112	n = 144	%
Education level at first migration							
Level 1	n = 60	51.7	13.3	10	6.7	18.3	100
Level 2	n = 233		42.9	18.5	21.9	16.7	100
Level 3	n = 106			37.7	27.4	34.9	100
Level 4	n = 49				57.1	42.9	100
Level 5	n = 36					100	100

Note: All values are row percentages.
Level 1: None, Level 2: Primary incomplete, Level 3: Primary complete, Level 4: Secondary incomplete, Level 5: Secondary complete or more

Table V-11. Education level at first migration disaggregated by education level attained in adulthood amongst participants who migrated at age 18 years-old or less

		Education level attained in adulthood					
		Level 1	Level 2	Level 3	Level 4	Level 5	Total
		n = 19	n = 92	n = 83	n = 101	n = 117	%
Education level at first migration							
Level 1	n = 48	39.6	16.8	12.5	8.3	22.9	100
Level 2	n = 213		39.4	19.7	23	17.8	100
Level 3	n = 99			35.4	27.3	37.4	100
Level 4	n = 40				52.5	47.5	100
Level 5	n = 12					100	100

Note: All values are row percentages.
Level 1: None, Level 2: Primary incomplete, Level 3: Primary complete, Level 4: Secondary incomplete, Level 5: Secondary complete or more

Table V-12. Education level at first migration disaggregated by education level attained in adulthood amongst participants who migrated at age more than 18 years-old

		Education level attained in adulthood					
		Level 1	Level 2	Level 3	Level 4	Level 5	Total
		n = 12	n = 16	n = 6	n = 11	n = 27	%
Education level at first migration							
Level 1	n = 12	100	0	0	0	0	100
Level 2	n = 20		80	5	10	5	100
Level 3	n = 7			71.4	28.6	0	100
Level 4	n = 9				77.8	22.2	100
Level 5	n = 24					100	100

Note: All values are row percentages.
Level 1: None, Level 2: Primary incomplete, Level 3: Primary complete, Level 4: Secondary incomplete, Level 5: Secondary complete or more

5.3.5. Remarks in relation to selection bias

This section addressed the potential of selection bias introduced by exploring whether or not migrants were a self-selected group. Before that, detailed attention was placed to ensure that the definition used for exposure groups were not prone to misclassification. Some key observations were:

i) the ascertainment of migration status, key factors to identify an individual as a migrant, was confirmed by reported place of origin (rural versus urban), by self-reported ability of language proficiency (mother tongue) and an established migration pattern from Ayacucho to Lima;

ii) additional classification of migration status by length of residence in an urban area and age at first migration can aid understanding of migration patterns as these two variable were not completely correlated;

iii) the profile of reasons for migration in the migrant study group does not show major variations in the driving forces associated with either first migration, arrival into Lima or any migration; and, terrorism was an important driving factor for migration in the migrant group studied;

iv) individual socioeconomic reasons, i.e. migration to pursue work or study opportunities, was referred as the most common reason for migration followed by terrorism as an important forcing factor for migration in the population studied;

v) half of migrants did "advance" in their socioeconomic position —measured by educational attainment at adulthood— following migration while the other half maintained the same socioeconomic position

All these observations point towards a case where misclassification of exposure was minimal, but, socioeconomic-driving migration factors cannot be removed from the population study. This is a difficulty inherent in any migration study [58, 59, 62]. The impact of socioeconomic factors can exert some influence on outcomes of interest and is discussed in the next section.

5.4. Socioeconomic indicators

This section presents the various socioeconomic indicators measured in the present study, its aggregation and the rationale for it. It differentiates indicators measured at the individual's current socioeconomic position, e.g. income, overcrowding, educational level attained and assets, from indicators of socioeconomic conditions in childhood, e.g. paternal and maternal education. It continues presenting a rationale for its aggregation into a single socioeconomic indicators as well as the final distribution of newer indicators. The main objective of this approach is to device a simple yet solid strategy that uses most of data for the treatment of socioeconomic indicators as confounders in the association of interest for the present study.

5.4.1. Variables measured

Table V-13 and Table V-14, pages 158 and 159, show the distribution of the current socioeconomic characteristics and socioeconomic conditions in childhood by study group, respectively. Current socioeconomic characteristics were measured through a variety of standard proxies of socioeconomic position including individual's education level and household's characteristics such as income, overcrowding and assets index (the methodology for the calculation of possessions weighted asset index is presented in Table II-1, page 40). Childhood socioeconomic characteristics were evaluated asking for paternal and maternal education attained when the study's participant was aged 12 years-old.

More than 60% of participants from the rural group did not complete primary level education and a gradient towards better attainment in educational level is observed across groups, from rural to migrant to urban people. In terms of household income, figures indicate a high degree of both, rural and urban poverty: 70% of the rural population lives with the equivalent of less than USD $50 dollars per month and nearly three-quarters of migrants and urban population live with a household income in the range of USD $ 50 – 250 dollars. Overcrowding, defined as more than three people per room [143], is much more prevalent in rural than urban areas, affecting

nearly 50% of rural participants. Possessions of assets is Remarkably low in rural group —97% of rural participants fall in the lowest tertile for the assets index— and a gradient towards increased ownership of assets is observed towards the urban group. Women, in both rural and migrant groups, were more likely to have lower attainment of educational level compared against men. None or some primary education was present in 88/106 females (83%) versus 44/95 males (46%) and 137/309 females (44%) versus 46/279 males (17%) in rural and migrant groups, respectively. In terms of household income, women from migrant and urban groups were more likely to have belonged to a family with lower household income compared to men. For example, household's income less than USD $150 dollars per month were reported in 93/289 (32%) of migrant females compared to 58/266 (22%) of migrant males and in 28/103 (27%) of urban females compared to 10/90 (11%) of urban males. In the rural group, such low level of household income was reported in both female and male participants, 70/76 (92%) and 71/82 (87%), respectively.

In terms of parental education, around 80% of participant's mothers in both rural and migrant groups did not have formal education compared to only 20% in the urban group. Although maternal education showed a similar pattern in both rural and migrant groups, this was not the case with paternal education. In the latter, a gradient is observed across groups, similar to the patter of individual's educational level. There were no differences in maternal and paternal educational status by gender across groups, except in the case of migrant's paternal educational status (data not shown). When aggregating both paternal and maternal education into a single variable of highest parental education level, a clear gradient towards increased parental educational level is observed from rural to urban people.

Table V-13. Current individual's socioeconomic indicators by study group

	Rural	Migrant	Urban	Missing data
	n = 201	n = 589	n = 199	n/989
Age, mean (±SD)	48.3 (13.1)	47.8 (11.7)	48.1 (11.9)	–
Individual's education level, n(%)				2 (0.2%)
None	68 (33.8 %)	59 (10 %)	2 (1 %)	
Primary incomplete	64 (31.8 %)	124 (21.1 %)	11 (5.6 %)	
Primary complete	30 (14.9 %)	99 (16.8 %)	23 (11.6 %)	
Secondary incomplete	16 (8 %)	126 (21.4 %)	50 (25.3 %)	
Secondary complete or more	23 (11.4 %)	180 (30.6 %)	112 (56.6 %)	
Household income, n(%)				83 (8.4%)
≤$50 US dollars	109 (69 %)	8 (1.4 %)	2 (1 %)	
$51-150 US dollars	32 (20.3 %)	143 (25.8 %)	36 (18.7 %)	
$151-250 US dollars	10 (6.3 %)	292 (52.6 %)	104 (53.9 %)	
$251-350 US dollars	4 (2.5 %)	82 (14.8 %)	40 (20.7 %)	
$351-450 US dollars	2 (1.3 %)	26 (4.7 %)	8 (4.2 %)	
≥$450 US dollars	1 (0.6 %)	4 (0.7 %)	3 (1.6 %)	
Number of people per household n(%)				6 (0.6%)
<2 people per room	34 (17 %)	217 (37.1 %)	68 (34.3 %)	
2-3 people per room	72 (36 %)	240 (41 %)	75 (37.9 %)	
3-4 people per room	43 (21.5 %)	78 (13.3 %)	37 (18.7 %)	
4 or more people per room	51 (25.5 %)	50 (8.6 %)	18 (9.1 %)	
Possessions weighted asset index				0
Lowest tertile	196 (97.5%)	110 (18.7%)	24 (12.1%)	
Middle	5 (2.5%)	259 (44%)	72 (36.2%)	
Highest tertile	0	220 (37.4%)	103 (51.8%)	

Table V-14. Socioeconomic conditions in childhood by study group

	Rural	Migrant	Urban	Missing data
	n = 201	n = 589	n = 199	n/989
Mother's education level, n(%)				120 (12.1%)
None	149 (85.1 %)	408 (78.8 %)	36 (20.5 %)	
Primary incomplete	21 (12 %)	64 (12.4 %)	31 (17.6 %)	
Primary complete	4 (2.3 %)	26 (5 %)	63 (35.8 %)	
Secondary incomplete	0 (0 %)	6 (1.2 %)	21 (11.9 %)	
Secondary complete or more	1 (0.6 %)	14 (2.7 %)	25 (14.2 %)	
Father's education level, n(%)				240 (24.3%)
None	82 (52.9 %)	173 (38.8 %)	8 (5.4 %)	
Primary incomplete	47 (30.3 %)	143 (32.1 %)	20 (13.5 %)	
Primary complete	21 (13.6 %)	60 (13.5 %)	56 (37.8 %)	
Secondary incomplete	3 (1.9 %)	29 (6.5 %)	17 (11.5 %)	
Secondary complete or more	2 (1.3 %)	41 (9.2 %)	47 (31.8 %)	
Highest parental education level, n(%)				0
None	122 (60.7%)	298 (50.6%)	30 (15.1%)	
Some primary	52 (25.9%)	147 (25%)	26 (13.1%)	
Primary complete or more	27 (13.4%)	144 (24.5%)	143 (71.9%)	

5.4.2. Aggregation of variables into a socioeconomic deprivation index

This section analyses in detail the rationale for the aggregation of individual socioeconomic characteristics into a single variable and its construction. This new variable, a socioeconomic deprivation index, was decided to be the single adulthood socioeconomic indicator to be added in multivariable statistical modelling as a confounder factor. Such procedure was not adopted for parental education, a proxy for childhood socioeconomic characteristics, because its influence in health outcomes occurs at a different point of the life-course. Therefore, parental education was also decided to be kept as a separate variable to control for in further analyses.

5.4.2.1. Rationale

It is well recognised that different socioeconomic indicators —income, wealth, educational attainment and occupational group— are all related to and help explain people's health status and that social circumstances across the life-course influence people's health and well-being [149]. In addition, educational attainment, as an indicator or socioeconomic position, is primarily related to health through the advantages it gives people in their later socioeconomic trajectories, not simply because education encourages healthy behaviours [149].

Following on the evidence from literature, it is evident that the information collected related to socioeconomic characteristics is important to be taken into account in the evaluation of any association of interest. In fact, such information was collected, as decided a priori and confirmed in the section, due to the confounding roles it may exert on the associations of interest of this study.

The pattern of the various adulthood socioeconomic-related variables in this study is clear —the rural group tend to have a consistent pattern of lower socioeconomic position, either by showing lower rates of education attained, lower income, low assets index or higher rates of overcrowding. The opposite is observed for the urban group. Such disparity places considerable challenges on the uses of data for multivariable modelling in this study.

Ideally, as recommended by the literature [139-141, 149], all indicators should be considered separately in the statistical multivariable modelling because of their independent contribution to health outcomes. Such possibility was explored by ruling-out colinearity between different individual socioeconomic variables and results suggest that all of them could be used in statistical models to be built (data not shown).

The main challenge, however, was that, as shown already in Table V-13 and Table V-14, various indicators had very low numbers of cases in the extreme cells, e.g. only one urban individual earns less than USD $50 dollars and only one rural individual reported to earn more than USD $450 dollars. In the same vein, 97% of rural individuals fall in the lowest tertile for possessions weighted asset index. These observations limit the spread of sufficient number of individuals in each cell or category for each variable, an important requirement to be able to run multivariable statistical models. Such concentration of characteristics into certain cells would have resulted in statistical models yielding wide confidence intervals. Even after shortening the aggregation of variables into smaller number of categories, still wide confidence intervals were observed.

The study had rich information on various proxies for socioeconomic position, but because of the patterns of the data, such richness could not be exploited to a maximum as separate variables in the statistical modelling. Thus, in terms of how to manage socioeconomic variables at the statistical modelling stage of the study, two scenarios were considered: either to choose a single indicator or to find out a reasonable way to aggregate all the available information.

It was considered an advantage of the present study to have such richness of data related to socioeconomic position. The option of selecting only one of the socioeconomic proxies would sacrifice most of the data gathered. Such alternative was discarded because, in addition to the “waste” of data collected due to non-usage, none of the indicators has been ascribed as the “best” or “gold standard” for measuring socioeconomic status.

It was thus decided to explore the maximisation of measured variables through the creation of a single proxy for socioeconomic status that could sustain multivariable modelling. There are no clear guidelines for aggregating indicators of socioeconomic position [142, 226]. However, the social sciences have a demonstrated track record of operationalising indicators, particularly for the measurement of poverty through deprivation indexes [143-153] and such have been adopted and recommended by international organisations including UNICEF [227] and the UN sponsored Expert Group on Poverty Statistics [143].

5.4.2.2. Construction of the socioeconomic deprivation index

As shown in the construction of deprivation indexes elsewhere [144, 145, 148, 150], all four individual proxies for socioeconomic status —education, income, assets and overcrowding— were grouped into deprivation categories following the operational definitions set out in Table V-15. These new variables were evaluated through inter-item correlations and Cronbach's alpha, where values of 0.7 are considered appropriate for research purposes [145, 228]. The four deprivation variables showed a Cronbach's alpha of 0.5668, and excluding overcrowding it increased to 0.6040. The later value was considered a reasonable trade-off that would enable the use of three socioeconomic indicators into a single aggregated variable.

Following this assessment, the equally weighted deprivation scores (0, 1) of education, income and assets were summed, with a maximum score of three and a minimum of zero. Higher scores reflect the experience of a larger number of deprivations simultaneously. A cut-off of two or more deprivations was considered the threshold to define socioeconomic deprivation: nearly 90% of the rural participants were socioeconomically deprived compared to 18% and 7% in migrants and urban people, respectively. If overcrowding were to be included in the calculation of the socioeconomic deprivation index, the prevalences of experiencing simultaneously two or more deprivations would remain the same largely because of the threshold level of two or more deprivations (data not shown).

Table V-15. Operational definitions of socioeconomic deprivation

Deprivation	Yes	No
Education	None or incomplete primary education	Primary complete or more
Income	Household income less than USD $150 dollars per month	Household income more than USD $150 dollars per month
Overcrowding	Three or more people per room	Less than three people per room
Assets	Lowest tertile of possessions weighted asset index	Middle and highest tertiles of possessions weighted asset index

Note: The categories presented in Table V-13 were used to create specific deprivation variables for each variable presented in this table

Table V-16. Distribution of specific deprivations and deprivation index by study groups

	Rural	Migrant	Urban	Missing data
	n = 201	n = 589	n = 199	n/989
Deprived by specific category, n(%)*				
Education	132 (65.7%)	183 (31.1%)	13 (6.57%)	0
Income	141 (89.2%)	151 (27.2%)	38 (19.7%)	
Overcrowding	94 (47%)	128 (21.9%)	55 (27.8%)	
Assets	196 (97.5%)	119 (20.2%)	32 (16.1%)	
Number of deprivations per individual, n(%)**				0
None	3 (1.5%)	265 (45%)	131 (65.8%)	
One deprivation	18 (9%)	217 (36.8%)	55 (27.6%)	
Two deprivations	89 (44.23%)	85 (14.4%)	11 (5.5%)	
Three deprivations	91 (45.3%)	22 (3.7%)	2 (1%)	
Socioeconomically deprived, n(%)***				
No	21 (10.5%)	482 (81.8%)	186 (93.5%)	0
Yes	180 (89.5%)	107 (18.2%)	13 (6.5%)	

Notes:
* Categories are based on the operational definitions presented in Table V-15.
** Number of deprivations is based on the sum of individual deprivations (education, income and assets) in the same individual.
*** Socioeconomic deprivation index, as explained in the text, was calculated based on individual deprivations except overcrowding. An individual was considered as socioeconomically deprived if had two or more deprivations.

5.4.3. Remarks in relation to socioeconomic indicators

The various characteristics —proxies for socioeconomic status at different points over time— have been, a priori, considered as potential confounders. A single aggregated variable of socioeconomic deprivation will be used as confounding factor in multivariable analyses in the present study. These observations are intriguing and worth summarising due to the potential impact these characteristics may exert in the associations of interest in the present study.

i) A gradient across groups is observed in terms of individual's educational attainment but not in household income. Participant's households in both rural and urban sites generate very low monthly incomes, a reflection of poverty in both study sites;

ii) Although aggregated household income may be similar in the urban site amongst migrants and urban populations, women are more likely to be worst-off within these groups;

iii) Socioeconomic status addressed by assets ownership or number of people per household is much lower in the rural group compared to their migrant and urban counterparts;

iv) A socioeconomic deprivation index —based on education, income and assets— was used as a simple yet solid strategy that uses most of data for the treatment of current individual socioeconomic indicators as confounders in the association of interest for the present study. This index identified as socioeconomically deprived those individuals that experience simultaneously two or more socioeconomic deprivations. A gradient in socioeconomic deprivation was observed from rural to urban groups;

v) Maternal educational level is very poor and almost identical in both rural and urban participants. On the contrary, paternal educational level and aggregated

highest paternal and maternal education show a gradient towards better education across groups, from rural to urban;

5.5. Summary

- Overall response rate at enrolment was 73.2% (1176/1606) and overall response rate at completion of the study was 61.6% (989 /1606). In both cases, response rate was lowest in the urban group compared to the other study groups;

- Twenty-five percent (107/430) of non-responders were because of failure to establish contact, while 75% (323/430) of non-responders were due to refusals to participate in the study;

- Information about non-responders gathered through a rejection form was obtained in 282/323 (87.3%) who refused to participate in the study. Urban non-responders reported higher levels of education attained (~70% reported secondary complete or more);

- In relation to selection bias, the ascertainment of migration status was confirmed by reported place of origin (rural versus urban), by self-reported ability of language proficiency (mother tongue) and an established migration pattern from Ayacucho to Lima;

- Additional classification of migration status by length of residence in an urban area and age at first migration can separately inform to the understanding of migration patterns as these two variable were not correlated;

- A socioeconomic deprivation index —based on education attained, current income and current assets' possession— identified as socioeconomically deprived as those individuals that experience simultaneously two or more socioeconomic deprivations. A gradient in socioeconomic deprivation was observed from rural to urban study groups, that is, the rural group was more deprived based on current measures of deprivation.

Chapter VI. Behavioural risk factors

This chapter presents the descriptive and multivariable analyses of tobacco and alcohol consumption. Of note, none of the measurements presented in this chapter were primary outcomes. However, for a clearer presentation of results, it was decided to present these secondary outcomes as a separate chapter.

Being this the first chapter that presents results obtained from statistical analyses, a section on how to approach all results chapters, from Chapter VI to Chapter X, is also presented within this chapter.

6.1. Presentation and structure of results chapters

From this section onwards, and for convenience in the presentation of results in a clear manner, individual but related risk factors were aggregated into separate chapters: behavioural, anthropometric, blood pressure and lipids and metabolic risk factors. A final result chapter deals with aggregation of risk factors as a single outcome. Each of these chapters were organised in smaller sections in order to address and answer clearly the main overall and specific research questions of this study.

As stated in section 2.1.4, the overall research question of this study reads as follow: "i) is there a difference in specific CVD risk factors in the rural-to-urban migrant group compared to those who did not migrate?" To address this question, analyses were carried out using the rural group as baseline group. Comparisons were made between migrant-to-rural group and urban-to-rural groups using, in most cases, information from all study participants (n = 989).

The distribution of migrants according to length of residence in urban area —either as absolute number of years or lifetime exposure to urban area— and age at first migration was presented in Table V-6, Table V-7 and Table V-8, respectively. Three separate specific research questions also form part of this research. These explore if the pattern of CVD risk factors amongst migrants vary by: ii) length of residence in urban environment, iii) lifetime exposure to urban environment, and iv) age at first migration. All these three specific research questions used information only from the migrant group. Comparisons were made using as baseline the lowest category of exposure created after the sub-classification of the migrant group. Unless otherwise stated, specific research questions ii) and iii) were evaluated with information available from 559/589 (95%) of migrants, whereas research question iv) was evaluated using information from 585/589 (99%) of migrants. Such losses of information occurred because not all migrants provided information on either number of years living on urban area or age at first migration. Such information was

necessary to proceed to migrant's group sub-classification. Subheadings will also be used to guide the reading of results chapters.

Also, most definitions and approaches to data aggregation were previously set out in Chapter II, Section 2.3.3 (specifically from Table II-1 to Table II-13, pages 40 to 63). Each of those tables provides the definitions used for general variables as well as exposures and outcomes. In the following results chapters conventional units of measurement are reported. Conversion factors to the international System of Units are provided at the beginning of this document, on page xxvii.

In relation to multivariable analyses, all models were elaborated in a step-wise approach, including a priori defined confounder variables in the models. All standardised mean difference (SMD) calculations and figures elaborated were made using the final fully adjusted model controlling for current socioeconomic status and parental education.

Each results chapter ends with a short discussion section that addresses the specific results, its strengths and weaknesses, and an interpretation of results presented. Later on, in a separate section, a discussion chapter is presented which in turn addresses, together as an overview, all results reported in this document.

6.2. Descriptive results

Table VI-1 presents the distribution of the variables relevant to this chapter. Alcohol consumption and volume of alcohol consumption in the same occasion were asked in relation to the last year. However, frequency of hangover was restricted to the last month only (see Table II-9 for a detailed description of definitions used in this section).

All study groups show similar patterns of alcohol consumption and heavy drinking. A clear gradient, however, towards higher prevalences of current smoking status from rural to urban groups was observed: 5% in rural people, 10% in migrants and 20% in urban people.

Table VI-1. Distribution of behavioural risk factor variables by study group

	Rural	Migrant	Urban	missing
	201	589	199	n/989 (%)
Alcohol				
Overall frequency of alcohol consumption				31 (3%)
Never	78 (39.2 %)	103 (18.2 %)	14 (7.3 %)	
One or less per month	114 (57.3 %)	419 (73.9 %)	159 (82.8 %)	
Two or more per month	7 (3.5 %)	45 (7.9 %)	19 (9.9 %)	
Volume (6+ drinks) in the same occasion				31 (3%)
Never	83 (41.9 %)	129 (22.7 %)	25 (13.1 %)	
Less than once per month	92 (46.5 %)	393 (69.1 %)	149 (78 %)	
Once or more per month	23 (11.6 %)	47 (8.3 %)	17 (8.9 %)	
Frequency of hangover in the last month				30 (3%)
Never	116 (58.3 %)	297 (52.3 %)	91 (47.4 %)	
Less than once per month	70 (35.2 %)	234 (41.2 %)	84 (43.8 %)	
Once or more per month	13 (6.5 %)	37 (6.5 %)	17 (8.9 %)	
Heavy drinkers				0
Yes	176 (87.6 %)	541 (91.9 %)	180 (90.5 %)	
No	25 (12.4 %)	48 (8.2 %)	19 (9.6 %)	
Tobacco				
Smoking status				0
Never	187 (93 %)	492 (83.5 %)	134 (67.3 %)	
Former	3 (1.5 %)	38 (6.5 %)	25 (12.6 %)	
Current	11 (5.5 %)	59 (10 %)	40 (20.1 %)	
Cigarette consumption in the last month*				75/110 (32%)*
n = 75	6	37	32	
Mean (IQR)	10 (1 - 20)	5 (3 - 20)	5.5 (1 - 26.5)	

* Cigarette consumption in the last month was only asked to current smokers (n = 75).

6.3. Multivariable analyses

All behavioural risk factor variables were re-grouped into binary variables (Table II-9) and analysed using logistic regressions. In the case of frequency-related alcohol questions, comparisons were made between low vs. high consumption. Heavy drinkers and current smoker status were considered as categorical binary yes/no variables.

6.3.1. Analysis by main exposure: migrant vs. non-migrant groups

Table VI-2 present OR (95% CI) for the associations of rural, migrant and urban groups with behavioural risk factors. No differences were observed between groups with regards to volume of alcohol consumption in the same occasion, frequency of hangover or heavy drinker status.

Compared to the rural group, and after adjustment for age and sex, migrants and urban people were 2.4 (95% CI 1.1 – 5.6) and 3.2 (95% CI 1.3 – 8) times more likely to have a higher overall frequency of alcohol consumption, respectively. However, such associations were slightly attenuated and became non significant after adjustment for socioeconomic indicators.

In relation to current smoking status, and after adjustment for age, sex, current individual's socioeconomic deprivation and parental education status, urban people were 2.5 (95% CI 1 – 6.7) times more likely to be current smokers than rural people. No differences were observed between rural and migrant participants in term of smoking status.

Table VI-2. Multivariable association of rural, migrant and urban groups with behavioural risk factors

Outcome Category	Rural	Migrants	Urban
Alcohol			
Frequency of consumption			
Model 1	1	2.4 (1.1 – 5.6)	3.2 (1.3 – 8)
Model 2	1	2.3 (0.8 – 6.6)	3 (0.9 – 9.8)
Model 3	1	2.4 (0.8 – 6.9)	2.9 (0.9 – 9.7)
Volume of consumption			
Model 1	1	0.7 (0.4 – 1.1)	0.7 (0.4 – 1.4)
Model 2	1	0.7 (0.3 – 1.5)	0.7 (0.3 – 1.9)
Model 3	1	0.7 (0.3 – 1.6)	0.7 (0.3 – 1.9)
Frequency of hangover			
Model 1	1	1 (0.5 – 1.9)	1.4 (0.7 – 3.1)
Model 2	1	0.9 (0.3 – 2.3)	1.3 (0.4 – 3.8)
Model 3	1	0.9 (0.4 – 2.4)	1.3 (0.4 – 3.9)
Heavy drinkers			
Model 1	1	0.6 (0.4 – 1)	0.7 (0.4 – 1.4)
Model 2	1	0.6 (0.3 – 1.3)	0.8 (0.3 – 1.9)
Model 3	1	0.7 (0.3 – 1.4)	0.8 (0.3 – 1.9)
Tobacco			
Current smoker			
Model 1	1	2 (1 – 3.9)	4.9 (2.4 – 10.2)
Model 2	1	1.1 (0.4 – 2.6)	2.5 (1 – 6.6)
Model 3	1	1.1 (0.4 – 2.7)	2.5 (1 – 6.7)

Notes:
All values are OR (95% CI). Total n = 989.
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education

6.3.2. Analysis by sub-classification *of migrants*

6.3.2.1. By length of residence in urban area

The group with lowest length of residency in an urban area, less than 20 years, was used as the baseline for all comparisons and results are presented in Table VI-3. A tendency towards increased OR for all outcomes, from lowest to highest length of residency in urban area, was observed. This was more pronounced for alcohol-related outcomes where nearly doubling OR were observed for the two groups with longest number of years living in an urban area. However, some of them were borderline significant while others were non significant. Also, it is notorious that most CI were wide, probably reflecting the low prevalence of each specific outcome in the all study groups and in the migrant group as well (see Table VI-1).

Table VI-3. Multivariable association of length of residence in urban area with behavioural risk factors in migrants

Outcome Category	<20y	20-29y	30-39y	≥40y
Alcohol				
Frequency of consumption				
Model 1	1	1.1 (0.3 – 3.8)	2.7 (0.7 – 10.3)	6.1 (1.1 – 33.7)
Model 2	1	1.1 (0.3 – 3.9)	2.7 (0.7 – 10.5)	6.1 (1.1 – 34.7)
Model 3	1	1.2 (0.3 – 4)	2.8 (0.7 – 10.9)	6.3 (1.1 – 35.7)
Volume of consumption				
Model 1	1	1.2 (0.4 – 4)	2.7 (0.7 – 10.2)	5.2 (1 – 28.1)
Model 2	1	1.2 (0.4 – 4)	2.7 (0.7 – 10.2)	5.2 (0.9 – 28.5)
Model 3	1	1.3 (0.4 – 4.3)	2.9 (0.7 – 11.2)	5.5 (1 – 30.9)
Frequency of hangover				
Model 1	1	1.3 (0.3 – 5.1)	2.7 (0.6 – 12.3)	6.9 (1.1 – 44.7)
Model 2	1	1.2 (0.3 – 4.9)	2.6 (0.6 – 11.7)	6.3 (1 – 41.4)
Model 3	1	1.3 (0.3 – 5.1)	2.7 (0.6 – 12.3)	6.5 (1 – 43.2)
Heavy drinkers				
Model 1	1	1.2 (0.3 – 3.9)	2.7 (0.7 – 10.2)	4.6 (0.9 – 24.5)
Model 2	1	1.2 (0.3 – 3.9)	2.7 (0.7 – 10.2)	4.6 (0.9 – 24.7)
Model 3	1	1.2 (0.4 – 4.2)	2.9 (0.8 – 11.2)	4.9 (0.9 – 26.9)
Tobacco				
Current smoker				
Model 1	1	0.7 (0.3 – 2)	0.8 (0.3 – 2.5)	1 (0.3 – 3.4)
Model 2	1	0.7 (0.3 – 2)	0.8 (0.3 – 2.4)	0.9 (0.3 – 3.3)
Model 3	1	0.7 (0.3 – 2.1)	0.8 (0.3 – 2.5)	0.9 (0.2 – 3.3)

Notes:
All values are OR (95% CI). Total n = 559/559.
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education

6.3.2.2. By lifetime exposure to urban area

This analysis was made using lifetime exposure to urban area in quartiles and the lowest quartile was the reference subgroup for comparisons. The results are presented in Table VI-4. Using this classification, no clear pattern of associations between exposure to urban area and behavioural risk factors were evident, and none of them were significant.

Table VI-4. Multivariable association of lifetime exposure to urban area with behavioural risk factors in migrants

Outcome Category	Q1* (lowest)	Q2	Q3	Q4 (highest)
Alcohol				
Frequency of consumption				
Model 1	1	1 (0.4 – 2.7)	1.7 (0.7 – 4.4)	1.4 (0.6 – 3.7)
Model 2	1	1 (0.4 – 2.7)	1.7 (0.7 – 4.4)	1.4 (0.6 – 3.7)
Model 3	1	1 (0.4 – 2.7)	1.8 (0.7 – 4.5)	1.4 (0.5 – 3.7)
Volume of consumption				
Model 1	1	1.1 (0.4 – 3)	1.7 (0.7 – 4.3)	1.6 (0.6 – 4.1)
Model 2	1	1.1 (0.4 – 3)	1.7 (0.7 – 4.3)	1.6 (0.6 – 4.1)
Model 3	1	1.1 (0.4 – 3)	1.8 (0.7 – 4.6)	1.5 (0.6 – 3.9)
Frequency of hangover				
Model 1	1	0.7 (0.2 – 2.2)	1.8 (0.7 – 4.7)	1.3 (0.5 – 3.6)
Model 2	1	0.7 (0.2 – 2.2)	1.7 (0.6 – 4.5)	1.2 (0.4 – 3.4)
Model 3	1	0.7 (0.2 – 2.2)	1.7 (0.7 – 4.6)	1.2 (0.4 – 3.3)
Heavy drinkers				
Model 1	1	1.3 (0.5 – 3.2)	1.7 (0.7 – 4.3)	1.6 (0.6 – 4.1)
Model 2	1	1.3 (0.5 – 3.2)	1.7 (0.7 – 4.2)	1.6 (0.6 – 4)
Model 3	1	1.3 (0.5 – 3.3)	1.7 (0.7 – 4.5)	1.5 (0.6 – 3.9)
Tobacco				
Current smoker				
Model 1	1	1 (0.4 – 2.3)	1.6 (0.7 – 3.4)	1.3 (0.6 – 3)
Model 2	1	1 (0.4 – 2.3)	1.5 (0.7 – 3.4)	1.3 (0.5 – 3)
Model 3	1	1 (0.4 – 2.3)	1.5 (0.7 – 3.4)	1.2 (0.5 – 2.8)

Notes:
All values are OR (95% CI). Total n = 559/559.
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%

6.3.2.3. By age at first migration

Migrants were sub-classified as below or above 12 years-old. Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults. Those individuals aged 12 years or less when first out-migrated were considered as baseline for comparisons. As shown in Table VI-5, no major differences in the behavioural factors studied in this section were observed between these two groups defined by age at first migration.

Table VI-5. Multivariable association of age at first migration with behavioural risk factors in migrants

Outcome Category	≤12yo*	>12yo
Alcohol		
Frequency of consumption		
Model 1	1	0.9 (0.5 – 1.8)
Model 2	1	0.9 (0.5 – 1.8)
Model 3	1	0.9 (0.5 – 1.8)
Volume of consumption		
Model 1	1	0.8 (0.4 – 1.6)
Model 2	1	0.8 (0.4 – 1.6)
Model 3	1	0.8 (0.4 – 1.6)
Frequency of hangover		
Model 1	1	0.9 (0.5 – 1.9)
Model 2	1	0.9 (0.5 – 1.9)
Model 3	1	0.9 (0.5 – 1.9)
Heavy drinkers		
Model 1	1	0.9 (0.5 – 1.6)
Model 2	1	0.9 (0.5 – 1.6)
Model 3	1	0.9 (0.5 – 1.6)
Tobacco		
Current smoker		
Model 1	1	0.7 (0.4 – 1.2)
Model 2	1	0.7 (0.4 – 1.2)
Model 3	1	0.7 (0.4 – 1.3)

Notes:
All values are OR (95% CI). Total n = 585/585.
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

6.4. Discussion of results

6.4.1. Summary

Prevalence of overall alcohol consumption increased from rural (4%) to migrant (8%) and to urban groups (10%) but there was no difference in heavy drinking status. Prevalences of abstemious status (never drink) were 39%, 18% and 7% for the same groups, respectively. In terms of tobacco consumption, a clear gradient towards higher prevalences of current smoking status from rural to urban groups was observed: 5% in rural people, 10% in migrants and 20% in urban people. Compared to rural population, urban group showed a borderline association in current smoking status (OR 2.5 (95% CI 1 – 6.7)).

6.4.2. Strengths

Strengths of this section rely on its design and the careful selection of study groups and comparison groups. Initially, during the planning, the main group of interest was the migrant group and a rural group for comparison purposes. By expanding the study to include an urban group it enabled the assessment of gradients in specific risk factors across study groups. Such gradients are notorious for overall alcohol consumption, never drinkers and current smokers.

An additional strength of the study was the adaptation and use of standardised instruments, such as the WHO STEPwise approach to chronic disease risk factor surveillance [201] and other tools to address behavioural risk factors. This enabled the use of data gathered in different aggregations that could ensure comparability with other published studies.

6.4.3. Weaknesses

The use of international established instruments for measurement of behavioural risk factors was also a weakness for the present study, especially because the limitations

on its applicability to a different context such as Peru in terms of alcohol consumption and tobacco use.

First, common to all behavioural risk factors evaluated in this chapter is the issue of affordability. In order to be able to engage either on regular alcohol consumption, tobacco use or leisure physical activities, people need the financial capacity to do so. The population groups studied, as shown in Chapter V in the section related to socioeconomic indicators, showed varied proportions of socioeconomic disadvantages. Despite this, it was also observed that frequencies of alcohol and cigarette consumption were low. It could be argued that better instruments for the assessment of alcohol or cigarette consumption in societies with low prevalence of such risk factors would be deemed more appropriate. However, such decision would imply a trade-off versus the comparability of data with other studies. The final decision was to favour, as much as possible, the use of international standards to maintain comparability of results with other population-based studies. Such decision was aided based on the fact that the main outcomes for this study were other variables and a general, and not detailed, picture of behavioural risk factors was only needed. Based on the same evaluation criteria, and following the pilot study, the nutritional component of the study was completely dropped-out.

Second, in case of alcohol consumption, most instruments investigate patterns related to regular alcohol consumption as well as the type of alcohol being consumed. This, in aggregation, provides information about frequency and units of alcohol being consumed. In Peru, and particularly in middle and low socioeconomic status, most people do not engage on "regular" patterns of alcohol consumption mostly because of costs. The social use of alcohol is usually bound to sporadic events during the year (final year and or personal or relative's birthday's celebrations) and is more of a binge-drinking type of alcohol consumption. These local customs do not fit well with the approach towards "regular" use of alcohol that is more common in other settings, particularly in developed societies. In order to overcome this the data generated was aggregated into a heavy drinking category, and as such, it was informative to find out that despite the differential profile of alcohol consumption across groups such profile was not present for heavy drinking status.

Third, in relation to cigarette consumption, it is accepted that number of cigarettes per day, that enable its aggregation into packs per year, would have been a better aggregated variable to use. However, smoking patterns follow, to certain extent, the complexities of alcohol consumption's patterns. Most people, if smokers, buy cigarettes per units and/or smoke sporadically. Therefore, our definition of current smokers (n = 110) was forced to include anyone who has smoked a cigarette within the last six months. As shown in Table VI-1, only 75 people out of 110 current smokers provided a number other than zero to the question of "number of cigarettes smoked in the last 30 days". Such trade-off, to expand the definition of current smoker used in this study to include people who have smoked any cigarette in the last six months, was considered acceptable ever since the question on the last 30 days could miss some sporadic occasional smokers.

In this section, greater detail has been placed on discussing specific limitations related to the choice of instrument made for measurements, data aggregation and its interpretation. Despite behavioural risk factors not being the main outcome of the study, it is considered that alternatives were explored and strategies were put in place to address and overcome, if possible, such limitations. Special emphasis was placed in ensuring comparability of data with other international CVD risk factor surveys.

6.4.4. Interpretation

Prevalences of overall alcohol consumption increased from rural (4%) to migrant (8%) and to urban groups (10%) but there was no difference in heavy drinking status. When aggregated into a single heavy drinking variable, prevalences for the same groups were 12%, 8% and 10%, respectively.

In terms of tobacco consumption, a clear gradient towards higher prevalences of current smoking status from rural to urban groups was observed: 5% in rural people, 10% in migrants and 20% in urban people.

These results suggest that following migration, migrants are more exposed to tobacco and alcohol. Nonetheless, multivariable analyses did not show a clear pattern of differential risk between groups, except for a borderline association in current smoking status amongst urban people (OR 2.5, 95% CI 1 – 6.7). In most cases,

multivariable analyses showed wide CI reflecting the low distribution of cases in each group.

Data from the migrant group analysed separately, using different exposure sub-classifications, did not show any association with behavioural risk factors evaluated in this chapter.

Alcohol and tobacco consumption are important contributors to burden of disease and mortality. They, together with high blood pressure, high cholesterol and obesity were identified by the World Health Report 2002 [229] within the top ten risks, globally and regionally, in terms of the burden of disease they cause. The rise and establishment of CVD epidemic in developed countries has been closely related to tobacco consumption, a major cause of lung cancer. However, the effect of tobacco on contributing towards a sustained increase in CVD and lung cancer-related mortality, as seen in developed countries [230, 231], should not necessarily be applicable to Peru. Ezzati and Lopez, in a study of smoking-attributable mortality, describe Peru as part of a region with a low total adult mortality caused by smoking, in the range of 2–4% [39]. Analyses of Peruvian mortality records show that lung cancer was not listed as a major cause of death [Luis Huicho & J. Jaime Miranda, personal communication].

Information from PAHO about Peru indicates that the annual prevalence of tobacco use was 37.5% (53.3% for men and 23.7% for women) [92]. Such figures do not provide detailed information about current smoking status or disaggregates by region, making difficult its interpretation in relation to the findings of this study. Of note, a clear pattern from rural to urban groups was observed: one of decreasing of never smokers, and at the same time, one of increasing of current smokers in the same direction. In this study, urban population was, although borderline, 2.5 more likely to be current smokers than their rural counterparts.

In terms of alcohol consumption, PAHO reported an annual prevalence of alcohol consumption of 75.1% which was greater in men than in women (79.8% and 71.0%, respectively) [92]. A more detailed PAHO multi-country study that included Peru's urban and rural areas, published in 2007, reports that the proportion of men and

women abstemious was 20 and 27%, respectively [232]. Results of this study provide further insight and indicate that such proportions vary considerably by study group. For example, in terms of overall frequency of alcohol consumption different proportions of abstemious (those who answered "never" to overall frequency of alcohol consumption) were found in rural (39%), migrant (18%) and urban (7%) areas. However, in the multivariable analyses, in all alcohol-related variables, no differences were found between different groups studied.

Despite the difficulties in assessing behavioural risk factors, results from this study provide evidence of differential exposures to potentially harmful behaviours. This was observed in terms of prevalences of alcohol consumption and current smoking status in the groups studied. However, none of these factors showed clear differential risks in multivariate analysis, thus limiting the extent of the conclusions that can be drawn from this part of this study.

6.5. Summary points

- Prevalence of overall alcohol consumption increased from rural (4%) to migrant (8%) and to urban groups (10%) but there was no difference in heavy drinking status. Prevalence of abstemious status (never drink) was 39%, 18% and 7% for the same groups, respectively.

- A clear gradient towards higher prevalences of current smoking status was observed: 5% in rural people, 10% in migrants and 20% in urban people.

- When urban group was compared to rural group, a borderline association was found in current smoking status (OR 2.5 (95% CI 1 – 6.7)). Such association was not significant in the migrant group.

- Multivariable analyses did not show a clear pattern of differential risk between groups in terms of behavioural risk factors.

Chapter VII. Anthropometry

This chapter describes the main outcomes body mass index (BMI), waist-to-hip ratio (WHR) and prevalence of obesity as well as prevalence of combined overweight and obesity. Additionally, information is provided about waist circumference and skinfolds. Since most outcomes in this chapter are continuous ones, results are presented for each specific outcome to provide all the information in context. That is, e.g. in the case of BMI, information on height, weight and BMI mean values are presented, followed by β regression coefficients, followed by its standardised mean difference (SMD). Reported outputs from SMD regressions are those derived from fully adjusted models taking into account age, sex, individual's deprivation index and parental education.

This strategy was selected to maintain clarity in presentation of results. Later on, and for comparative purposes addressing the extent of variation between different but related outcome variables and study groups, a specific section is devoted to the visual presentation of SMD.

7.1. Body mass index

7.1.1. BMI in all study groups

Compared to the rural group, mean BMI was higher in both migrant and urban groups (Table VII-1). Within each specific study group females had higher mean BMI than males.

In multivariable regression, after adjustment for age, sex and socioeconomic confounders, BMI in the migrant population was 3.2 Kg/m^2 (95% CI 2.4 – 4.1) higher than in the rural population. Similarly, BMI was 4.3 Kg/m^2 (95% CI 3.2 – 5.3) higher in the urban population than in rural group.

The difference between groups compared to the rural population in terms of SD units were 0.8 (95% CI 0.6 - 1) and 1 (95% CI 0.8 – 1.3) SD units for migrant and urban people, respectively (Table VII-1).

Table VII-1. Descriptive and multivariable analyses of body mass index by study groups

	Rural	Migrant	Urban	
Descriptive statistics	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Missing data**
Weight (Kg)	53.9 (8.2)	64.1 (10.6)	69.4 (14.5)	2 (0.2%)
Height (m)	1.5 (0.1)	1.5 (0.1)	1.6 (0.1)	2 (0.2%)
BMI (Kg/m^2)				2 (0.2%)
All	23.2 (2.7)	27 (4.3)	28.3 (5.4)	
Female	23.5 (3.2)	28 (4.7)	29.5 (6.1)	
Male	22.9 (2.1)	25.9 (3.5)	26.8 (4)	
Multivariable BMI		**β coefficient (95% CI)**	**β coefficient (95% CI)**	**R^2**
Model 1	Reference	3.8 (3.1 – 4.5)	5 (4.2 – 5.9)	0.18
Model 2	Reference	3.2 (2.3 – 4.1)	4.4 (3.3 – 5.4)	0.19
Model 3	Reference	3.2 (2.4 – 4.1)	4.3 (3.2 – 5.3)	0.19
	Within groups SD	**SMD (95% CI)**	**SMD (95% CI)**	**R^2**
SMD BMI	4.2	0.8 (0.6 – 1)	1 (0.8 – 1.3)	0.19

Notes: All regressions calculated using the rural group as baseline (Total n = 987/989).
Model 1: Adjusted for age and sex; Model 2: As model 1 plus individual socioeconomic deprivation; Model 3: As model 2 plus highest parental education (this final model was also used for SMD regressions)

7.1.2. BMI in migrant subgroups

To address the specific research questions of this study, sub-classifications by length of residence in urban area, lifetime exposure to an urban area and age at first migration were analysed in the migrant population only. Table VII-2 shows BMI's β coefficients (95% CI) in Kg/m^2 units of comparisons between BMI and various categories used.

Most CI observed did cross the value zero, thus indicating that those β coefficients are not significantly different from their respective baseline comparison group. There were two isolated exceptions where observations CI did not cross value zero. First, people who were living 30-39 years in an urban area had in average 2 Kg/m^2 (95% CI 0.7 – 3.4) more than those who were living less than 20 years in an urban area. The other exception was observed in the second quartile of lifetime exposure to urban environment compared to the lowest quartile, who had 1.6 Kg/m^2 (95% CI 0.7 – 2.6) units higher.

Thus, no clear pattern of variation of risk profile, in terms of BMI, was observed within the migrant group using various sub-classifications.

Table VII-2. Multivariable analyses of migrant sub-classifications and body mass index

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	1.3 (0 – 2.5)	1.2 (0 – 2.5)	1.2 (-0.1 – 2.5)
30-39 years in urban area	2.1 (0.8 – 3.5)	2.1 (0.7 – 3.4)	2.0 (0.7 – 3.4)
≥40 years in urban area	1.3 (-0.2 – 2.8)	1.2 (-0.3 – 2.8)	1.2 (-0.3 – 2.7)
R^2	0.08	0.08	0.08
Migrants, by lifetime exposure to urban area*, n = 558/559			
Q1, lowest	Reference	Reference	Reference
Q2	1.7 (0.7 – 2.6)	1.7 (0.7 – 2.6)	1.6 (0.7 – 2.6)
Q3	0.9 (-0.1 – 1.8)	0.8 (-0.2 – 1.7)	0.8 (-0.2 – 1.7)
Q4, highest	0 (-1 – 0.9)	-0.1 (-1.1 – 0.9)	-0.1 (-1.1 – 0.9)
R^2	0.09	0.09	0.09
Migrants, by age at first migration**, n = 584/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	0.7 (0 – 1.4)	0.7 (0 – 1.4)	0.7 (0 – 1.4)
R^2	0.06	0.07	0.07

Note: All values, except R^2 values, are β coefficients (95% CI) in Kg/m^2 units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

7.2. Overweight and Obesity

The distribution of BMI aggregated into standard categories —underweight, normal, overweight and obesity— is presented in Table VII-3. As shown in this table, it is clear that underweight is not a major problem for the populations studied. Furthermore, it is striking the imbalance of normal versus overweight/obesity. While nearly 80% of rural population falls within the normal category according to BMI, only 30% does so in the migrant and urban groups. Both, migrant and urban populations have much higher prevalences of overweight and obesity than their rural counterparts.

Obesity was one of the main outcomes defined for this study. However, due to the important magnitude of overweight and obesity in the groups studied, it was decided to explore the combination of both, overweight and obesity, as a joint outcome.

Table VII-3. Prevalence of underweight, overweight and obesity by study groups

BMI category	Rural	Migrant*	Urban
Underweight	2 (1%)	3 (0.5%)	2 (1%)
Normal	160 (79.6%)	189 (32.2%)	56 (28.1%)
Overweight	33 (16.4%)	271 (46.2%)	73 (36.7%)
Obesity	6 (3%)	124 (21.1%)	68 (34.2%)

Note: Underweight defined as BMI < 18.5 Kg/m^2; Overweight as BMI ≥ 25 Kg/m^2 and BMI < 30 Kg/m^2; and, Obesity as BMI ≥ 30 Kg/m^2 (Total n = 987/989, 2 missing people on the migrant group).

7.2.1. Obesity

Table VII-4 presents the distribution and OR (95% CI) of obesity in all study groups and according to migrant subgroups.

Obesity, defined as BMI ≥ 30 Kg/m^2, was prevalent in 3%, 21% and 34% of rural, migrant and urban population, respectively. In the fully adjusted model compared to the rural group, migrant and urban people were 9.5 (95% CI 3.8 – 23.4) and 20.1 (95% CI 7.6 – 53.3) times more likely to be obese than rural people, respectively. Such wide CI can be related to the fact that only 6/201 subjects in the rural group were obese (Table VII-4). OR were similar in crude and fully adjusted models.

When exploring the association of obesity within migrants, a tendency towards increase in OR was observed in migrants by number of years living in urban area. That is, compared to those migrants living less than 20 years in an urban area, migrants living 20-29 years in urban area were 3.7 (95% CI 1.2 – 11) times more likely to be obese. Similarly, migrants living 30-39 years in urban area were 4.2 (95% CI 1.4 – 12.9) times more likely to be obese. Such association was not observed for those living ≥40 years in urban area (2.7 (95% CI 0.8 – 9)).

No clear associations between obesity and lifetime exposure to urban area or age at first migration exposures were observed.

Table VII-4. Association between obesity and migration by exposure groups

Exposure Groups	Obese n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Rural, Migrant, Urban	**(n = 987/989)**			
Rural	6 (3%)	1	1	1
Migrant	124 (21.1%)	9.4 (4 – 21.7)	9.2 (3.7 – 22.7)	9.5 (3.8 – 23.4)
Urban	68 (34.2%)	18.7 (7.8 – 44.7)	18.3 (7.1 – 47.7)	20.1 (7.6 – 53.3)
Migrants, by years in urban area	**(n = 558/559)**			
<20 years in urban area	12 (16%)	1	1	1
20-29 years in urban area	48 (25.3%)	3.9 (1.3 – 11.4)	3.9 (1.3 – 11.7)	3.7 (1.2 – 11)
30-39 years in urban area	31 (20%)	4.2 (1.4 – 12.9)	4.4 (1.4 – 13.5)	4.2 (1.4 – 12.9)
≥40 years in urban area	27 (18.5%)	2.7 (0.8 – 9)	2.8 (0.8 – 9.4)	2.7 (0.8 – 9)
Migrants, by lifetime exposure to urban area*	**(n = 558/559)**			
Q1, lowest	24 (21.1%)	1	1	1
Q2	40 (35.1%)	2.1 (1.2 – 3.9)	2.1 (1.2 – 3.9)	2 (1.1 – 3.7)
Q3	29 (25.4%)	1.4 (0.7 – 2.5)	1.4 (0.7 – 2.6)	1.4 (0.7 – 2.5)
Q4, highest	21 (18.4%)	0.9 (0.4 – 1.7)	0.9 (0.5 – 1.7)	0.9 (0.4 – 1.7)

Table VII-4. (continued)

Exposure Groups	Obese n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Migrants, by age at first migration**	**(n = 584/585)**			
≤ 12 yo at first migration	40 (17.9%)	1	1	1
> 12 yo at first migration	84 (23.4%)	0.8 (0.4 – 1.3)	0.8 (0.4 – 1.3)	0.7 (0.4 – 1.3)

Notes: All values are OR (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

7.2.2. Overweight and obesity

Table VII-5 presents the distribution and OR (95% CI) of an aggregated outcome, overweight and obesity (BMI ≥ 25 Kg/m^2), in all study groups and according to migrant sub-classifications.

Overweight or obesity was prevalent in 19.4%, 67.3% and 70.9% of rural, migrant and urban population, respectively. Compared to rural group, crude OR for overweight or obesity were 8.8 (95% CI 6 – 13.1) for migrants and 10.4 (95% CI 6.5 – 16.7) for urban people. Following adjustment for socioeconomic indicators, such OR were attenuated but remained significant. In the fully adjusted model compared to the rural group, migrant and urban people were 5.9 (95% CI 3.7 – 9.4) and 5.7 (95% CI 3.2 – 10.2) times more likely to be overweight or obese than rural people, respectively.

When exploring the association of overweight or obesity within migrants in terms of number of years in an urban area, migrants living 20-29 years in urban area were 80% more likely to be overweight or obese compared to those migrants living less than 20 years in an urban area (OR 1.8 (95% CI 1 – 3.5)). This association was borderline significant and was not observed for the other sub-groups in the same category.

Using a different sub-classification of migrants, in quartiles of lifetime of exposure to urban area, only the second quartile showed increased odds for overweight or obesity compared to the lowest quartile group (OR 2.2 (95% CI 1.3 – 3.9)). No associations were observed in the other quartiles evaluated.

In terms of age at first migration, a borderline significant association of increased odds was observed amongst those aged 12 or more at first migration compared to those who migrated at earlier ages (OR 1.4 (95% CI 1 – 2)).

Table VII-5. Association between overweight or obesity and migration by exposure groups

Exposure Groups	Overweight or Obese n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Rural, Migrant, Urban	**(n = 987/989)**			
Rural	39 (19.4%)	1	1	1
Migrant	395 (67.3%)	8.8 (6 – 13.1)	5.8 (3.6 – 9.2)	5.9 (3.7 – 9.4)
Urban	141 (70.9%)	10.4 (6.5 – 16.7)	6.3 (3.6 – 11.1)	5.7 (3.2 – 10.2)
Migrants, by years in urban area	**(n = 558/559)**			
<20 years in urban area	29 (54.7%)	1	1	1
20-29 years in urban area	140 (69%)	1.9 (1 – 3.6)	1.8 (1 – 3.5)	1.8 (1 – 3.5)
30-39 years in urban area	118 (69.8%)	2 (1 – 3.9)	1.8 (0.9 – 3.6)	1.8 (0.9 – 3.6)
≥40 years in urban area	86 (64.7%)	1.5 (0.7 – 3.2)	1.3 (0.6 – 2.8)	1.3 (0.6 – 2.8)
Migrants, by lifetime exposure to urban area*	**(n = 558/559)**			
Q1, lowest	90 (63.8%)	1	1	1
Q2	110 (79.1%)	2.2 (1.3 – 3.8)	2.2 (1.3 – 3.9)	2.2 (1.3 – 3.9)
Q3	89 (63.1%)	1 (0.6 – 1.6)	0.9 (0.5 – 1.5)	0.9 (0.5 – 1.5)
Q4, highest	84 (61.3%)	0.9 (0.5 – 1.5)	0.8 (0.5 – 1.3)	0.8 (0.5 – 1.3)

Table VII-5. (continued)

Exposure Groups	Overweight or Obese n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Migrants, by age at first migration**	**(n = 584/585)**			
≤ 12 yo at first migration	141 (62.7%)	1	1	1
> 12 yo at first migration	525 (70.2%)	1.4 (1 – 2)	1.4 (1 – 2)	1.4 (1 – 2)

Notes: All values are OR (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

7.3. Waist-to-hip ratio

7.3.1. WHR in all study groups

Table VII-6 shows mean (SD) of waist circumference, hip circumference and WHR. Means of waist and hip circumferences show a gradual increase from rural to urban group. Mean of their ratio for both sexes, WHR, the chosen outcome of the study, was similar in migrant and urban population (0.93 and 0.92, respectively) but lower in rural group (0.87).

Following international recommendations for cut-offs of increased risk when WHR >0.9 in male and WHR>0.85 for females [159], it was observed that mean WHR in both sexes of all groups fell at the increased risk category except for rural women (WHR 0.83). Such data disaggregated by sex was presented for informative purposes. However, exploring sex differences was neither a main nor a specific research question for this study.

Linear regression analyses showed that, compared to the rural group, migrant people had a higher WHR (β coefficient 0.06 (95% CI 0.04 – 0.07)) as did urban people (β coefficient 0.04 (95% CI 0.03 – 0.06)). In the calculations of SMD, migrant and urban population showed significantly differences in SD units compared to the rural population (Table VII-6). However, such differences were quite small, in the order of decimal points, due to the very nature of the index WHR whose values are usually below 1.

Table VII-6. Descriptive and multivariable analyses of waist-to-hip ratio by study groups

	Rural	Migrant	Urban	
Descriptive statistics	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Missing**
Waist circumference (cm)	76.1 (8.4)	88.1 (9.9)	91.4 (12.1)	6 (0.6%)
Hip circumference (cm)	87.4 (5.3)	94.4 (7.5)	98.9 (10.6)	8 (0.8%)
WHR				8 (0.8%)
All	0.87 (0.07)	0.93 (0.07)	0.92 (0.07)	
Female	0.83 (0.07)	0.90 (0.06)	0.90 (0.06)	
Male	0.91 (0.05)	0.97 (0.05)	0.95 (0.05)	
Multivariable WHR		**β coefficient (95% CI)**	**β coefficient (95% CI)**	**R^2**
Model 1	Reference	0.06 (0.05 – 0.07)	0.05 (0.04 – 0.07)	0.35
Model 2	Reference	0.06 (0.04 – 0.07)	0.04 (0.03 – 0.06)	0.35
Model 3	Reference	0.06 (0.04 – 0.07)	0.04 (0.03 – 0.06)	0.35
	Within group SD	**SMD (95% CI)**	**SMD (95% CI)**	**R^2**
SMD WHR	0.06	1 (0.8 – 1.2)	0.7 (0.5 – 1)	0.35
SMD Waist circumference	10	1 (0.8 – 1.2)	1.3 (1 – 1.5)	0.23

Notes: All regressions calculated using the rural group as baseline (Total n = 981/989).
Model 1: Adjusted for age and sex; Model 2: As model 1 plus individual socioeconomic deprivation; Model 3: As model 2 plus highest parental education (this final model was also used for SMD regressions)

7.3.2. WHR in migrant subgroups

WHR was also evaluated within migrant group using different sub-classifications (Table VII-7). A borderline pattern of increased WHR with change in migration category was observed in migrants with increased length of residence in urban areas as measured by absolute number of years spent in an urban area, as demonstrated by β coefficients that did not cross value zero. Such pattern was not evident when migrants were divided by lifetime exposure to urban area or age at first migration.

Table VII-7. Multivariable analyses of migrant sub-classifications and waist-to-hip ratio

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 557/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	0.01 (0 – 0.03)	0.01 (0 – 0.03)	0.01 (0 – 0.03)
30-39 years in urban area	0.03 (0.02 – 0.05)	0.03 (0.01 – 0.05)	0.03 (0.01 – 0.05)
≥40 years in urban area	0.03 (0.01 – 0.05)	0.03 (0.01 – 0.05)	0.03 (0.01 – 0.05)
R^2	0.31	0.31	0.31
Migrants, by lifetime exposure to urban area*, n = 557/559			
Q1, lowest	Reference	Reference	Reference
Q2	0.01 (0 – 0.02)	0.01 (0 – 0.02)	0.01 (0 – 0.02)
Q3	0.01 (-0.01 – 0.02)	0.01 (-0.01 – 0.02)	0.01 (-0.01 – 0.02)
Q4, highest	0 (-0.02 – 0.01)	-0.01 (-0.02 – 0.01)	-0.01 (-0.02 – 0.01)
R^2	0.29	0.3	0.3
Migrants, by age at first migration**, n = 583/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	0.01 (0 – 0.02)	0.01 (0 – 0.02)	0.01 (0 – 0.02)
R^2	0.3	0.3	0.3

Note: All values, except R^2 values, are β coefficients (95% CI) of WHR.
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

7.4. Skinfolds

7.4.1. Skinfolds in all study groups

The primary outcome for this was a single skinfolds indicator which summed all individual skinfolds as presented earlier in Table II-10 (page 55). However, for informative purposes the distributions of each individual skinfold site measurement are provided as median (IQR) in Table VII-8, acknowledging their skewed distribution.

Table VII-8 shows consistently that mean of sum of all skinfolds doubles from rural to migrant group and it was even higher in the urban group. Such observations were confirmed by higher β coefficients as well as higher SMD in both migrant and urban populations compared to the rural group.

Table VII-8. Descriptive and multivariable analyses of skinfolds by study groups

	Rural	Migrant	Urban	
Descriptive statistics				**Missing**
Biceps (mm), median (IQR)	4.2 (3.1)	8.5 (8.4)	13.1 (13.6)	9 (0.9%)
Triceps (mm), median (IQR)	10.2 (8.3)	21.1 (19.5)	30 (21.4)	9 (0.9%)
Subscapular (mm), median (IQR)	11.3 (6.9)	19.7 (11.2)	23.3 (12)	9 (0.9%)
Suprailiac (mm), median (IQR)	9.7 (10.3)	27.1 (11.3)	29.1 (14.4)	13 (1.3%)
Sum of all skinfolds (mm)				13 (1.3%)
All, mean (SD)	42 (20.1)	81.4 (30.6)	96 (35.1)	
Female, mean (SD)	49.1 (20.3)	97 (26.9)	111.3 (33)	
Male, mean (SD)	34.1 (16.6)	63.7 (24.4)	77.9 (28.3)	
Multivariable All skinfolds		**β coefficient (95% CI)**	**β coefficient (95% CI)**	**R^2**
Model 1	Reference	39 (34.9 – 43.1)	53.4 (48.4 – 58.4)	0.47
Model 2	Reference	34.4 (29.2 – 39.7)	48.1 (41.8 – 54.5)	0.48
Model 3	Reference	34.8 (29.5 – 40.1)	45.8 (39.3 – 52.3)	0.48
	Within group SD	**SMD (95% CI)**	**SMD (95% CI)**	**R^2**
SMD All skinfolds	25.3	1.4 (1.2 – 1.6)	1.8 (1.6 – 2.1)	0.48

Notes: All regressions calculated using the rural group as baseline (Total n = 976/989).
Model 1: Adjusted for age and sex; Model 2: As model 1 plus individual socioeconomic deprivation; Model 3: As model 2 plus highest parental education (this final model was also used for SMD regressions)

7.4.2. Skinfolds in migrant subgroups

When all skinfolds were evaluated in migrant subgroups, all β coefficients observed were higher compared to their specific baseline group. However, in all of them except in two occasions, such coefficients were not significant (did overlap value zero).

Table VII-9. Multivariable analyses of migrant sub-classifications and skinfolds

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 551/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	6.4 (-1.4 – 14.2)	6.1 (-1.7 – 13.9)	6.3 (-1.5 – 14.1)
30-39 years in urban area	12.1 (3.9 – 20.4)	11.6 (3.2 – 19.9)	11.8 (3.4 – 20.1)
≥40 years in urban area	10.2 (0.9 – 19.6)	9.5 (0.1 – 18.9)	9.3 (-0.1 – 18.7)
R^2	0.32	0.33	0.33
Migrants, by lifetime exposure to urban area*, n = 551/559			
Q1, lowest	Reference	Reference	Reference
Q2	7.1 (1.1 – 13.2)	7.2 (1.2 – 13.2)	7.4 (1.3 – 13.4)
Q3	6.7 (0.6 – 12.7)	6.1 (0 – 12.1)	5.9 (-0.2 – 11.9)
Q4, highest	1.6 (-4.4 – 7.6)	1.1 (-5 – 7.1)	0.8 (-5.3 – 6.8)
R^2	0.32	0.32	0.33
Migrants, by age at first migration**, n = 577/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	2.9 (-1.5 – 7.3)	2.9 (-1.5 – 7.3)	3.2 (-1.2 – 7.6)
R^2	0.3	0.31	0.31

Note: All values, except R^2 values, are β coefficients (95% CI) in mm units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

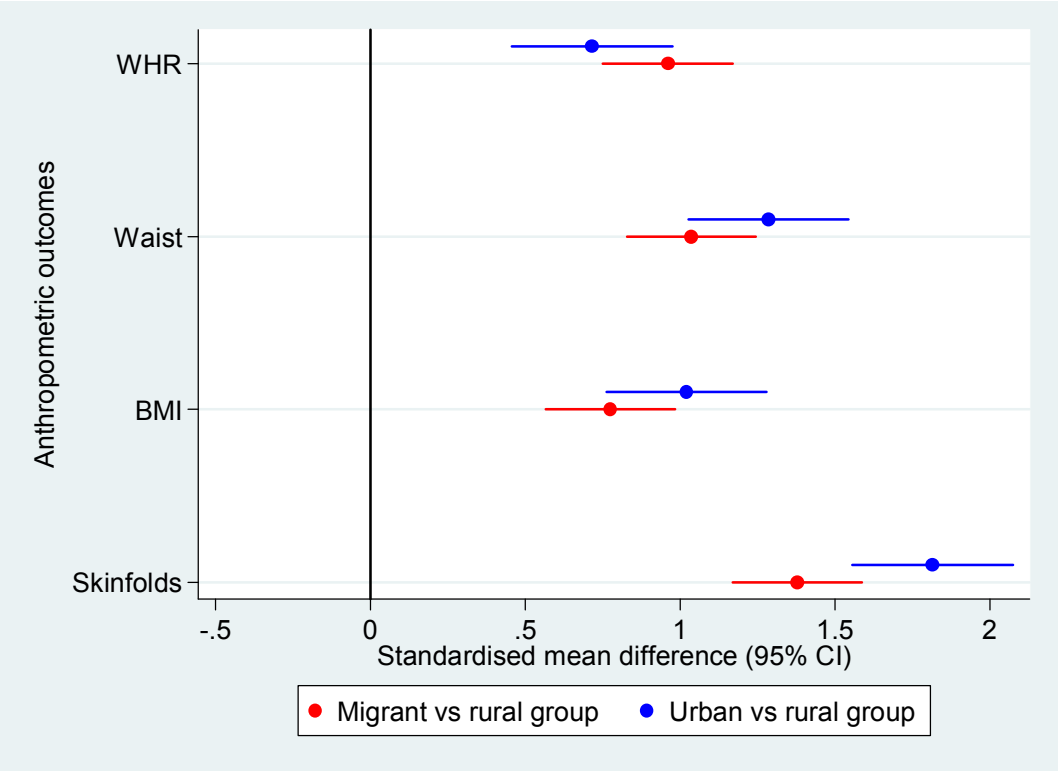
7.5. SMD in anthropometric outcomes

In Chapter II, the rationale for presenting SMD was elaborated (see Section 2.4). Figure VII-1 shows the SMD comparisons against rural group of all continuous outcomes presented in this chapter.

All anthropometric SMD for migrant and urban people were greater than 0.5 compared to the rural population, a notorious difference in terms of risk profile. Compared to the rural population, the single risk factor with most variation between groups was skinfolds, with SMD of 1.4 and 1.8 for the migrant and urban groups, respectively.

Compared to rural people, the "size" of difference observed was usually bigger in urban than in migrant population, except with WHR SMD which suggested that this was the only risk factor where SMD in migrant group was higher than in the urban one. In all cases, with the exception of skinfolds, however, SMD's CI's for migrant and urban groups did overlap. This observation makes it difficult to ascertain whether or not a specific risk factor's SMD was much higher in either the migrant or urban group when compared to the rural population.

Figure VII-1. Standardised mean differences of anthropometric outcomes in migrant and urban population compared to rural population



Note: Although not a specified outcome of the study, SMD for waist circumference are also presented in this graph for visual comparison with WHR. Waist circumference's means for each specific group are presented in Table VII-6, alongside median for WHR.

7.6. Discussion of results

7.6.1. Summary

This was the first chapter where results of main outcomes of the study were presented. All anthropometric risk factors evaluated showed a pattern of higher risk profiles among migrant and urban groups compared to the rural group. This study found very high OR for the association between obesity —as well as various other anthropometric indicators— and migration in Peru. Migrant and urban people were 9.5 (95% CI 3.8 – 23.4) and 20.1 (95% CI 7.6 – 53.3) times more likely to be obese than rural people. OR for overweight or obesity were 5.9 (95% CI 3.7 – 9.4) and 5.7 (95% CI 3.2 – 10.2) for migrant and urban people, respectively.

7.6.2. Strengths

The strength of this chapter is that it reports well established risk factors for CVD, such as obesity and WHR, and describes its differences between groups studied. Such strengths are notorious in two areas: its analytical approach and use of data, as well as standardised data collection for comparability purposes.

In additional to traditional assessment of statistical associations using linear or logistic regression, this study used SMD. The use and presentation of SMD added clarity when addressing the magnitude of the difference in specific risk factors. All anthropometric outcomes evaluated in this study have specific units, and thus it can be difficult to appreciate the magnitude of the difference between different outcomes. For example, in the case of WHR —which has a range of values around 1— most of its variation between individuals and populations occurs in decimal points. On the contrary, in BMI most changes or variations occur in single units, as it does with skinfolds. Because of the nature of such indicators, such changes in units rather than decimals can mislead to the interpretation of overall change between these risk factors and SMD can be a useful tool to address such limitations. SMD can provide a comparative view in terms of each outcome mean and SD distribution, in a unit free scale. As an example, the absolute mean difference in WHR between migrant and

rural people was 0.06 (difference between means: 0.93 minus 0.87). When expressed in SMD, such difference was around one SD, very similar to the change observed in BMI. Thus, SMD can visually inform variations within single risk factors compared to other risk factors, despite their difference in units of measurement (see Figure VII-1).

An additional strength of this section was that it uses widely accepted indicators measured using standardised methods that can make the data comparable to other populations or studies. Both, the strength of methods used and comparability of data with other studies is further elaborated later in this chapter.

7.6.3. Weaknesses

One of the limitations of the analysis plan set out for this study is that it does not addresses in depth potential differences that may exist in the profile of risk factors in terms of sex. For example, it is widely accepted that WHR cut-offs are set differently for male and female population. When appropriate, baseline means and SD were presented disaggregated by sex in the descriptive tables. However, it was decided that the main question of this study was to address if there was any difference in CVD risk profile between study groups, and more specifically if those differences vary by patterns of migration. Now that these questions have been answered, demonstrating a clear differential profile between groups but not in terms of patterns of migration within migrants, this study opens the avenue for further analyses, such as potential differences in sex in specific CVD risk factors.

7.6.4. Interpretation

This section concentrates on the discussion of results presented in this chapter as well as comparability with other relevant published information for Peru.

It has been argued as one of the strengths of the analysis strategy is that SMD constitute a simple and clear way to present magnitude of the difference across risk factors. The next question would be to address whether the magnitude of differences in specific risk factors is uniform or not. In this chapter, SMD for anthropometric indicators compared to rural population suggest that, at least in terms of magnitude

of the difference, a non-uniform pattern of change was observed. Differences — including their respective CI— ranged from 0.5 to 2 SD (see Figure VII-1).

Cochrane guidelines suggest that "rules of thumb exist for interpreting SMD (or 'effect sizes'), which have arisen mainly from researchers in the social sciences. One example is as follows: 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 a large effect... Variations exist (for example, <0.41 = small, 0.40 to 0.70 = moderate, >0.70 = large)" [175]. This interpretation was suggested for the evaluation of effect sizes after specific interventions (e.g. a randomised controlled trial), hence, a better use of the term 'effect sizes' in this study would be 'difference sizes'.

These rules of thumb, suggest that the differences observed in risk profile between rural and urban populations compared to the rural one were in the range of large differences: all SMD for BMI, WHR, waist circumferences and skinfolds were of the magnitude of 0.8 or greater. This constitutes an important observation since a priori such magnitude of difference was not known nor expected.

Such observations in magnitude of differences are important to address since the hypothesis of this study suggests that CVD risk factors increase following migration. The extent of such change can consider the plausibility that the profile of risk factors for migrants are similar to or, potentially could be, much worse than their urban counterparts. The latter observation seemed to arise in the case of WHR, in that, when compared to the rural group, migrants had a higher WHR's SMD than urban people. Such observation was not significant because of the overlap of SMD's CI between migrant and urban groups. It does raise, however, the point that the magnitude of the difference and the direction of change for specific CVD risk factors might not be uniform across risk factors.

Following on this, the observations made in this chapter raises important questions and challenges for future research in these populations, as well as the development of future interventions. For example, when addressing the epidemic of obesity in Peru, should research and policies concentrate in all groups or in specific groups? This study shows that a difference in patterns of anthropometric risk factors was observed and results obtained will clearly inform future Peruvian health policies at the national

level. In the same vein, when establishing further research on interventions or evaluating the impact of health policies, locally or internationally, the question may arise as to which anthropometric indicator is the best suited to assess the effect or impact of those interventions? Is it changes in BMI, WHR or waist circumference? Controversy has been raised in relation to the adequacy of BMI as a marker for obesity [233, 234]. Such point is raised since, despite the differential profile established for the groups studied, the difference between outcomes in each population was not uniform in terms of magnitude of difference as expressed by SMD. Skinfolds' SMD in both migrant and rural population were the ones with greater difference than the ones for WHR, BMI or waist circumference.

Moving on to comparability of results with other published data, it is worth devoting a few lines to obesity. In the case of obesity, when comparing the results of this study with data from other Peruvian sources (see Table I-2, page 21), it is notorious that the prevalences reported in previous studies —mostly urban ones— do not represent the range of observations found in this study. This study clearly identified that for a much defined group of rural people the prevalence of obesity was very low (around 3%). However, prevalence of obesity in migrants was 21% and 34% for urban people. Peruvian literature suggests that prevalence ranges from 10 to 20% for broadly defined groups. Taking into account the limitations in generalising the results observed because of the populations involved, this study provides further insights into understanding the profile of anthropometric risk factors in Peru emphasizing that Peru is not a uniform group but an aggregation of different populations undergoing different stages of the nutritional and epidemiological transition.

7.7. Summary points

- People who had migrated from rural areas to urban areas had higher mean BMI, WHR and skinfolds compared with people living in the rural area from where migrants had originated.

- The anthropometric profiles of rural-to-urban migrants were remarkably similar to those of people who had always lived in the urban area.

- Obesity was prevalent in 3%, 21% and 34% of rural, migrant and urban population, respectively. Migrant and urban people were 9.5 (3.8 – 23.4) and 20.1 (7.6 – 53.3) times more likely to be obese than rural people, respectively.

- Migrant and urban people were 5.9 (95% CI 3.7 – 9.4) and 5.7 (95% CI 3.2 – 10.2) times more likely to be overweight or obese than rural people, respectively.

- All anthropometric SMD for migrant and urban people were greater than 0.5 compared to the rural population, a notorious difference in terms of risk profile. Compared to the rural population, the single risk factor with most variation between groups was skinfolds, with SMD of 1.4 and 1.8 for the migrant and urban groups, respectively.

- In terms of migrant sub-classification —by number of years living in urban area, lifetime exposure to urban area or age at first migration— all anthropometric outcomes evaluated in this chapter did not show a clear pattern of variation compared to their baseline group.

Chapter VIII. Blood pressure, lipid profile and inflammation markers

This chapter presents results related to systolic blood pressure (SBP), diastolic blood pressure (DBP), hypertension and lipid profile as well as inflammation markers including C-reactive protein (CRP) and fibrinogen. Conversion factors to the international System of Units are provided at the beginning of this document, on page xxvii.

Results are presented for each specific outcome to provide all the information in context. Hypertension is the only categorical outcome in this chapter. For general informative purposes and due to the biological differences in lipid and inflammation markers by gender, descriptive statistics are presented disaggregated by sex in those specific cases.

Additionally, this is the first chapter where results from a log transformed variable, i.e. triglycerides, CRP and fibrinogen, are presented. In the latter cases, median and interquartile range (IQR) for descriptive statistics, geometric means and ratios for multivariable analyses, and standardised mean differences (SMD) of the log transformed variable, are shown. Geometric means and ratios were chosen over β regression coefficients of the log transformed variable because of the convenience of their interpretation [171, 172]. For all other continuous normally-distributed outcomes mean values, β regression coefficients and SMD are presented. Reported outputs from SMD regressions are those derived from fully adjusted models taking into account age, sex, individual's deprivation index and parental education.

Later on, and for comparative purposes addressing the extent of variation between different outcome variables and study groups, a specific section is devoted to the visual presentation of SMD.

8.1. Blood pressure

8.1.1. Statistical analysis

This section is very specific and relevant only for the evaluation of blood pressure as a continuous outcome. As recommended by Tobin et al., "a population-based study of a quantitative trait may be seriously compromised when the trait is subject to the effects of a treatment" [173]. That is, the observed or measured blood pressure (BP) may not reflect the underlying BP because of a treatment effect. Instead of ignoring this fact or exclude treated subjects in linear regression analyses, censored normal regression analyses were used as recommended [173]. This approach was only used in the case of blood pressure outcomes but not in lipid outcomes as no subjects reported to be on statins or other lipid medication.

8.1.2. SBP and DBP in all study groups

Results presented were derived from censored normal regressions as described in section 2.4.2. However, they did not differ from "normal" linear regression analyses, probably due to the smaller percentage of participants currently taking a medication and with controlled hypertension (proportions are presented in Section 8.2).

Migrants had lower mean SBP and lower mean DBP compared to the rural group, but differences were quite small. The opposite, higher mean SBP and DBP compared to rural ones, was observed in urban population (Table VIII-1). Although not an objective of this study, mean SBP and DBP was lower in females compared to male population.

In multivariable regression, after adjustment for age, sex and socioeconomic confounders, SBP did not differ between rural and migrant groups (β coefficient -0.7 mm Hg (95% CI -2.7 to 4.2)). That is, in average migrants have 0.7 mm Hg less than rural people, but CI did cross value of zero. On the contrary, SBP was 9.1 mm Hg (95% CI 4.8 – 13.3) higher in the urban population than in rural group.

Similarly, in multivariable regression and after adjustment for age, sex and socioeconomic confounders, DBP in the migrant population was significantly lower than in the rural group (β coefficient -3.4 mm Hg (95% CI -5.4 to -1.4)) while DBP did not differ between rural and urban groups.

The difference in SBP and DBP between groups compared to the rural population in terms of SMD followed the same inconsistent pattern described earlier. Compared to the rural group, SMD were significantly different only for SBP in urban and DBP in migrant groups (Table VIII-1).

Table VIII-1. Descriptive and multivariable analyses of systolic and diastolic blood pressure by study groups

	Rural	Migrant	Urban	
Descriptive statistics	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Missing data**
SBP (mm Hg)				
All	120.9 (18.7)	119.9 (16.4)	128.2 (22.9)	1 (0.1%)
Female	117.1 (19.1)	116.6 (16.9)	124.7 (22.2)	
Male	125.2 (17.4)	123.5 (14.9)	132.2 (23.2)	
DBP (mm Hg)				
All	74.2 (9.2)	71.3 (9.3)	76.2 (11.5)	1 (0.1%)
Female	72.5 (9.2)	69 (8.7)	73.7 (10.6)	
Male	76.1 (8.9)	74 (9.2)	79.1 (9.9)	
Multivariable SBP*		**β coefficient (95% CI)**	**β coefficient (95% CI)**	
Model 1	Reference	-0.4 (-3.1 – 2.2)	8.2 (4.9 – 11.4)	
Model 2	Reference	0.6 (-2.8 – 4)	9.3 (5.2 – 13.4)	
Model 3	Reference	0.7 (-2.7 – 4.2)	9 (4.8 – 13.3)	
Multivariable DBP*				
Model 1	Reference	-2.6 (-4.2 – -1.1)	2.5 (0.6 – 4.4)	
Model 2	Reference	-3.6 (-5.6 – -1.6)	1.4 (-1 – 3.8)	
Model 3	Reference	-3.4 (-5.4 – -1.4)	1.3 (-1.2 – 3.7)	

Table VIII-1. (continued)

	Rural	Migrant	Urban	
	Within groups SD	**SMD (95% CI)**	**SMD (95% CI)**	**R^2**
SMD SBP	16.1	0 (-0.2 – 0.2)	0.5 (0.3 – 0.8)	0.26
SMD DBP	9.4	-0.4 (-0.6 – -0.2)	0.1 (-0.1 – 0.4)	0.11

Notes: All regressions calculated using the rural group as baseline (Total n = 988/989).
Model 1: Adjusted for age and sex; Model 2: As model 1 plus individual socioeconomic deprivation; Model 3: As model 2 plus highest parental education (this final model was also used for SMD regressions)
* Multivariable analyses for SBP and DBP done with censored normal regressions, which do not calculate an R^2 as with other linear regressions.

8.1.3. SBP and DBP in migrant subgroups

Sub-classifications by length of residence in urban area, lifetime exposure to an urban area and age at first migration were analysed in the migrant population only. Table VIII-2 and Table VIII-3 show BMI's β coefficients (95% CI) in mm Hg units of comparisons between various categories used within migrants and SBP and DBP.

All CI observed did cross the value zero, indicating that those β coefficients were not significantly different from their respective baseline comparison group. Thus, no clear pattern of variation of risk profile, in terms of SBP and DBP, was observed within the migrant group using various sub-classifications.

Table VIII-2. Multivariable analyses of migrant sub-classifications and systolic blood pressure

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	-0.9 (-5.3 – 3.4)	-0.8 (-5.1 – 3.5)	-0.8 (-5.1 – 3.5)
30-39 years in urban area	-0.7 (-5.3 – 3.9)	-0.3 (-4.9 – 4.3)	-0.4 (-4.9 – 4.2)
≥40 years in urban area	-2.2 (-7.4 – 3)	-1.7 (-7 – 3.5)	-1.9 (-7.1 – 3.3)
Migrants, by lifetime exposure to urban area*, n = 558/559			
Q1, lowest	Reference	Reference	Reference
Q2	1.5 (-1.9 – 4.8)	1.4 (-1.9 – 4.7)	1.3 (-2 – 4.6)
Q3	0 (-3.3 – 3.3)	0.3 (-3.1 – 3.6)	0.2 (-3.1 – 3.5)
Q4, highest	0.8 (-2.6 – 4.1)	1.1 (-2.3 – 4.4)	0.8 (-2.5 – 4.2)
Migrants, by age at first migration**, n = 584/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	-0.7 (-3.2 – 1.8)	-0.7 (-3.2 – 1.7)	-0.7 (-3.2 – 1.7)

Note: All values are β coefficients (95% CI) in mm Hg units
Multivariable analyses done with censored normal regressions, which do not calculate an R^2 as with other linear regressions
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table VIII-3. Multivariable analyses of migrant sub-classifications and diastolic blood pressure

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	1.3 (-1.4 – 4.1)	1.3 (-1.5 – 4.1)	1.3 (-1.5 – 4.1)
30-39 years in urban area	2.8 (-0.1 – 5.8)	2.8 (-0.2 – 5.8)	2.8 (-0.2 – 5.7)
≥40 years in urban area	2 (-1.4 – 5.3)	1.9 (-1.5 – 5.3)	1.8 (-1.6 – 5.2)
Migrants, by lifetime exposure to urban area*, n = 558/559			
Q1, lowest	Reference	Reference	Reference
Q2	1.6 (-0.6 – 3.7)	1.6 (-0.6 – 3.7)	1.5 (-0.7 – 3.6)
Q3	1.7 (-0.4 – 3.9)	1.7 (-0.5 – 3.8)	1.6 (-0.5 – 3.8)
Q4, highest	0.8 (-1.4 – 2.9)	07(-1.4 – 2.9)	0.6 (-1.5 – 2.8)
Migrants, by age at first migration**, n = 584/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	-0.4 (-1.9 – 1.2)	-0.3 (-1.9 – 1.2)	-0.4 (-1.9 – 1.2)

Note: All values are β coefficients (95% CI) in mm Hg units
Multivariable analyses done with censored normal regressions, which do not calculate an R^2 as with other linear regressions
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

8.2. Hypertension

Hypertension was defined as SBP ≥ 140 mm Hg or DBP ≥ 90 mm Hg or currently receiving antihypertensive medication, as shown earlier in Chapter II and following standard European and US guidelines [155, 156].

Breakdown of hypertension defined by blood pressure readings or report of current specific antihypertensive treatment is provided in Table VIII-4. Prevalence of hypertension was 16.2% (160/988). Amongst hypertensives, 37/160 (23.1%) were on antihypertensive medication and 11/37 (29.7%) were controlled, i.e. had a BP within normal ranges.

Table VIII-4. Distribution of hypertension-related categories by study group

	Rural	Migrant	Urban	Total	Missing
	201	589	199		
Hypertension defined					
(A) by direct measurement	24/201	70/588	55/199	149/988	1/989
	11.9%	11.9%	27.6%	15.1%	0.1%
(B) currently on treatment*	0/201	7/588	4/199	11/988	
	0%	1.2%	2%	1.1%	
Hypertension (A + B)	24/201	78/588	59/199	160/988	1/989
	11.9%	13.3%	29.7%	16.2%	0.1%

* This group included only those hypertensives currently on medication and with blood pressure within normal ranges. They reported to be on antihypertensive medication and provided the specific name of the drug they were taking. In other denominations these individuals would be classified as "aware and controlled hypertensives".

Table VIII-5 presents the distribution and OR (95% CI) of hypertension in all study groups and according to migrant subgroups.

Hypertension was prevalent in 12%, 13% and 30% of rural, migrant and urban population, respectively. In the fully adjusted model compared to the rural group, migrant and urban people were 1.5 (95% CI 0.8 – 2.8) and 5 (95% CI 2.3 – 10.6) times more likely to be hypertensives than rural people, respectively. OR were similar in crude and fully adjusted models. By examining CI, this association was only significant for the comparison between urban versus rural populations, and it was not significant for the comparison between migrant and rural groups.

When exploring the association of hypertension within migrants, no clear associations between hypertension and number of years in an urban area, lifetime exposure to urban area or age at first migration exposures were observed.

Table VIII-5. Multivariable association of hypertension and migration by exposure groups

Exposure Groups	Hypertensives n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Rural, Migrant, Urban	**(n = 988/989)**			
Rural	24 (11.9%)	1	1	1
Migrant	77 (13.1%)	1.3 (0.8 – 2.2)	1.5 (0.8 – 2.8)	1.5 (0.8 – 2.9)
Urban	59 (29.7%)	4.2 (2.3 – 7.4)	5.1 (2.4 – 10.5)	5 (2.3 – 10.6)
Migrants, by years in urban area	**(n = 558/589)**			
<20 years in urban area	4 (7.6 %)	1	1	1
20-29 years in urban area	16 (7.9 %)	1.1 (0.3 – 3.7)	1.1 (0.3 – 3.7)	1 (0.3 – 3.6)
30-39 years in urban area	19 (11.2 %)	1.1 (0.3 – 3.6)	1.1 (0.3 – 3.7)	1.1 (0.3 – 3.6)
≥40 years in urban area	27 (20.3 %)	1.1 (0.3 – 3.7)	1.1 (0.3 – 3.8)	1.1 (0.3 – 3.7)
Migrants, by lifetime exposure to urban area*	**(n = 558/589)**			
Q1, lowest	14 (9.9 %)	1	1	1
Q2	17 (12.2 %)	1.8 (0.8 – 4.1)	1.8 (0.8 – 4.1)	1.7 (0.7 – 3.8)
Q3	13 (9.2 %)	1 (0.4 – 2.3)	1 (0.4 – 2.4)	1 (0.4 – 2.3)
Q4, highest	22 (16.1 %)	1.9 (0.9 – 4.2)	2 (0.9 – 4.4)	1.9 (0.9 – 4.2)

Table VIII-5. (continued)

Exposure Groups	Hypertensives n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Migrants, by age at first migration**	**(n = 584/589)**			
≤ 12 yo at first migration	28 (12.4 %)	1	1	1
> 12 yo at first migration	47 (13.1 %)	0.7 (0.4 – 1.3)	0.8 (0.4 – 1.3)	0.7 (0.4 – 1.3)

Notes: All values are OR (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

8.3. Lipid profile

8.3.1. Lipid profile in all study groups

Table VIII-6 presents the descriptive analyses of total cholesterol, HDL, LDL, triglycerides and total cholesterol / HDL (TC/HDL) ratio disaggregated by sex and followed by SMD in each of the variables. None of the participants reported to be on any specific treatment for lipids.

All three groups studied —rural, migrant and urban populations—had a mean total cholesterol below 200 mg/dL, mean HDL above 40 mg/dL and mean LDL below 129 mg/dL, the recommended thresholds in US guidelines [165]. TC/HDL ratio in all groups was also within recommended ranges, i.e. below 5.

Within groups and for all lipid markers evaluated, except HDL, a consistent pattern was shown. First, no gradient of progressive increase or decline from rural to migrant to urban was observed. Secondly, lipid marker's means or medians were very similar in migrant and urban groups. Thirdly, lipid markers in migrant and urban groups were higher when compared to the rural population. In the case of HDL, no difference was observed between rural, migrant or urban groups.

Compared to male group, rural and migrant females, but not urban females, had higher mean values of total cholesterol and LDL. In the case of HDL and TC/HDL ratio, it was migrant and urban females, but not rural females, the ones who had higher mean values compared to male group. In the case of triglycerides, a 20mg/dL difference by gender was observed in the urban group but not in the other groups.

In terms of SMD, compared to the rural group, no differences were observed for HDL and triglycerides in both migrant and urban populations. However, both migrants and urban groups had around 1 SD unit higher than the rural group for total cholesterol and LDL and around 0.6 SD units higher for TC/HDL cholesterol (Table VIII-6).

Table VIII-6. Descriptive analyses and standardised mean differences of lipid profile by study groups

	Rural	Migrant	Urban	
Descriptive statistics				**Missing data**
Total cholesterol (mg/dL)				
All, mean (SD)	155.7 (33.3)	190.8 (39.5)	194.9 (40)	1 (0.1%)
Female, mean (SD)	156.2 (33.7)	192.7 (41.5)	194.1 (39)	
Male, mean (SD)	155.2 (33)	188.6 (37)	195.8 (41.2)	
HDL (mg/dL)				
All, mean (SD)	44.1 (13.1)	44 (11.2)	44.4 (11)	1 (0.1%)
Female, mean (SD)	43 (12.2)	45.3 (11.4)	46.4 (11.1)	
Male, mean (SD)	45.2 (14.1)	42.5 (10.8)	42 (10.5)	
LDL (mg/dL)				
All, mean (SD)	85.6 (27.1)	115.9 (33)	119.8 (34.2)	1 (0.1%)
Female, mean (SD)	86.6 (26.3)	117.1 (34.8)	119.1 (33.8)	
Male, mean (SD)	84.5 (28.1)	114.5 (30.8)	120.6 (34.9)	
Triglycerides (mg/dL)				
All, median (IQR)	113 (71)	133 (95.5)	135 (109)	1 (0.1%)
Female, median (IQR)	113 (73)	132 (92)	125 (103)	
Male, median (IQR)	114 (68)	133 (100)	145.5 (103.5)	

Table VIII-6. (continued)

	Rural	Migrant	Urban	
TC/HDL ratio				
All, mean (SD)	3.8 (1.2)	4.6 (1.4)	4.7 (1.6)	1 (0.1%)
Female, mean (SD)	3.9 (1.4)	4.5 (1.4)	4.4 (1.4)	
Male, mean (SD)	3.6 (1)	4.7 (1.4)	5 (1.8)	
	Within groups SD	**SMD (95% CI)**	**SMD (95% CI)**	**R^2**
Total cholesterol	37.9	0.9 (0.7 – 1.1)	1 (0.7 – 1.2)	0.16
HDL	11.6	0 (-0.3 – 0.2)	0 (-0.3 – 0.3)	0.01
LDL	32	1 (0.8 – 1.2)	1 (0.8 – 1.3)	0.15
Triglycerides (log)	0.5	0.2 (0 – 0.4)	0.2 (-0.1 – 0.4)	0.04
TC/HDL ratio	1.4	0.6 (0.4 – 0.8)	0.7 (0.4 – 0.9)	0.07

Notes: All regressions calculated using the rural group as baseline (Total n = 988/989). SMD regression models included adjustment for age, sex, individual socioeconomic deprivation and highest parental education

In multivariable regression analyses (Table VIII-7), after adjustment for age, sex and socioeconomic confounders, there was no difference in HDL in both, migrant and urban groups compared to the rural group. On the contrary, both migrant and urban groups had significantly higher mean total cholesterol, LDL and TC/HDL ratio compared to the rural group. In average, for total cholesterol and LDL, migrant and urban populations had over 30 mg/dL units higher than their rural counterparts.

Triglycerides were log transformed for regression analyses and thus, its output was presented as geometric means (95% CI) for the baseline rural group and ratios (95% CI) for the comparison migrant and urban groups. Ratios require two decimal points for their interpretation as percentages relative to baseline group. In the same vein, for ratios to be significantly different their 95% CI should not overlap the value of one, which implies no difference from baseline [171, 172].

After adjustment for age, sex and socioeconomic confounders the rural group had a geometric mean of 116.2 mg/dL (95% CI 102.2 – 132.2). The geometric mean of the migrant group was equal to 112% (95% CI 101% – 125%) of the geometric mean of the rural group or, in other words, 12% (95% CI 1% – 25%) greater than the geometric mean of rural group. In this case, CI was near to value of 1 which places such comparison as borderline significant. Such significant difference was not observed for the comparison of triglycerides between urban and rural groups (Table VIII-7).

Table VIII-7. Multivariable analyses of lipid profile by study groups

	Rural	Migrant	Urban	
Multivariable Total cholesterol		**β coefficient (95%CI)**	**β coefficient (95%CI)**	**R^2**
Model 1	Reference	35.4 (29.3 – 41.5)	39.2 (31.8 – 46.7)	0.15
Model 2	Reference	34.8 (26.9 – 42.7)	38.6 (29.1 – 48)	0.15
Model 3	Reference	34.9 (27 – 42.8)	36.2 (26.4 – 46)	0.15
Multivariable HDL				
Model 1	Reference	-0.1 (-2 – 1.8)	0.3 (-2 – 2.5)	0.01
Model 2	Reference	-0.5 (-2.9 – 1.9)	-0.2 (-3.1 – 2.7)	0.01
Model 3	Reference	-0.5 (-2.9 – 1.9)	-0.1 (-3.1 – 2.9)	0.01
Multivariable LDL				
Model 1	Reference	30.5 (25.4 – 35.6)	34.3 (28 – 40.5)	0.15
Model 2	Reference	31.1 (24.4 – 37.7)	35 (27 – 42.9)	0.15
Model 3	Reference	31.1 (24.5 – 37.8)	32.9 (24.7 – 41.2)	0.15
Multivariable TC/HDL ratio				
Model 1	Reference	0.8 (0.6 – 1.1)	0.9 (0.7 – 1.2)	0.07
Model 2	Reference	0.9 (0.6 – 1.2)	1 (0.7 – 1.4)	0.07
Model 3	Reference	0.9 (0.6 – 1.2)	0.9 (0.6 – 1.3)	0.07

Table VIII-7. (continued)

	Rural	Migrant	Urban	
Multivariable Triglycerides (log)	**Geometric mean (95% CI)**	**Ratio (95% CI)**	**Ratio (95% CI)**	**R^2**
Model 1	114 (105.4 – 123.3)	1.15 (1.06 – 1.25)	1.15 (1.04 – 1.27)	0.04
Model 2	118.1 (104.8 – 133)	1.12 (1.01 – 1.25)	1.11 (0.98 – 1.27)	0.04
Model 3	116.2 (102.2 – 132.2)	1.12 (1.01 – 1.25)	1.09 (0.96 – 1.25)	0.04

Notes: All regressions calculated using the rural group as baseline (Total n = 988/989).
Model 1: Adjusted for age and sex; Model 2: As model 1 plus individual socioeconomic deprivation; Model 3: As model 2 plus highest parental education

8.3.2. Lipid profile in migrant subgroups

Sub-classifications by length of residence in urban area, lifetime exposure to an urban area and age at first migration were analysed in the migrant population only. Table VIII-8 to Table VIII-12 (pages 238 to 242) show β regression coefficients (95% CI) or geometric means / ratios (95% CI) of comparisons between various categories used within migrants and total cholesterol, HDL, LDL, triglycerides and TC/HDL ratio.

A significant difference was observed for total cholesterol (Table VIII-8, page 238) and LDL (Table VIII-10, page 240) by number of years living in an urban area. Migrants with 20 or more years living in an urban area had in average 15 mg/dL units of total cholesterol and 10 mg/dL units of LDL higher than those with less than 20 years in urban area.

No clear pattern of difference was observed using other migrant sub-classifications or in other lipid markers.

Table VIII-8. Multivariable analyses of migrant sub-classifications and total cholesterol

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	13.4 (1.6 – 25.1)	13.4 (1.6 – 25.2)	13.5 (1.7 – 25.4)
30-39 years in urban area	16.6 (4.1 – 29.1)	16.8 (4.2 – 29.4)	16.9 (4.3 – 29.5)
≥40 years in urban area	16.3 (2.1 – 30.4)	16.5 (2.2 – 30.8)	16.4 (2.1 – 30.8)
R^2	0.06	0.06	0.06
Migrants, by lifetime exposure to urban area*, n = 558/559			
Q1, lowest	Reference	Reference	Reference
Q2	7.2 (-1.9 – 16.4)	7.2 (-1.9 – 16.4)	7.4 (-1.8 – 16.6)
Q3	5.8 (-3.4 – 14.9)	5.8 (-3.5 – 15)	5.7 (-3.5 – 15)
Q4, highest	3.9 (-5.3 – 13.1)	3.9 (-5.4 – 13.1)	3.8 (-5.5 – 13.1)
R^2	0.05	0.05	0.05
Migrants, by age at first migration**, n = 584/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	5.7 (-0.9 – 12.4)	5.7 (-0.9 – 12.4)	6 (-0.7 – 12.7)
R^2	0.05	0.05	0.05

Note: All values, except R^2 values, are β coefficients (95% CI) in mg/dL units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table VIII-9. Multivariable analyses of migrant sub-classifications and high-density lipoprotein cholesterol

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	1.8 (-1.6 – 5.2)	1.8 (-1.6 – 5.1)	1.7 (-1.7 – 5.1)
30-39 years in urban area	-1.5 (-5.1 – 2.1)	-1.7 (-5.3 – 1.9)	-1.7 (-5.3 – 1.9)
≥40 years in urban area	-0.1 (-4.2 – 4)	-0.3 (-4.4 – 3.8)	-0.3 (-4.4 – 3.8)
R^2	0.04	0.04	0.04
Migrants, by lifetime exposure to urban area*, n = 558/559			
Q1, lowest	Reference	Reference	Reference
Q2	-0.4 (-3 – 2.2)	-0.4 (-3 – 2.2)	-0.4 (-3.1 – 2.2)
Q3	-0.6 (-3.2 – 2)	-0.7 (-3.4 – 1.9)	-0.7 (-3.4 – 1.9)
Q4, highest	-0.8 (-3.5 – 1.8)	-0.9 (-3.6 – 1.8)	-0.9 (-3.6 – 1.7)
R^2	0.03	0.03	0.03
Migrants, by age at first migration**, n = 584/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	0.5 (-1.4 – 2.4)	0.5 (-1.4 – 2.4)	0.5 (-1.4 – 2.4)
R^2	0.02	0.02	0.02

Note: All values, except R^2 values, are β coefficients (95% CI) in mg/dL units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table VIII-10. Multivariable analyses of migrant sub-classifications and low-density lipoprotein cholesterol

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	9.6 (-0.3 – 19.4)	9.8 (0 – 19.7)	10 (0.1 – 19.8)
30-39 years in urban area	11 (0.5 – 21.4)	11.6 (1.1 – 22.2)	11.7 (1.2 – 22.3)
≥40 years in urban area	12.6 (0.7 – 24.4)	13.4 (1.4 – 25.3)	13.3 (1.4 – 25.3)
R^2	0.05	0.05	0.05
Migrants, by lifetime exposure to urban area*, n = 558/559			
Q1, lowest	Reference	Reference	Reference
Q2	4.4 (-3.2 – 12.1)	4.4 (-3.3 – 12)	4.5 (-3.1 – 12.2)
Q3	4.1 (-3.6 – 11.7)	4.5 (-3.2 – 12.2)	4.5 (-3.2 – 12.2)
Q4, highest	3.1 (-4.6 – 10.8)	3.5 (-4.2 – 11.2)	3.4 (-4.3 – 11.2)
R^2	0.04	0.04	0.04
Migrants, by age at first migration**, n = 584/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	3.8 (-1.8 – 9.4)	3.8 (-1.8 – 9.4)	4.1 (-1.5 – 9.7)
R^2	0.04	0.04	0.04

Note: All values, except R^2 values, are β coefficients (95% CI) in mg/dL units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table VIII-11. Multivariable analyses of migrant sub-classifications and triglycerides

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
Baseline			
<20 years in urban area	119.3 (102.5 – 138.7)	121.5 (103.9 – 142.1)	121.4 (102.8 – 143.4)
Ratios			
20-29 years in urban area	1.02 (0.87 – 1.20)	1.02 (0.87 – 1.20)	1.02 (0.87 – 1.20)
30-39 years in urban area	1.24 (1.04 – 1.47)	1.23 (1.03 – 1.46)	1.23 (1.03 – 1.46)
≥40 years in urban area	1.13 (0.93 – 1.37)	1.12 (0.92 – 1.36)	1.12 (0.92 – 1.36)
R^2	0.99	0.99	0.99
Migrants, by lifetime exposure to urban area*, n = 558/559			
Baseline			
Q1, lowest	124 (112.5 – 136.7)	126.8 (114.3 – 140.6)	126.8 (112.9 – 142.4)
Ratios			
Q2	1.10 (0.97 – 1.25)	1.11 (0.98 – 1.25)	1.11 (0.98 – 1.26)
Q3	1.05 (0.93 – 1.12)	1.04 (0.92 – 1.18)	1.04 (0.92 – 1.18)
Q4, highest	1.07 (0.94 – 1.21)	1.06 (0.93 – 1.20)	1.06 (0.93 – 1.20)
R^2	0.99	0.99	0.99
Migrants, by age at first migration**, n = 584/585			
Baseline			
≤ 12 yo at first migration	128.9 (118.7 – 139.9)	130.5 (119.7 – 142.2)	131.4 (118.5 – 145.8)
Ratios			
> 12 yo at first migration	1.04 (0.95 – 1.14)	1.04 (0.95 – 1.14)	1.04 (0.95 – 1.14)
R^2	0.99	0.99	0.99

Note: All baseline values, except ratios and R^2 values, are geometric means (95% CI) in mg/dL units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table VIII-12. Multivariable analyses of migrant sub-classifications and total cholesterol / high-density lipoprotein cholesterol ratio

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
<20 years in urban area	Reference	Reference	Reference
20-29 years in urban area	0.2 (-0.2 – 0.7)	0.2 (-0.2 – 0.7)	0.3 (-0.2 – 0.7)
30-39 years in urban area	0.6 (0.2 – 1)	0.6 (0.2 – 1.1)	0.6 (0.2 – 1.1)
≥40 years in urban area	0.5 (0 – 1)	0.5 (0 – 1)	0.5 (0 – 1)
R^2	0.06	0.06	0.06
Migrants, by lifetime exposure to urban area*, n = 558/559			
Q1, lowest	Reference	Reference	Reference
Q2	0.3 (0 – 0.6)	0.3 (0 – 0.6)	0.3 (0 – 0.6)
Q3	0.2 (-0.2 – 0.5)	0.2 (-0.1 – 0.5)	0.2 (-0.1 – 0.5)
Q4, highest	0.2 (-0.2 – 0.5)	0.2 (-0.2 – 0.5)	0.2 (-0.2 – 0.5)
R^2	0.05	0.05	0.05
Migrants, by age at first migration**, n = 584/585			
≤ 12 yo at first migration	Reference	Reference	Reference
> 12 yo at first migration	0.1 (-0.1 – 0.3)	0.1 (-0.1 – 0.3)	0.1 (-0.1 – 0.4)
R^2	0.04	0.04	0.04

Note: All values, except R^2 values, are β coefficients (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

8.4. Inflammation markers

8.4.1. CRP and Fibrinogen in all study groups

Both, median CRP and median fibrinogen were similar between migrant and urban groups but higher when compared to rural one, particularly in the case of CRP (Table VIII-13). Compared to males, CRP and fibrinogen levels were higher amongst females except in the case of CRP in rural females.

After multivariate adjustment and compared to the rural group which had a CRP geometric mean of 0.6 mg/L (95% CI 0.4 – 0.8), both migrant and urban populations had significantly higher geometric means of CRP. The CRP geometric means of migrant and urban groups were 196% (95% CI 120% – 298%) and 224% (95% CI 124% – 367%) greater than the geometric mean of the rural group, respectively. In the case of fibrinogen, the geometric mean of the migrant group was 11% (95% CI 6% – 16%) greater than the geometric mean of rural group. A similar geometric mean ratio, 11% higher (95% CI 1.05 – 1.17), was observed in the rural-urban comparison of fibrinogen.

In terms of SMD, compared to the rural group, both migrants and urban groups showed similar significant differences for CRP and fibrinogen (Table VIII-13).

Table VIII-13. Descriptive and multivariable analyses of inflammation markers by study groups

	Rural	Migrant	Urban	
Descriptive statistics	**Median (IQR)**	**Median (IQR)**	**Median (IQR)**	**Missing data**
CRP (mg/L)				
All	0.7 (1.6)	1.6 (2.7)	1.6 (2.7)	1 (0.1%)
Female	0.6 (1.4)	1.9 (3.1)	1.9 (2.8)	
Male	0.8 (1.6)	1.2 (1.9)	1.5 (2.4)	
Fibrinogen (mg/dL)				
All	351.4 (115.7)	386.9 (96.5)	383 (77.5)	1 (0.1%)
Female	376.5 (112)	401.7 (76)	397.9 (80)	
Male	340 (113.3)	366.7 (124.4)	365.1 (95.4)	
Multivariable CRP (log)	**Geometric mean (95% CI)**	**Ratio (95% CI)**	**Ratio (95% CI)**	**R^2**
Model 1	0.6 (0.5 – 0.7)	2.79 (2.22 – 3.50)	3.10 (2.34 – 4.10)	0.13
Model 2	0.5 (0.4 – 0.7)	3.02 (2.24 – 4.06)	3.40 (2.38 – 4.85)	0.13
Model 3	0.6 (0.4 – 0.8)	2.96 (2.20 – 3.98)	3.24 (2.24 – 4.67)	0.14
Multivariable Fibrinogen (log)				
Model 1	351.4 (340.1 – 363)	1.10 (1.06 – 1.14)	1.10 (1.06 – 1.15)	0.19
Model 2	346.5 (329.7 -364.1)	1.11 (1.06 – 1.16)	1.12 (1.06 – 1.18)	0.19
Model 3	349.9 (331.7 – 369.1)	1.11 (1.06 – 1.16)	1.11 (1.05 – 1.17)	0.20

Table VIII-13. (continued)

	Rural	Migrant	Urban	
	Within groups SD	**SMD (95% CI)**	**SMD (95% CI)**	**R^2**
SMD CRP (log)	1.4	0.8 (0.6 – 1)	0.8 (0.6 – 1.1)	0.14
SMD Fibrinogen (log)	0.2	0.5 (0.3 – 0.7)	0.5 (0.2 – 0.7)	0.20

Notes: All regressions calculated using the rural group as baseline (Total n = 988/989).
Model 1: Adjusted for age and sex; Model 2: As model 1 plus individual socioeconomic deprivation; Model 3: As model 2 plus highest parental education (this final model was also used for SMD regressions)

8.4.2. CRP and Fibrinogen in migrant subgroups

As with other outcomes, sub-classifications by length of residence in urban area, lifetime exposure to an urban area and age at first migration were analysed in the migrant population only. Table VIII-14 and Table VIII-15 show geometric means / ratios (95% CI) of comparisons between various categories used within migrants and CRP and fibrinogen.

No clear pattern of difference was observed using various migrant sub-classifications in inflammation markers. All CI for geometric mean ratios in crude and adjusted multivariable models did overlap the value 1 in both, CRP and fibrinogen.

Table VIII-14. Multivariable analyses of migrant sub-classifications and C-reactive protein

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
Baseline			
<20 years in urban area	1.9 (1.3 – 2.8)	1.9 (1.3 – 2.8)	2 (1.3 – 3)
Ratios			
20-29 years in urban area	0.83 (0.55 – 1.24)	0.83 (0.55 – 1.24)	0.83 (0.55 – 1.24)
30-39 years in urban area	0.91 (0.59 – 1.40)	0.91 (0.59 – 1.40)	0.91 (0.59 – 1.40)
≥40 years in urban area	0.95 (0.59 – 1.55)	0.95 (0.58 – 1.55)	0.96 (0.59 – 1.57)
R^2	0.13	0.13	0.14
Migrants, by lifetime exposure to urban area*, n = 558/559			
Baseline			
Q1, lowest	1.8 (1.4 – 2.3)	1.8 (1.4 – 2.4)	1.9 (1.4 – 2.6)
Ratios			
Q2	0.93 (0.68 – 1.27)	0.93 (0.68 – 1.27)	0.94 (0.69 – 1.29)
Q3	0.91 (0.67 – 1.25)	0.91 (0.67 – 1.25)	0.92 (0.67 – 1.25)
Q4, highest	0.92 (0.67 – 1.25)	0.92 (0.68 – 1.26)	0.93 (0.68 – 1.27)
R^2	0.13	0.13	0.13
Migrants, by age at first migration**, n = 584/585			
Baseline			
≤ 12 yo at first migration	1.7 (1.4 – 2)	1.7 (1.4 – 2.1)	1.8 (1.4 – 2.3)
Ratios			
> 12 yo at first migration	1.05 (0.84 – 1.31)	1.05 (0.84 – 1.32)	1.05 (0.84 – 1.31)
R^2	0.14	0.14	0.14

Note: All baseline values, except ratios and R^2 values, are geometric means (95% CI) in mg/L units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table VIII-15. Multivariable analyses of migrant sub-classifications and fibrinogen

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
Baseline			
<20 years in urban area	380 (356.5 – 405)	378.9 (354.6 – 404.8)	380.1 (354.4 – 407.8)
Ratios			
20-29 years in urban area	0.98 (0.92 – 1.05)	0.98 (0.92 – 1.05)	0.98 (0.92 – 1.05)
30-39 years in urban area	1.02 (0.95 – 1.10)	1.02 (0.95 – 1.10)	1.03 (0.95 – 1.10)
≥40 years in urban area	1.06 (0.97 – 1.15)	1.06 (0.98 – 1.15)	1.06 (0.98 – 1.15)
R^2	0.99	0.99	0.99
Migrants, by lifetime exposure to urban area*, n = 558/559			
Baseline			
Q1, lowest	383.1 (367.5 – 399.1)	382.3 (366.1 – 399.3)	383.6 (365.4 – 402.8)
Ratios			
Q2	0.98 (0.93 – 1.03)	0.98 (0.92 – 1.03)	0.98 (0.93 – 1.03)
Q3	1.00 (0.95 – 1.06)	1.00 (0.95 – 1.06)	1.00 (0.95 – 1.06)
Q4, highest	1.03 (0.98 – 1.09)	1.03 (0.98 – 1.09)	1.03 (0.98 – 1.09)
R^2	0.99	0.99	0.99
Migrants, by age at first migration**, n = 584/585			
Baseline			
≤ 12 yo at first migration	389.2 (376.3 – 402.6)	388.9 (375.4 - 403)	389.5 (373.2 – 406.5)
Ratios			
> 12 yo at first migration	0.98 (0.94 – 1.02)	0.98 (0.94 – 1.02)	0.98 (0.94 – 1.02)
R^2	0.99	0.99	0.99

Note: All baseline values, except ratios and R^2 values, are geometric means (95% CI) in mg/dL units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

8.5. SMD in blood pressure, lipid profile and inflammation markers

Figure VIII-1, Figure VIII-2 and Figure VIII-3 show the SMD comparisons against rural group of blood pressure, lipid profile and inflammation markers, respectively.

Compared to rural people, the "size" of difference observed in cardiovascular risk factors presented in this chapter were mixed.

Blood pressure did not show a clear gradient of difference as other risk factors presented before. The migrant group had similar SBP but lower DBP than rural group. The urban group had higher SBP but similar DBP than rural group. In the case of lipid profile, no difference was observed between groups for HDL and triglycerides. Total cholesterol, LDL and TC/HDL ratio did, however, show important size of differences (> 0.5 SD) between groups, as did both inflammation markers studied.

Figure VIII-1. Standardised mean differences in systolic and diastolic blood pressure in migrant and urban population compared to rural population

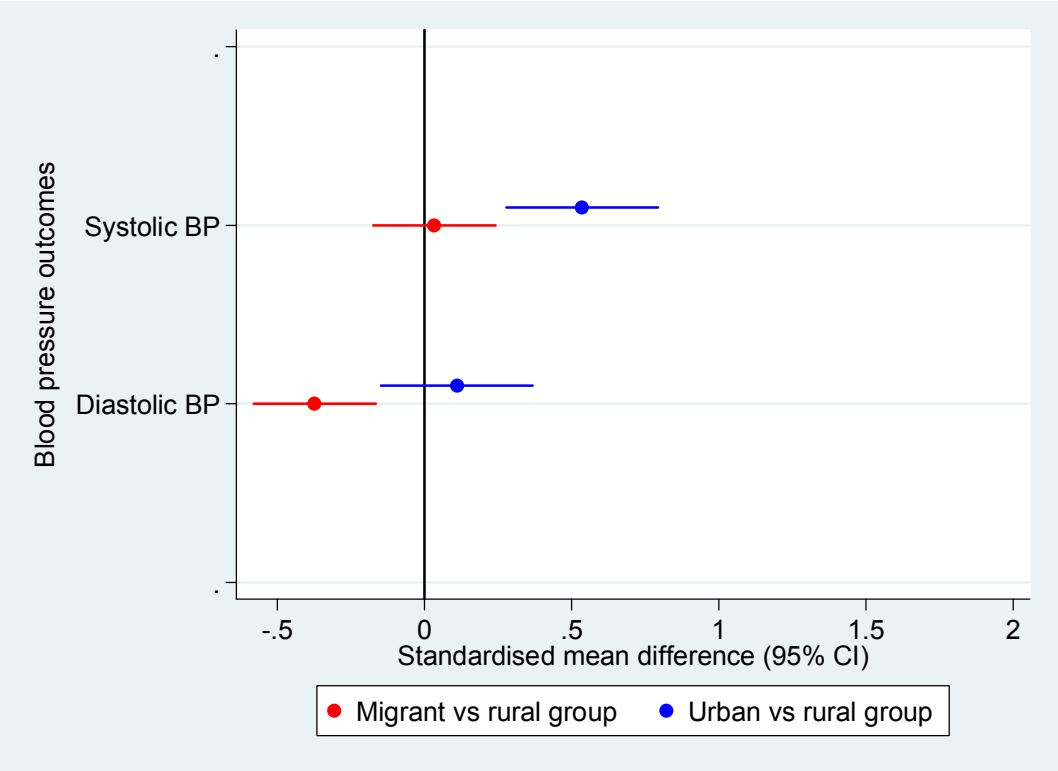


Figure VIII-2. Standardised mean differences in lipid profile in migrant and urban population compared to rural population

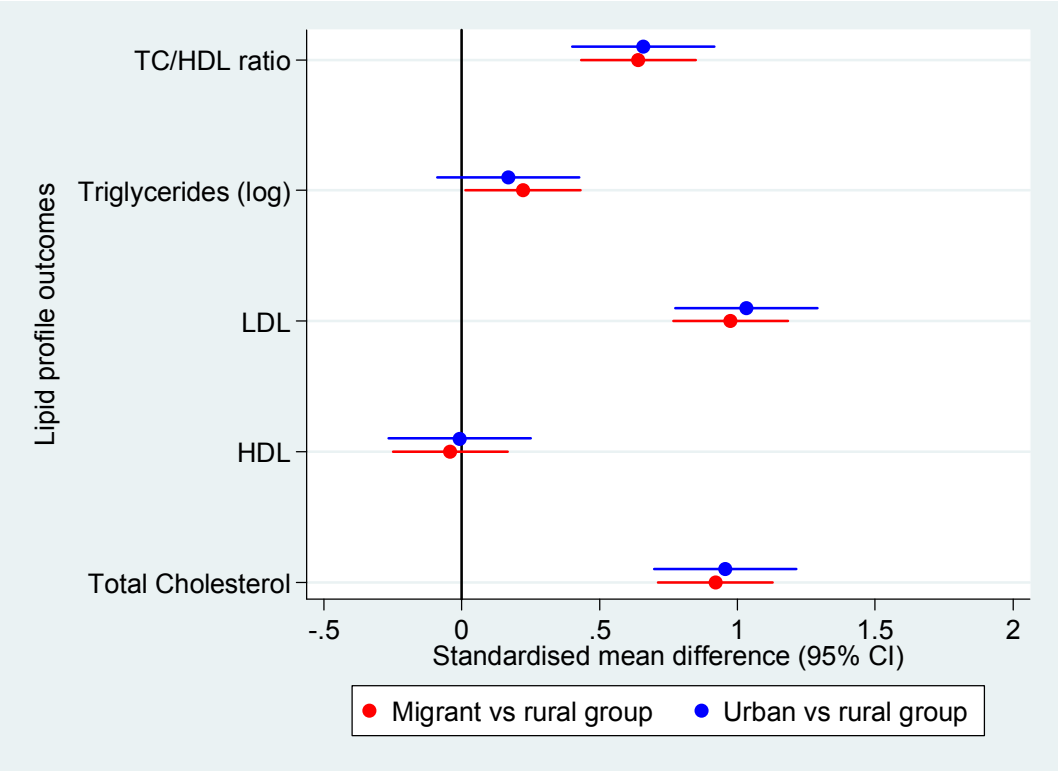
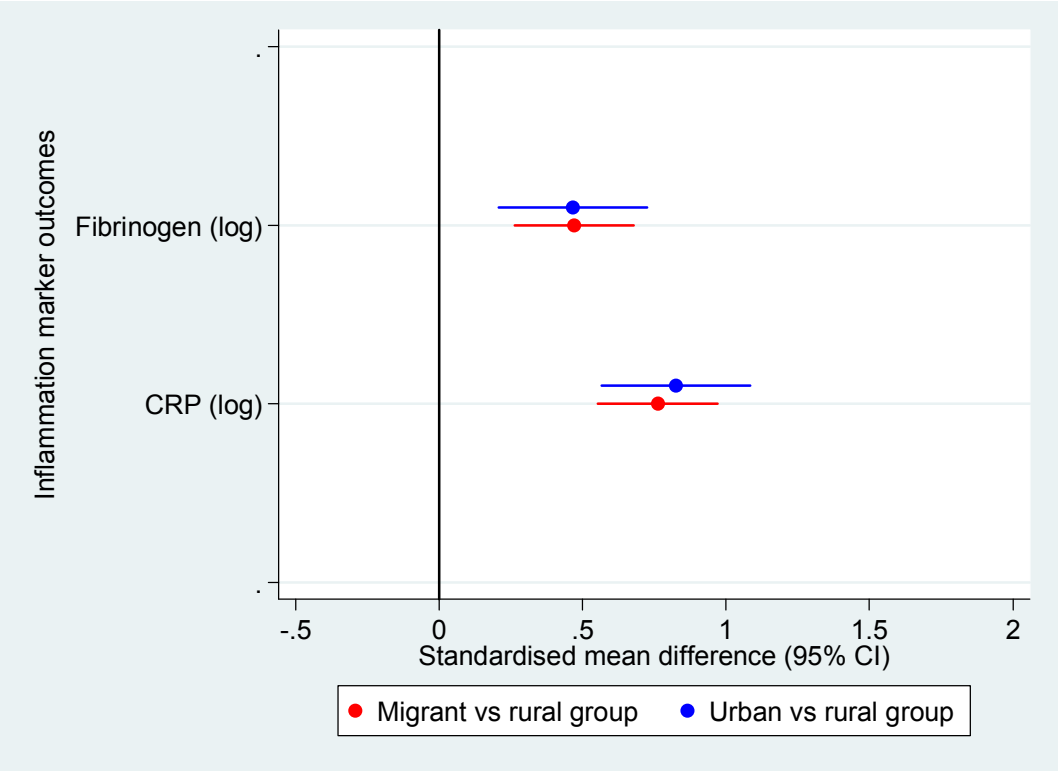


Figure VIII-3. Standardised mean differences in inflammation markers in migrant and urban population compared to rural population



8.6. Discussion of results

8.6.1. Summary

Blood pressure did not show a clear gradient of difference as other risk factors presented before. The migrant group had similar SBP but lower DBP than rural group. The urban group had higher SBP but similar DBP than rural group. Compared to the rural group, migrant and urban people were 1.5 (95% CI 0.8 – 2.9) and 5 (95% CI 2.3 – 10.6) times more likely to be hypertensives than rural people, respectively. In the case of lipid profile, no difference was observed between groups for HDL and triglycerides. Total cholesterol, LDL and TC/HDL as well as CRP and fibrinogen were much higher in migrant and urban population compared to rural group.

8.6.2. Strengths

As with anthropometric outcomes, one of the major strengths of this section is that it reports well established risk factors for CVD, including blood pressure, lipid profile and inflammation markers, describing also its differences between groups studied.

The findings of marked differences in some instances such as CRP, non-uniform differences as in BP or the absence of difference as in HDL or triglycerides, provides further insights to the description of the complexity of CVD risk profile in low- and middle-income countries (LMIC).

Another strength that deserves consideration was the diagnosis criteria established for hypertension. The risk of overestimation of self-reported diagnosis of hypertension in this population was high considering Peru's fragmented health system, the lack of continuous follow-up and long-term medication. Thus, during the interview, three questions were asked to be able to identify those with an established diagnosis and with ongoing antihypertensive medication. First, participants were asked about their medical history of hypertension, specifically to report if he/she had been previously diagnosed with hypertension, if they were currently on treatment and to name the specific antihypertensive medication they were currently taking. Those

who reported to be on antihypertensive medication and provided the specific name of the drug they were taking were considered as "currently on treatment". This was considered a far more stringent criterion than physician —or any other health professional— diagnosis of hypertension, ever since the diagnosis of hypertension requires repeated measurements in separate visits. In such circumstances and without the availability of medical records to ascertain such diagnosis, the strong possibility remained that people were told of such diagnosis at single random visits, either at private practices or pharmacies, which are common health seeking behaviours in Peru. By following this established criterion for hypertension, only one person was left out of the final analyses (the subject, a migrant, reported to be hypertensive, to have been diagnosed by a physician but did not provide the name of their antihypertensive medication).

The evaluation of a complete set of lipid profile studied and the up-to-date evaluation of important inflammation markers such as CRP and fibrinogen constitute an additional strength of this study. Such measurements were made in fasting conditions and also permitted the calculation of TC/HDL ratio. As recommended, current guidelines for coronary heart disease (CHD) prevention [165, 235] emphasize the use of total cholesterol and LDL in CHD risk assessment.

8.6.3. Weaknesses

Apo B:Apo A-I ratio has been reported as a better predictor of CHD [236] and, following results derived from the INTERHEART study [237, 238] which indicated Apo B:Apo A-I ratio as a strong predictor of myocardial infarction, it could be argued that such ratio should have be measured as part of lipid profile. However, such assertion is contentious and alternative evidence from other population-based studied, including Framingham, suggest otherwise. Ingelsson et al. report that "the performance of the Apo B:Apo A-I ratio for CHD risk prediction was comparable with that of other lipid ratios with respect to model discrimination, calibration, and reclassification characteristics in both sexes. Furthermore, Apo B:Apo A-I does not provide incremental value for CHD risk prediction over established risk factors, including total cholesterol:HDL-C" [239].

Considering the evidence presented above and taking into account the increased costs associated with the measurement of Apo A-I and Apo B, it was decided that the lipid profile evaluated in the study was solid enough to provide risk information in the groups studied.

As stated as a limitation in the previous chapter related to anthropometric outcomes, the analysis plan set out for this study did not address in depth potential differences that may exist in the profile of risk factors in terms of sex. In most cases, baseline means and SD (or medians and IQR) were presented disaggregated by sex in the descriptive tables. However, it was decided that the main question of this study was to address if there was any difference in CVD risk profile between study groups, and more specifically if those differences vary by patterns of migration. In the same vein, when maintaining a uniform approach to the analysis, further adjustment beyond socioeconomic indicators was not pursued, for example in the case of BP. In this later instance, it is acknowledged that BMI is associated with BP [240, 241] and also high altitude has been reported to be associated with BP [86, 200]. While further adjustment for this variable would be of interest, and will be explored in separate analyses at a later stage, it detracts from the main research question of this study which is to address differences on various CVD risk factors following migration.

8.6.4. Blood pressure

Surprisingly, SBP of migrants was similar to and DBP was lower than in rural counterparts. In the case of urban group, their SBP was higher than in rural group, but DBP was not different because of its overlapping CI. This raises two potential circumstances that could explain such findings. First, the possibility that BP in the rural baseline group did not behave as a "healthier" profile, in a population left-shifted curve, similar to anthropometric indicators presented in the previous chapter or to inflammation markers presented in this chapter. Related to this point is that further adjustment for known confounding factors that exert an effect on BP were not applied, particularly BMI, smoking status and tobacco consumption. Secondly, on the contrary, this study might also well raise the possibility that following migration certain protective mechanisms take place that enables a better profile, at least in terms of BP, in the migrant group. This study is neither capable of evaluating the

temporality of events nor causality needed to address the latter scenario because of its specific study design-related limitation, a cross-sectional study. However, by looking at some published results we could explore the first point, whether or not the BP findings on the rural population fall alongside a healthier or not-so-healthy profile, which thus will serve as the basis for comparisons with the migrant and urban groups of this study.

Of note, crude mean SBP in all three groups ranged from 120 – 128 mm Hg. SBP after adjustment for age, sex and socioeconomic variables was 113.1 mm Hg (95% CI 109.1 – 117.1) in rural baseline group, and β coefficients (95% CI) for migrant and urban groups were 0.7 mm Hg (95% CI -2.7 to 4.2) higher and 9.1 mm Hg (95% CI 4.8 – 13.3) higher than in rural group, respectively (see Table VIII-1).

8.6.4.1. Comparison with previous studies

Observed crude SBP mean values in this study differ from others reported in other national or international studies. For example, the ENINBSC Survey carried out by the Peruvian National Institute of Health between 2004-2005 (Table I-1) [108] found that SBP for rural Sierra and Lima were 112.3 mm Hg (95% CI 107.7 – 116.9) and 115.1 mm Hg (95% CI 112.7 – 117.6), respectively. Such values for BP, however, represent single measurements compared to the average of the last two out of three measurements used in this study and do not report adjustment for age, sex or other variables. The PREVENCION study, another well-conducted population-based study in a different but urban Andean city of Peru found a similar mean SBP to the rural and migrant groups of this study, with an unadjusted mean SBP of 121 and 120 mm Hg for males and females, respectively [112]. The literature within Peru suggests that SBP results obtained in the rural and migrant populations of this study do not differ greatly.

Nonetheless, the international literature shows lower, similar but also higher crude unadjusted means of SBP. For example, SBP in rural and migrant Chinese in the Yi People Study was 105 and 112 mm Hg, respectively [84-86]. Remote indigenous populations from Brazil, Papua New Guinea and Kenya that took part on the INTERSALT study had a combined SBP, average of four centres, of 103 mm Hg

compared to 120 mm Hg in the remaining 48 INTERSALT centres [242, 243]. On the contrary, the Kenyan Luo migration study found a mean SBP of 121 and 112 mm Hg in rural males and females, contrasted with 129 and 119 mm Hg in urban males and females, respectively [78, 244, 245]. In Cameroon, mean SBP was 125 and 119 mm Hg for rural male and female groups, and 136 and 127 mm Hg for urban males and females, respectively [83]. A rural-to-urban migration study in Iran found much higher values of mean SBP with figures of 126 mm Hg for rural people contrasted with 138 mm Hg in migrants from Azerbaijan and urban non-migrants [80]. Although it is difficult to establish the comparability between rural and migrant populations of cited examples with the respective groups of this study except for their common condition of being born in a rural place in their country of origin, the common pattern is that despite the location of study all rural populations have a lower SBP than migrants or urban counterparts. Of interest to the question posed above, it is clear that the range of rural baseline mean SBP varies widely in each example, having the Chinese rural population and other indigenous groups in a much "healthier", left-shifted population mean than other groups.

The observation that migrants from a wide age range, after a sustained process of migration and establishment into an urban environment for a number of years, have similar SBP to their rural counterparts, to the best of our knowledge, has not been previously reported. However, supporting evidence of surprising decrease in BP following migration in a much younger cohort and within 6 months of migration have recently been reported in Tanzania [82].

In the case of DBP, the results from this study do not differ much from Peru's ENINBSC Survey [108] which reports mean DBP of 72 mm Hg (95% CI 70 – 74) for rural Sierra and 70 mm Hg (95% CI 69 – 72) for urban Lima, but the PREVENCION study reports a slightly higher mean DBP of 79 and 77 mm Hg for males and females, respectively [112]. The Chinese Yi Migrant Study reports a mean DBP of 66 mm Hg for rural people and 71 mm Hg for migrants and urban groups [84]. The INTERSALT study reports a mean DBP of 57-67 mm Hg in four remote areas compared to an average of 74 mm Hg in the 48 remaining sites [242]. The Kenyan Luo migration study reports a mean DBP of 60 and 62 mm Hg in the rural and urban groups, respectively [78]. On the contrary, the Iran rural-to-urban

migration report a higher mean DBP of 72 versus 84 mm Hg in rural and migrant population, respectively [80]. In Cameroon, even higher mean DBP are reported for both rural (female/male 76/81 mm Hg) and urban dwellers (female/male 80/86 mm Hg) [83].

8.6.4.2. Remarks on blood pressure

The comparison with previous published studies indicates that BP profile of the rural group in this study is not so unhealthy. Indeed, rural group's BP falls within a left-shifted population curve for blood pressure and can serve as a good "healthy" baseline group for the comparisons of interest, similar to other risk factors already explored in this study. Thus, a surprising finding is that the migrant group also shows a healthy BP profile, as good as the rural group in the case of SBP or even a better profile in the case of DBP.

Similar findings of surprising decrease in BP following migration in a much younger cohort and within 6 months of migration have recently been reported in Tanzania [82]. As suggested by Unwin et al. in relation to these findings on BP, this "suggest that the pattern of change on rural to urban migration may be more complex than commonly thought and is worthy of further study" [82].

8.6.5. Hypertension

Hypertension was prevalent in 12%, 13% and 30% of rural, migrant and urban population, respectively. Such prevalences fit within the wide range of reported hypertension in Peru for various contexts (see Table I-2) with a clear pattern of lower hypertension rates in rural areas [109, 117, 118]. Again, surprisingly, migrants have a similar rate of hypertension to rural counterparts despite their considerable exposition to an urban environment. Migrants appear to have increased odds of being hypertensive than rural population but CI overlap the value 1 (OR 1.5 (95% CI 0.8 – 2.9)). The urban group, in contrast, was 5 (95% CI 2.3 – 10.6) times more likely to be hypertensive than the rural group. As with other risk factors, it would have been expected to find a gradient in this condition. This finding adds to the observations noted above on BP, that rural people are not necessarily a not-so-healthy group to be

compared against, but indeed, it is the migrant population described in this study that behaves in a very healthy, left-shifted population curved, in terms of BP and hypertension.

8.6.6. Lipid profile

In relation to lipid profile, the very first observation that yields from this study is that, overall, lipid markers in all three groups studied fits into a healthy lipid profile when compared to Western populations in developed countries of North America and Europe. All groups had a desirable level (below 200 mg/dL) for mean total cholesterol, above 40 mg/dL for HDL and optimal (below 100 mg/dL) or near or above optimal (100-129 mg/dL) for LDL [165]. European guidelines suggest that "in general, total plasma cholesterol should be below 5mmol/L (190mg/dL), and LDL cholesterol should be below 3mmol/L (115mg/dL)" [235]. Following European cut-offs, the rural and migrant groups population means are the ones within ranges proposed by European guidelines. The US National Health and Nutrition Examination Survey 2003-2004 reported the following overall mean values for lipid markers: total cholesterol 200.3 mg/dL, LDL 118.7 mg/dL, HDL 54.3 mg/dL, and triglycerides 129.5 mg/dL [246]. In contrast with these US results, the results from this study indicate that rural, migrant and urban populations have not only lower mean total cholesterol levels but also lower mean HDL levels. Migrant and urban but not rural people's mean LDL (116, 120 and 86 mg/dL, respectively) were similar to US levels (119 mg/dL). Interestingly, migrant and urban groups but not rural showed a higher mean triglyceride level (around 133 mg/dL) compared to the US population (129.5 mg/dL). Chinese individuals have also been described as a low lipid profile population [247]. The Yi People Study, showed better profile than this study for total cholesterol, HDL and LDL but not for triglycerides (range 136 – 152 mg/dL) [84].

Of interest in this study, in terms of SMD and comparing against the rural group, no differences were observed for HDL and triglycerides in both migrant and urban populations. However, both migrants and urban groups had around 1 SD unit higher than the rural group for total cholesterol and LDL and around 0.6 SD units higher for TC/HDL cholesterol (Table VIII-6). An explanation for such difference in total cholesterol and LDL but not in HDL can rely on the impact of diet, which was not

measured in this study. It is well established that high cholesterol intake significantly influences serum cholesterol levels but not HDL [248-256]. Such effects of diet on lipid profile may advocate the evaluation of dietary patterns in the present study aiming to have a better understanding and profiling of CVD risk factors. However, the evidence on the contribution of fat diet to CVD is inconclusive. As Hooper concludes, on a systematic review on the topic, "despite decades of effort and many thousands of people randomised, there is still only limited and inconclusive evidence of the effects of modification of total, saturated, monounsaturated, or polyunsaturated fats on cardiovascular morbidity and mortality" [257]. For the present study, the main aim was to explore whether or not differences existed in CVD risk factors. Now that these have been documented with an interesting pattern, room for improvement in future surveys that address dietary components exist.

TC/HDL has been recommended for risk prediction [258-260] as it "captures the protective effect of HDL cholesterol as well as the harmful effects of non-HDL cholesterol in a single parameter" [261]. The European SCORE risk estimation system [262], recommended by European CVD prevention guidelines [235], also considers the TC/HDL as an option for risk scoring. It argues that "persons with multiple risk factors tend to have lower HDL cholesterol levels and there is therefore a concern that failing to take HDL cholesterol into account will underestimate risk in those most at risk [258, 259]. A number of clinicians therefore, have expressed interest in a risk estimation system based on cholesterol/HDL ratio" [262]. However, TC/HDL ratio was found to offer no advantage over cholesterol alone as a single index of lipid level [262].

When lipid profile was explored within migrant sub-classifications, migrant with 20 or more years living in an urban area had in average 15 mg/dL units of total cholesterol and 10 mg/dL units of LDL higher than those with less than 20 years in urban area. This is the only significant result, so far, that found out a difference by length of migration.

8.6.7. Inflammation markers

CRP [263] and fibrinogen [264] have been explored widely as inflammation markers in Caucasian population and demonstrate modest associations with CVD. This is the first study in Peru to report inflammation markers as CVD risk factors. Surprisingly, CRP was markedly high in both migrant and urban populations compared to rural groups. Not only that, the degree of change in CRP was much higher than the change observed with fibrinogen. Both, migrants and urban groups had a CRP geometric mean 196% (95% CI 120 - 298) and 224% (124 – 367) higher than the rural people, respectively. Fibrinogen was also higher in both groups compared to rural people but only 11% more. Despite these differences, when explored in terms of SMD, CRP in both groups was around 0.8 SD units higher while fibrinogen was 0.5 SD units higher, which serves best for comparison purposes.

8.6.8. Interpretation

This chapter presented results on BP, lipid profile and inflammation markers. Findings of this study on the impact of migration, at least within Peru, challenges common views and suspicions that following migration all risk factors amongst migrants will mirror the urban population. These findings suggest that the impact of migration on cardiovascular risk profile is not uniform across risk factors and this study add to the understanding of the complexity of such process.

The interpretation of SMD was outlined in the previous chapter, indicating that it is a simple and clear way to present magnitude of the difference across risk factors. SMD presented in this chapter for inflammation markers and some lipid markers indicate that difference sizes compared to the rural group are important, with differences — including their respective CI— ranging from 0.5 to 1 SD. Interestingly, no major differences were observed between migrant or urban groups compared to the rural population in terms of BP, HDL or triglycerides.

SBP and DBP of the rural and migrant populations of this study appear to be within the range of other studies and certainly within range of other Peruvian groups. In addition, other studies of migration, particularly in China [86], Cameroon [83] and

Kenya [78], found that BP rises following migration. One crucial difference from the cited examples is that length of migration was more established in this study, i.e. a mean of 32 years living in an urban area and a mean of 67% of lifetime exposed to an urban area (see Table V-6 and Table V-7). Within migrants no difference was observed when BP was explored by migration sub-classifications. It is plausible to postulate a short-term rise on BP following migration, below 20 years the lowest category, may have occurred, but reversed later in years.

This study thus generates a valid and important question that remains unanswered: why is that, following migration, migrant's anthropometric and inflammation markers, but not BP, shape like the urban group? Genetic differences may not explain these differences as it is widely accepted a high degree of admixture in Peruvian race. Thus the question remains opens and deserves further exploration. The first step in further analyses, although not an objective of the present research, would be to take advantage of the data collected and concentrate on BP as a single outcome and evaluate the association allowing for further adjustment for BMI, smoking and smoking status.

8.7. Summary points

- Blood pressure did not show a clear gradient of difference between migrants and non-migrants. The migrant group had similar SBP but lower DBP than rural group. The urban group had higher SBP but similar DBP than rural group. However, differences were relatively small.

- Compared to the rural group, migrant and urban people were 1.5 (95% CI 0.8 – 2.9) and 5 (95% CI 2.3 – 10.6) times more likely to be hypertensives than rural people, respectively.

- All three groups studied —rural, migrant and urban populations—had a mean total cholesterol below 200 mg/dL, mean HDL above 40 mg/dL and mean LDL below 129 mg/dL which are the recommended thresholds in US guidelines [165]. TC/HDL ratio in all groups was also within recommended ranges, i.e. below 5.

- In the case of lipid profile, no difference was observed between groups for HDL and triglycerides. Total cholesterol, LDL and TC/HDL ratio did, however, show important size of differences (> 0.5 SD) between groups, as did both inflammation markers studied. Migrant and urban groups were similar, and both groups had substantially more adverse lipid profiles and higher levels of inflammation markers.

- When lipid profile was explored within migrant sub-classifications, migrant with 20 or more years living in an urban area had in average 15 mg/dL units of total cholesterol and around 12 mg/dL units of LDL higher than those with less than 20 years in urban area. This is the only significant result, so far, that found out a difference by length of migration.

- CRP geometric means of migrant and urban groups were 196% (95% CI 120% – 298%) and 224% (95% CI 124% – 367%) greater than the geometric mean of the rural group, respectively. In the case of fibrinogen, the geometric mean of the

migrant group was 11% (95% CI 6% – 16%) greater than the geometric mean of rural group and similar values were observed in the urban group.

Chapter IX. Diabetes and metabolic risk factors

This chapter presents results related to fasting glucose and fasting glycosylated haemoglobin (HbA_{1c}), diabetes and impaired fasting glycaemia (IFG), fasting insulin and insulin resistance, and metabolic syndrome. Conversion factors to the international System of Units are provided at the beginning of this document, on page xxvii. All blood samples were taking in fasting conditions early in the morning.

Results are presented for each specific outcome to provide all the information in context. For general informative purposes, descriptive statistics are presented disaggregated by sex when appropriate.

Diabetes and metabolic syndrome are categorical outcomes in this chapter. All other outcomes reported in this chapter are continuous outcomes that were log transformed prior to their analyses. Accordingly, median and interquartile range (IQR) for descriptive statistics, geometric means and ratios for multivariable analyses, and standardised mean differences (SMD) of the log transformed variable, are shown. As discussed previously, geometric means and ratios were chosen over β regression coefficients of the log transformed variable because of the convenience of their interpretation [171, 172]. Reported outputs from SMD regressions are those derived from fully adjusted models taking into account age, sex, individual's deprivation index and parental education.

Later on, and for comparative purposes addressing the extent of variation between different outcome variables and study groups, a specific section is devoted to the visual presentation of SMD. Hereto SMD were calculated on the log transformed variable.

9.1. Fasting glucose and glycosylated haemoglobin

9.1.1. Fasting glucose and glycosylated haemoglobin in all study groups

Compared to the rural group, both migrants and urban populations had higher median fasting glucose but such differences were not observed for HbA_{1c}. There were no differences by sex in these two markers (Table IX-1).

Adjustment for the effects of treatment of antidiabetic medication on fasting glucose or HbA_{1c} was not done as with blood pressure and antihypertensive medication. Glucose distribution in all participants did not have a normal distribution, thus violating the assumption of similar distributions of values measured between treated and untreated subjects [173]. Out of 24 diabetics in total, nine subjects were on treatment and only one had controlled glucose levels (below 126 mg/dL). Arguably, the only one individual on treatment and with controlled glucose could affect the regression estimates unadjusted for treatment effects as long as the remaining 8/9 on treatment had higher glucose levels. This was confirmed by exploring means and medians for glucose levels amongst subjects treated with antidiabetic medications (9/988) and those without treatment (979/988). Such means (and also medians) were in the order of 88 versus 228 mg/dL for untreated versus treated subjects, respectively. Furthermore, mean (median) glucose amongst diabetics without treatment (15/24) and diabetics on treatment (9/24) were 211 (202) and 228 (226 mg/dL), respectively. This finding confirms that the small group of diabetics on treatment had higher glucose means or medians compared to the rest of participants in this study and to the rest of diabetics.

Therefore, based on these circumstances, all analyses, unless otherwise stated in this chapter, were done without excluding those diabetics currently on medication. Additional sensitivity regression analyses excluding those taking antidiabetic medications were also performed for the main analyses of glucose and HbA_{1c} by

study groups as shown Table IX-1. These sensitivity analyses had a minor impact only in the urban group estimates decreasing glucose and HbA_{1c} ratios in the order of 2 to 3%.

After multivariate adjustment and compared to the rural group which had, excluding diabetics on medication, a fasting glucose geometric mean of 79.9 mg/dL (95% CI 77 –83), both migrant and urban populations had significantly higher geometric means of fasting glucose. The glucose geometric means of migrant and urban groups were 9% (95% CI 5% – 12%) and 13% (95% CI 8% – 17%) greater than the geometric mean of the rural group, respectively.

In the case of HbA_{1c}, in both scenarios, with and without the exclusion of diabetics on medication, there was no difference in median HbA_{1c} or geometric mean ratio between migrant and rural populations. However, in the fully adjusted model, urban group showed a geometric mean of HbA_{1c} 3 – 5% higher than the rural group.

In terms of SMD, compared to the rural group, both migrants and urban groups had higher significant differences for glucose but not for HbA_{1c} (Table VIII-13).

Table IX-1. Descriptive and multivariable analyses of fasting glucose and glycosylated haemoglobin by study groups

	Rural	Migrant	Urban	
Descriptive statistics	**Median (IQR)**	**Median (IQR)**	**Median (IQR)**	**Missing data**
Glucose (mg/dL)				1 (0.1%)
All	**80 (12)**	**86 (12)**	**88 (13)**	
Female	81 (12)	85 (13)	89 (12)	
Male	78 (12)	87 (12)	87 (13)	
HbA_{1c} (%)				0
All	**5.7 (0.5)**	**5.5 (0.5)**	**5.7 (0.6)**	
Female	5.7 (0.5)	5.6 (0.4)	5.7 (0.5)	
Male	5.7 (0.6)	5.5 (0.4)	5.7 (0.6)	
Multivariable Glucose (log)	**Geometric mean (95% CI)**	**Ratio (95%CI)**	**Ratio (95%CI)**	**R^2**
Model 1	79.8 (77.7 – 81.9)	1.09 (1.06 – 1.12)	1.16 (1.12 – 1.2)	0.09
Model 2	79.5 (76.4 – 82.8)	1.09 (1.06 – 1.14)	1.16 (1.11 – 1.22)	0.09
Model 3	80 (76.6 – 83.6)	1.09 (1.05 – 1.13)	1.16 (1.11 – 1.22)	0.09
Multivariable Glucose (log) excluding treated*				
Model 1	79.7 (77.9 – 81.6)	1.09 (1.06 – 1.11)	1.13 (1.1 – 1.16)	0.09
Model 2	79.7 (77 – 82.5)	1.09 (1.05 – 1.12)	1.13 (1.09 – 1.17)	0.09
Model 3	79.9 (77 – 83)	1.09 (1.05 – 1.12)	1.13 (1.08 – 1.17)	0.09

Table IX-1. (continued)

	Rural	Migrant	Urban	
Multivariable HbA_{1c} (log)				
Model 1	5.7 (5.6 – 5.8)	0.99 (0.97 – 1.01)	1.04 (1.01 – 1.06)	0.07
Model 2	5.6 (5.5 – 5.8)	1.00 (0.98 – 1.03)	1.04 (1.01 – 1.08)	0.07
Model 3	5.7 (5.5 – 5.9)	1.00 (0.97 – 1.03)	1.05 (1.02 – 1.08)	0.08
Multivariable HbA_{1c} (log) excluding treated*				
Model 1	5.7 (5.6 – 5.8)	0.99 (0.97 – 1.01)	1.02 (0.99 – 1.04)	0.07
Model 2	5.6 (5.5 – 5.8)	1.00 (0.98 – 1.02)	1.02 (1.00 – 1.05)	0.07
Model 3	5.7 (5.5 – 5.8)	1.00 (0.98 – 1.02)	1.03 (1.00 – 1.06)	0.07
	Within group SD	**SMD (95%CI)**	**SMD (95%CI)**	**R^2**
SMD Glucose (log)	0.2	0.5 (0.3 – 0.7)	0.9 (0.6 – 1.1)	0.09
SMD HbA_{1c} (log)	0.1	0 (-0.2 – 0.2)	0.4 (0.1 – 0.6)	0.07

Notes: All regressions calculated using the rural group as baseline (Total n = 988/989).
Model 1: Adjusted for age and sex; Model 2: As model 1 plus individual socioeconomic deprivation; Model 3: As model 2 plus highest parental education (this final model was also used for SMD regressions)
* Only in these cases, regression analyses were carried out excluding those diabetics currently on treatment (n = 9/988)

9.1.2. Fasting glucose and glycosylated haemoglobin in migrant subgroups

Sub-classifications by length of residence in urban area, lifetime exposure to an urban area and age at first migration were analysed in the migrant population only. All analyses reported in this section were carried out without excluding those diabetics currently on medication. Geometric means / ratios (95% CI) of comparisons between various categories used within migrants and glucose and HbA_{1c} are shown in Table IX-2 and Table IX-3.

Migrants with longer periods of residence in an urban area, compared to those living in an urban area < 20 years, appeared to have a borderline significant gradient towards increased levels of fasting glucose. A difference in glucose was also observed in those who migrated older than 12 years-old that had a 4% (95% CI 1% – 7%) higher ratio of blood glucose compared to those who migrated younger than 12 years-old. These were borderline significant associations, and possibly likely to be due to chance. Hereto, since blood glucose is central for the determination of diabetes, IFG and metabolic syndrome, this difference on glucose levels observed by age at migration will also have an impact in those categorisations to be presented later in this chapter.

Table IX-2. Multivariable analyses of migrant sub-classifications and fasting glucose

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 559/559			
Baseline			
<20 years in urban area	84.2 (80.2 – 88.4)	83.8 (79.7 – 88.2)	83.8 (79.4 – 88.5)
Ratios			
20-29 years in urban area	1.03 (0.97 – 1.08)	1.03 (0.98 – 1.08)	1.03 (0.98 – 1.08)
30-39 years in urban area	1.06 (1.00 – 1.12)	1.06 (1.00 – 1.12)	1.06 (1.00 – 1.12)
≥40 years in urban area	1.05 (0.99 – 1.12)	1.05 (0.99 – 1.12)	1.05 (0.99 – 1.12)
R^2	0.03	0.03	0.03
Migrants, by lifetime exposure to urban area*, n = 559/559			
Baseline			
Q1, lowest	86 (83.4 – 88.8)	85.9 (83.1 – 88.8)	85.8 (82.6 – 89.1)
Ratios			
Q2	1.04 (1.00 – 1.08)	1.04 (1.00 – 1.08)	1.04 (1.00 – 1.08)
Q3	1.02 (0.98 – 1.06)	1.02 (0.98 – 1.06)	1.02 (0.98 – 1.06)
Q4, highest	1.00 (0.96 – 1.04)	1.00 (0.96 – 1.04)	1.00 (0.96 – 1.04)
R^2	0.03	0.03	0.03
Migrants, by age at first migration**, n = 585/585			
Baseline			
≤ 12 yo at first migration	85.1 (83 – 87.3)	85 (82.7 – 87.3)	84.7 (82 – 87.5)
Ratios			
> 12 yo at first migration	1.04 (1.01 – 1.07)	1.04 (1.01 – 1.07)	1.04 (1.01 – 1.07)
R^2	0.03	0.03	0.03

Note:
All baseline values, except ratios and R^2 values, are geometric means (95% CI) in mg/dL units
Subjects with diabetes and currently on treatment (n = 9) were not excluded from this analyses
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults.

Table IX-3. Multivariable analyses of migrant sub-classifications and glycosylated haemoglobin

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 559/559			
Baseline			
<20 years in urban area	5.6 (5.4 – 5.8)	5.6 (5.4 – 5.8)	5.6 (5.4 – 5.9)
Ratios			
20-29 years in urban area	1.00 (0.96 – 1.03)	1.00 (0.96 – 1.03)	1.00 (0.96 – 1.03)
30-39 years in urban area	1.03 (0.99 – 1.07)	1.03 (0.99 – 1.08)	1.03 (0.99 – 1.08)
≥40 years in urban area	1.01 (0.97 – 1.06)	1.02 (0.97 – 1.06)	1.02 (0.97 – 1.06)
R^2	0.07	0.07	0.07
Migrants, by lifetime exposure to urban area*, n = 559/559			
Baseline			
Q1, lowest	5.6 (5.5 – 5.8)	5.6 (5.5 – 5.8)	5.7 (5.5 – 5.8)
Ratios			
Q2	1.02 (0.99 – 1.05)	1.02 (0.99 – 1.05)	1.02 (0.99 – 1.05)
Q3	1.01 (0.99 – 1.04)	1.02 (0.99 – 1.05)	1.02 (0.99 – 1.05)
Q4, highest	1.00 (0.97 – 1.03)	1.00 (0.97 – 1.03)	1.00 (0.97 – 1.03)
R^2	0.06	0.06	0.06
Migrants, by age at first migration**, n = 585/585			
Baseline			
≤ 12 yo at first migration	5.6 (5.5 – 5.7)	5.6 (5.5 – 5.7)	5.6 (5.5 – 5.7)
Ratios			
> 12 yo at first migration	1.02 (1.00 – 1.05)	1.02 (1.00 – 1.05)	1.02 (1.00 – 1.05)
R^2	0.06	0.06	0.06

Note: All baseline values, except ratios and R^2 values, are geometric means (95% CI) in % units

Subjects with diabetes and currently on treatment (n = 9) were not excluded from this analyses

Model 1: Adjusted for age and sex

Model 2: As model 1 plus individual socioeconomic deprivation

Model 3: As model 2 plus highest parental education

* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%

** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults.

9.2. Diabetes and impaired fasting glycaemia

9.2.1. Diabetes

Following WHO 1999 guidelines, diabetes was defined as fasting glucose ≥ 126 mg/dL (or ≥ 7 mmol/L) [159]. Additionally, subjects that self reported a physician diagnosis of diabetes and who were currently receiving antidiabetic medication were also considered as diabetics.

Breakdown of diabetes defined by blood glucose measurement or report of current specific antidiabetic treatment is provided in Table IX-4. None of the participants reported type I diabetes mellitus. Prevalence of diabetes mellitus type II, hereafter referred as diabetes unless otherwise stated, was 2.4% (24/988). Amongst diabetics, 8/24 (33.3%) were on antidiabetic medication and only 1/8 was controlled, i.e. had a glucose within normal ranges. Of those on antidiabetic medication, 2/8 were currently using insulin: two migrant females, aged 51 and 52 years, and thus very unlikely to correspond to type I diabetes if taking into account survival of this type of diabetes in developing countries.

Table IX-4. Distribution of diabetes-related categories by study group

	Rural	Migrant	Urban	Total	Missing
	201	589	199		
Diabetes defined					
(A) by blood glucose	1/200	12/589	10/199	23/988	1/989
	0.5%	2%	5%	2.3%	0.1%
(B) currently on treatment*	0/200	1/589	0/199	1/988	
	0%	0.2%	0%	0.1%	
Diabetes (A + B)	1/200	13/589	10/199	24/988	1/989
	0.5%	2.2%	5%	2.4%	0.1%

* This group included only those diabetics currently on medication and with blood glucose levels within normal ranges. They reported to be on antidiabetic medication and provided the specific name of the drug they were taking. In other denominations these individuals would be classified as “aware and controlled diabetics”.

Table IX-5 presents the distribution and OR (95% CI) of diabetes in all study groups and according to migrant subgroups.

As stated before, 24/988 subjects were classified as diabetics. Diabetes was prevalent in 0.5%, 2.2% and 5% of rural, migrant and urban population, respectively. In the fully adjusted model compared to the rural group, migrant and urban people were 6 (95% CI 0.7 – 51.9) and 15.9 (95% CI 1.6 – 159.2) times more likely to be diabetics than rural people, respectively. By examining CI, this association was only significant for the comparison between urban versus rural populations, and it was not significant for the comparison between migrant versus rural groups. This significant association had, however, a wide CI due to the fact that only 1/201 rural people were classified as diabetics.

When exploring the association of diabetes within migrants, no clear associations between diabetes and number of years in an urban area or lifetime exposure to urban area exposures were observed. Those who migrated aged 12-years or older were 7.5 (95% CI 0.9 – 59.7) times more likely to be diabetics than those who migrated younger with a borderline non-significant CI. However, such wide CI was due the huge imbalance of 1/225 (0.4%) versus 12/360 (3.3%) subjects with diabetes in each group.

Table IX-5. Association between diabetes and migration by exposure groups

Exposure Groups	Diabetes n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Rural, Migrant, Urban	**(n = 988/989)**			
Rural	1 (0.5%)	1	1	1
Migrant	13 (2.2%)	4.6 (0.6 – 35.7)	6.3 (0.7 – 54.4)	6 (0.7 – 51.9)
Urban	10 (5%)	10.7 (1.4 – 84.7)	15.8 (1.7 – 149.5)	15.9 (1.6 – 159.2)
Migrants, by years in urban area	**(n = 551/559)**			
<20 years in urban area	1 (1.9%)	1	1	1
20-29 years in urban area	2 (1%)	0.5 (0 – 5.9)	0.5 (0 – 6.2)	0.5 (0 – 6.3)
30-39 years in urban area	7 (4.1%)	2.1 (0.2 – 18.3)	2.3 (0.3 – 21.3)	2.4 (0.3 – 21.5)
≥40 years in urban area	3 (2.2%)	0.9 (0.1 – 11.1)	1 (0.1 – 13.2)	1 (0.1 – 13.3)
Migrants, by lifetime exposure to urban area*	**(n = 551/559)**			
Q1, lowest	2 (1.4%)	1	1	1
Q2	6 (4.3%)	3.4 (0.7 – 17.4)	3.3 (0.6 – 17)	3.5 (0.7 – 17.8)
Q3	3 (2.1%)	1.5 (0.2 – 9.3)	1.6 (0.3 – 9.9)	1.6 (0.3 – 10)
Q4, highest	2 (1.5%)	1 (0.1 – 7.4)	1.1 (0.1 – 7.8)	1.1 (0.1 – 7.9)

Table IX-5. (continued)

Exposure Groups	Diabetes n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Migrants, by age at first migration**	**(n = 585/585)**			
≤ 12 yo at first migration	1 (0.4%)	1	1	1
> 12 yo at first migration	12 (3.3%)	7.4 (0.9 – 58.6)	7.4 (0.9 – 58.1)	7.5 (0.9 – 58.7)

Notes: All values are OR (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

9.2.2. Impaired fasting glycaemia or diabetes

Following WHO 1999 guidelines, IFG was defined as fasting plasma glucose ≥ 110 mg/dL (or ≥ 6.1 mmol/L) and < 126 mg/dL (or < 7 mmol/L) [159]. In this section, IFG was considered together with diabetes, therefore anybody with fasting glucose ≥ 110 mg/dL (or ≥ 6.1 mmol/L).

IFG or diabetes was prevalent in 1.5%, 4% and 8% of rural, migrant and urban population, respectively. In the fully adjusted model compared to the rural group, migrant and urban people were 3.5 (95% CI 0.9 – 13.4) and 8.9 (95% CI 2 –39.1) times more likely to have IFG than rural people, respectively.

When exploring the association of IFG within migrants, no clear associations between diabetes and number of years in an urban area or lifetime exposure to urban area exposures were observed. However, those who migrated aged 12-years or older were 6.4 (95% CI 1.5 – 27.8) times more likely to have IFG than those who migrated younger.

Table IX-6. Association between impaired fasting glycaemia or diabetes and migration by exposure groups

Exposure Groups	IFG or Diabetes n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Rural, Migrant, Urban	**(n = 988/989)**			
Rural	3 (1.5%)	1	1	1
Migrant	23 (3.9%)	2.8 (0.8 – 9.3)	3.7 (1 – 14)	3.5 (0.9 – 13.4)
Urban	16 (8%)	5.9 (1.7 – 20.7)	8.4 (2 – 35.5)	8.9 (2 – 39.1)
Migrants, by years in urban area	**(n = 559/559)**			
<20 years in urban area	2 (3.8%)	1	1	1
20-29 years in urban area	6 (3%)	0.8 (0.1 – 4)	0.8 (0.2 – 4.1)	0.8 (0.1 – 4)
30-39 years in urban area	9 (5.3%)	1.2 (0.3 – 6.2)	1.4 (0.3 – 6.8)	1.3 (0.3 – 6.6)
≥40 years in urban area	6 (4.5%)	0.8 (0.1 – 4.6)	0.9 (0.1 – 5.1)	0.9 (0.1 – 5.1)
Migrants, by lifetime exposure to urban area*	**(n = 559/559)**			
Q1, lowest	4 (4.3%)	1	1	1
Q2	8 (5.8%)	1.5 (0.5 – 4.5)	1.5 (0.5 – 4.4)	1.5 (0.5 – 4.5)
Q3	7 (4.9%)	1.2 (0.4 – 3.6)	1.2 (0.4 – 3.9)	1.3 (0.4 – 3.9)
Q4, highest	2 (1.5%)	0.3 (0.1 – 1.7)	0.3 (0.1 – 1.7)	0.4 (0.1 – 1.8)

Table IX-6. (continued)

Exposure Groups	IFG or Diabetes n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Migrants, by age at first migration**	**(n = 585/585)**			
≤ 12 yo at first migration	2 (0.9%)	1	1	1
> 12 yo at first migration	21 (5.8%)	6.5 (1.5 – 28.5)	6.5 (1.5 – 28.4)	6.4 (1.5 – 27.8)

Notes: All values are OR (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

9.3. Fasting insulin and Insulin Resistance (HOMA model)

Insulin resistance was calculated using HOMA calculator (Oxford Centre for Diabetes, Endocrinology & Metabolism, Diabetes Trials Unit, http://www.dtu.ox.ac.uk/) [160] and excluding those with diabetes.

9.3.1. Fasting insulin and insulin resistance in all study groups

A marked gradient of increasing median in fasting insulin from rural to migrant to urban groups was observed in fasting insulin and insulin resistance. Within each specific study group females had 2-3 μIU/mL of insulin and 0.3 points of insulin resistance higher than males (Table IX-7). Means in fasting insulin by gender were significantly different (t test, p <0.0001). However, the normal range for insulin is 2 to 20 μIU/mL [265], and both female and males had insulin values within the normal range.

After multivariate adjustment and compared to the rural group which had a fasting insulin geometric mean of 2.7 μIU/mL (95% CI 2.3 – 3.2), both migrant and urban populations had significantly higher geometric means of insulin. The insulin geometric means of migrant and urban groups were 193% (95% CI 135% – 266%) and 251% (95% CI 167% – 362%) greater than the geometric mean of the rural group, respectively.

In the case of HOMA insulin resistance, the geometric mean of the migrant group was 202% (95% CI 156% – 257%) greater than the geometric mean of rural group. Similarly, the geometric mean of the urban group and 245% (95% CI 181% – 324%) greater than the geometric mean of rural group.

In terms of SMD, compared to the rural group, both migrants and urban groups had similar and very high significant differences in the order of 1 SD for insulin and 1.5 SD for insulin resistance (Table IX-7).

Table IX-7. Descriptive and multivariable analyses of fasting insulin and insulin resistance by study groups

	Rural	Migrant	Urban	
Descriptive statistics	**Median (IQR)**	**Median (IQR)**	**Median (IQR)**	**Missing data**
Insulin (μIU/mL)				11 (1.1%)
All	2.5 (3.9)	6.7 (5.9)	8.5 (7.8)	
Female	3.4 (3.7)	8.2 (7.4)	9.9 (10.3)	
Male	1.5 (2.7)	5.3 (4.6)	6.7 (6.6)	
HOMA-IR				12 (1.2%)
All,	0.3 (0.5)	0.9 (0.8)	1.1 (1.1)	
Female	0.5 (0.5)	1 (1)	1.3 (1.3)	
Male	0.2 (0.4)	0.7 (0.6)	0.9 (0.9)	
Multivariable Insulin (log)	**Geometric mean (95% CI)**	**Ratio (95%CI)**	**Ratio (95%CI)**	**R^2**
Model 1	2.7 (2.3 – 3.2)	3.09 (2.60 – 3.67)	4.05 (3.28 – 4.99)	0.23
Model 2	2.9 (2.3 – 3.8)	2.89 (2.31 – 3.6)	3.74 (2.86 – 4.88)	0.23
Model 3	2.6 (2 – 3.3)	2.93 (2.35 – 3.66)	3.51 (2.67 – 4.62)	0.23
Multivariable HOMA-IR (log)				
Model 1	0.3 (0.3 – 0.4)	3.27 (2.88 – 3.71)	4.02 (3.43 – 4.71)	0.36
Model 2	0.4 (0.3 – 0.4)	2.99 (2.53 – 3.53)	3.62 (2.96 – 4.42)	0.36
Model 3	0.3 (0.3 – 0.4)	3.02 (2.56 – 3.57)	3.45 (2.81 – 4.24)	0.36

Table IX-7. (continued)

	Rural	Migrant	Urban	
	Within group SD	**SMD (95%CI)**	**SMD (95%CI)**	**R^2**
SMD Insulin (log)	1.1	1 (0.8 – 1.2)	1.2 (0.9 – 1.5)	0.23
SMD HOMA-IR (log)	0.8	1.4 (1.2 – 1.7)	1.6 (1.3 – 1.9)	0.36

Notes: All regressions calculated using the rural group as baseline (Total n = 988/989).
Model 1: Adjusted for age and sex; Model 2: As model 1 plus individual socioeconomic deprivation; Model 3: As model 2 plus highest parental education (this final model was also used for SMD regressions)

9.3.2. Fasting insulin and insulin resistance in migrant subgroups

As with other outcomes, sub-classifications by length of residence in urban area, lifetime exposure to an urban area and age at first migration were analysed in the migrant population only. Table IX-8 and Table IX-9 show geometric means / ratios (95% CI) of comparisons between various categories used within migrants and insulin and insulin resistance.

Migrants with longer periods of residence in an urban area, compared to those living in an urban area < 20 years, appeared to have a borderline significant gradient towards increased levels of HOMA insulin resistance but not fasting insulin.

The difference observed for fasting glucose by age at migration was not observed for insulin or insulin resistance. No clear pattern of difference was observed using remaining migrant sub-classifications in fasting insulin or insulin resistance. All CI for geometric mean ratios in crude and adjusted multivariable models did overlap the value of one in both, insulin and insulin resistance.

Table IX-8. Multivariable analyses of migrant sub-classifications and fasting insulin

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 558/559			
Baseline			
<20 years in urban area	7.5 (5.4 – 10.4)	7.5 (5.3 – 10.6)	6.9 (4.8 – 9.9)
Ratios			
20-29 years in urban area	1.04 (0.73 – 1.48)	1.04 (0.73 – 1.48)	1.05 (0.73 – 1.49)
30-39 years in urban area	1.23 (0.85 – 1.79)	1.23 (0.84 – 1.79)	1.23 (0.85 – 1.8)
≥40 years in urban area	1.02 (0.67 – 1.56)	1.02 (0.66 – 1.56)	1.01 (0.66 – 1.55)
R^2	0.06	0.06	0.07
Migrants, by lifetime exposure to urban area*, n = 558/559			
Baseline			
Q1, lowest	7.5 (6 – 9.2)	7.6 (6 – 9.5)	7 (5.4 – 9)
Ratios			
Q2	1.16 (0.88 – 1.52)	1.16 (0.88 – 1.52)	1.16 (0.88 – 1.53)
Q3	1.04 (0.79 – 1.37)	1.04 (0.79 – 1.37)	1.03 (0.78 – 1.36)
Q4, highest	1.13 (0.86 – 1.48)	1.12 (0.85 – 1.48)	1.11 (0.84 – 1.46)
R^2	0.06	0.06	0.06
Migrants, by age at first migration**, n = 584/585			
Baseline			
≤ 12 yo at first migration	8.4 (7 – 9.9)	8.4 (7 – 10.1)	7.8 (6.2 – 9.7)
Ratios			
> 12 yo at first migration	0.95 (0.78 – 1.15)	0.95 (0.78 – 1.15)	0.95 (0.78 – 1.16)
R^2	0.06	0.06	0.06

Note: All baseline values, except ratios and R^2 values, are geometric means (95% CI) in µIU/mL units
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table IX-9. Multivariable analyses of migrant sub-classifications and HOMA insulin resistance

	Model 1	Model 2	Model 3
Migrants, by years in urban area, n = 536/559			
Baseline			
<20 years in urban area	0.9 (0.7 – 1.1)	0.9 (0.7 – 1.1)	0.8 (0.7 – 1)
Ratios			
20-29 years in urban area	1.07 (0.87 – 1.32)	1.07 (0.87 – 1.32)	1.07 (0.87 – 1.32)
30-39 years in urban area	1.27 (1.02 – 1.58)	1.26 (1.01 – 1.58)	1.26 (1.01 – 1.58)
≥40 years in urban area	1.30 (1.01 – 1.66)	1.29 (1.00 – 1.66)	1.28 (1.00 – 1.65)
R^2	0.11	0.11	0.12
Migrants, by lifetime exposure to urban area*, n = 536/559			
Baseline			
Q1, lowest	0.9 (0.8 – 1)	0.9 (0.8 – 1)	0.9 (0.8 – 1)
Ratios			
Q2	1.16 (0.99 – 1.37)	1.16 (0.99 – 1.37)	1.16 (0.99 – 1.37)
Q3	1.20 (1.02 – 1.42)	1.20 (1.02 – 1.41)	1.20 (1.02 – 1.41)
Q4, highest	1.16 (0.98 – 1.36)	1.15 (0.98 – 1.35)	1.14 (0.97 – 1.35)
R^2	0.11	0.11	0.11
Migrants, by age at first migration**, n = 562/585			
Baseline			
≤ 12 yo at first migration	1 (0.9 – 1.1)	1 (0.9 – 1.2)	1 (0.9 – 1.1)
Ratios			
> 12 yo at first migration	0.98 (0.87 – 1.10)	0.98 (0.87 – 1.10)	0.98 (0.87 – 1.10)
R^2	0.10	0.10	0.10

Note: All baseline values, except ratios and R^2 values, are geometric means (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

9.4. Metabolic syndrome

This syndrome has been defined in different ways by different bodies [159, 162, 163, 165], but all include (central) obesity, dyslipidaemia, hypertension and glucose intolerance. In this Chapter, as presented in Table II-12 (page 58), three definitions were used for the classification of metabolic syndrome in the sample studied and are presented again in Table IX-10.

First, the World Health Organization (WHO) 1999 definition [159], which considers a fasting plasma glucose ≥ 110 mg/dL (6.1 mmol/L) or diabetes as the central component, plus two other risk factors.

Second, International Diabetes Federation (IDF) 2005 definition [162, 163] that uses central obesity as a must criteria, and uses a lower plasma glucose (≥ 100 mg/dL or ≥ 5.6 mmol/L) threshold and a lower blood pressure cut-off than in the WHO definition. The advantage of this classification is that it acknowledges ethnic specific markers of obesity. However, due to the paucity of data on obesity from Latin America, South Asian cut-offs are recommended for this group.

Third, the US American Heart Association / National Heart, Lung and Blood Institute (AHA/NHLBI) 2005 definition [164], which is an updated and revised definition from the US National Cholesterol Education Programme Adult Treatment Panel III (NCEP ATP III) guidelines [165]. In relation to the later US guidelines, this updated classification uses a lower plasma glucose threshold (≥ 100 mg/dL or ≥ 5.6 mmol/L), similar to the IDF definition, and considers the individual's treatment status for specific conditions.

Table IX-10. Definitions of metabolic syndrome

Parameters	WHO 1999	IDF 2005	AHA/NHLBI 2005
Required	Insulin resistance in top 25 percent*; glucose ≥6.1 mmol/L (110 mg/dL); 2-hour glucose ≥7.8 mmol/L (140 mg/dL)	Waist ≥94 cm (men) or ≥80 cm (women)**	
Number of abnormalities	And ≥2 of:	And ≥2 of:	≥3 of:
Glucose		≥5.6 mmol/L (100 mg/dL) or diagnosed diabetes	≥5.6 mmol/L (100 mg/dL) or drug treatment for elevated blood glucose
HDL cholesterol	<0.9 mmol/L (35 mg/dL) (men); <1.0 mmol/L (40 mg/dL) (women)	<1.0 mmol/L (40 mg/dL) (men); <1.3 mmol/L (50 mg/dL) (women) or drug treatment for low HDL-C	<1.0 mmol/L (40 mg/dL) (men); <1.3 mmol/L (50 mg/dL) (women) or drug treatment for low HDL-C***
Triglycerides	or ≥1.7 mmol/L (150 mg/dL)	≥1.7 mmol/L (150 mg/dL) or drug treatment for high triglycerides	≥1.7 mmol/L (150 mg/dL) or drug treatment for elevated triglycerides***
Obesity	Waist/hip ratio >0.9 (men) or >0.85 (women) or BMI ≥30 kg/m^2		Waist ≥102 cm (men) or ≥88 cm (women)
Hypertension	≥140/90 mmHg	≥130/85 mmHg or drug treatment for hypertension	≥130/85 mmHg or drug treatment for hypertension

* Insulin resistance measured using insulin clamp; ** For South Asia and Chinese patients, waist ≥90 cm (men) or ≥80 cm (women); for Japanese patients, waist ≥90 cm (men) or ≥80 cm (women); *** Treatment with one or more of fibrates or niacin.
Table adapted from: http://www.uptodate.com/patients/content/image.do?file=endo_pix/defini17.htm&view=print

For clarity purposes, prevalences of metabolic syndrome according to the different definitions used are presented together in Table IX-11. It was not the purpose of this study to compare the performance of specific guidelines to each other.

The table shows two consistent findings. Firstly, using any of the definitions, a gradient of increasing prevalence of metabolic syndrome is observed from rural to migrant to urban groups. This gradient was not observed in the different sub-classifications within the migrant group. Secondly, prevalence estimates of metabolic syndrome, in all main groups and migrant sub-classifications, differ importantly and are substantially higher with the IDF or AHA/NHLBI guidelines compared to the WHO definition.

Table IX-11. Prevalence of metabolic syndrome by exposure groups

Exposure Groups	MS WHO n (%)	MS IDF n (%)	MS AHA/NHLBI n (%)
Rural, Migrant, Urban (n = 989/989)			
Rural	1 (0.5%)	13 (6.5%)	16 (8%)
Migrant	16 (2.7%)	176 (29.9%)	144 (24.5%)
Urban	14 (7%)	85 (42.7%)	74 (37.2%)
Migrants, by years in urban area (n = 559/559)			
<20 years in urban area	1 (1.9%)	10 (18.9%)	7 (13.2%)
20-29 years in urban area	2 (1%)	51 (25.1%)	29 (14.3%)
30-39 years in urban area	8 (4.7%)	58 (34.3%)	52 (30.8%)
≥40 years in urban area	5 (3.7%)	45 (33.6%)	44 (32.8%)
Migrants, by lifetime exposure to urban area (n = 559/559)			
Q1, lowest	3 (2.1%)	42 (29.8%)	30 (21.3%)
Q2	6 (4.3%)	52 (37.4%)	38 (27.3%)
Q3	5 (3.5%)	38 (26.8%)	35 (24.7%)
Q4, highest	2 (1.5%)	32 (23.4%)	29 (27.2%)
Migrants, by age at first migration (n = 585/585)			
≤ 12 yo at first migration	1 (0.4%)	49 (21.8%)	40 (17.8%)
> 12 yo at first migration	15 (4.2%)	126 (35%)	103 (28.6%)

Multivariate analyses using metabolic syndrome as a categorical outcome are presented in Table IX-12, Table IX-13 and Table IX-14 for the WHO, IDF and AHA/NHLBI definitions, respectively. Compared to the rural group, both migrant and urban population were more likely to have metabolic syndrome. The most consistent associations, with narrower CI, were observed with the AHA/NHLBI definition. Using this definition, migrants had an OR 5.1 (95% CI 2.7 – 9.6) and urban people had an OR 8.7 (95% CI 4.2 – 17.9) times more likely to have metabolic syndrome than the rural group. Similar OR (95% CI), also compared to the rural group, using the WHO definition were 5.4 (0.6 – 47.1) and 13.8 (1.4 – 134.2) and using the IDF definition were 7.2 (3.7 – 14) and 12 (5.7 – 25.3) for migrants and urban people, respectively. The wide CI observed with the WHO definition is directly related to the fact that within the rural group only 1/201 subject (0.5%) classified as having metabolic syndrome.

Within migrant sub-classifications, the only consistent significant association was observed when migrants were divided by age at migration. Those who migrated aged 12 years-old or more had increased odds of having metabolic syndrome in all definitions. Such association however was borderline significant, with a lower CI close to one. No clear pattern of association was observed when migrants were classified by length of migration or lifetime exposure to urban environment.

Table IX-12. Association between WHO definition of metabolic syndrome and migration by exposure groups

Exposure Groups	MS WHO n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Rural, Migrant, Urban	**(n = 989/989)**			
Rural	1 (0.5%)	1	1	1
Migrant	16 (2.7%)	5.7 (0.8 – 43.5)	5.6 (0.6 – 48.9)	5.4 (0.6 – 47.1)
Urban	14 (7%)	15.5 (2 – 119.1)	15.1 (1.6 – 141)	13.8 (1.4 – 134.2)
Migrants, by years in urban area	**(n = 559/559)**			
<20 years in urban area	1 (1.9%)	1	1	1
20-29 years in urban area	2 (1%)	0.5 (0 – 5.9)	0.5 (0 – 5.8)	0.5 (0 – 6)
30-39 years in urban area	8 (4.7%)	2.3 (0.3 – 18.9)	2.2 (0.3 – 18.2)	2.2 (0.3 – 18.8)
≥40 years in urban area	5 (3.7%)	1.2 (0.1 – 12.3)	1.2 (0.1 – 11.7)	1.2 (0.1 – 12.2)
Migrants, by lifetime exposure to urban area*	**(n = 559/559)**			
Q1, lowest	3 (2.1%)	1	1	1
Q2	6 (4.3%)	2.5 (0.6 – 10.5)	2.6 (0.6 – 11.1)	2.7 (0.6 – 11.9)
Q3	5 (3.5%)	1.8 (0.4 – 7.7)	1.7 (0.4 – 7.4)	1.7 (0.4 – 7.7)
Q4, highest	2 (1.5%)	0.7 (0.1 – 4.3)	0.7 (0.1 – 4.2)	0.7 (0.1 – 4.3)

Table IX-12. (continued)

Exposure Groups	MS WHO n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Migrants, by age at first migration**	**(n = 585/585)**			
≤ 12 yo at first migration	1 (0.4%)	1	1	1
> 12 yo at first migration	15 (4.2%)	8.6 (1.1 – 66.4)	8.6 (1.1 – 66.6)	8.8 (1.1 – 68.2)

Notes: All values are OR (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table IX-13. Association between IDF definition of metabolic syndrome and migration by exposure groups

Exposure Groups	MS IDF n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Rural, Migrant, Urban	**(n = 989/989)**			
Rural	13 (6.5%)	1	1	1
Migrant	176 (29.9%)	6.8 (3.7 – 12.3)	7.4 (3.8 – 14.4)	7.2 (3.7 – 14)
Urban	85 (42.7%)	11.9 (6.3 – 22.6)	13.2 (6.4 – 27.3)	12 (5.7 – 25.3)
Migrants, by years in urban area	**(n = 559/559)**			
<20 years in urban area	10 (18.9%)	1	1	1
20-29 years in urban area	51 (25.1%)	1.4 (0.7 – 3.2)	1.4 (0.7 – 3.2)	1.5 (0.7 – 3.2)
30-39 years in urban area	58 (34.3%)	1.9 (0.9 – 4.3)	1.9 (0.9 – 4.3)	1.9 (0.9 – 4.3)
≥40 years in urban area	45 (33.6%)	1.2 (0.5 – 2.9)	1.2 (0.5 – 2.9)	1.2 (0.5 – 2.9)
Migrants, by lifetime exposure to urban area*	**(n = 559/559)**			
Q1, lowest	42 (29.8%)	1	1	1
Q2	52 (37.4%)	1.6 (0.9 – 2.6)	1.6 (0.9 – 2.6)	1.6 (0.9 – 2.7)
Q3	38 (26.8%)	0.9 (0.5 – 1.5)	0.8 (0.5 – 1.5)	0.9 (0.5 – 1.5)
Q4, highest	32 (23.4%)	0.7 (0.4 – 1.2)	0.7 (0.4 – 1.2)	0.7 (0.4 – 1.2)

Table IX-13. (continued)

Exposure Groups	MS IDF n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Migrants, by age at first migration**	**(n = 585/585)**			
≤ 12 yo at first migration	49 (21.8%)	1	1	1
> 12 yo at first migration	126 (35%)	1.7 (1.1 – 2.5)	1.7 (1.1 – 2.5)	1.7 (1.1 – 2.6)

Notes: All values are OR (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

Table IX-14. Association between AHA/NHLBI definition of metabolic syndrome and migration by exposure groups

Exposure Groups	MS AHA/NHLBI n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Rural, Migrant, Urban	**(n = 989/989)**			
Rural	16 (8%)	1	1	1
Migrant	144 (24.5%)	4.1 (2.4 – 7.2)	5.1 (2.7 – 9.5)	5.1 (2.7 – 9.6)
Urban	74 (37.2%)	7.7 (4.2 – 14)	9.8 (4.8 – 19.7)	8.7 (4.2 – 17.9)
Migrants, by years in urban area	**(n = 559/559)**			
<20 years in urban area	7 (13.2%)	1	1	1
20-29 years in urban area	29 (14.3%)	1.1 (0.4 – 2.7)	1.1 (0.4 – 2.8)	1.1 (0.5 – 2.8)
30-39 years in urban area	52 (30.8%)	2.5 (1 – 6.3)	2.7 (1.1 – 6.7)	2.7 (1.1 – 6.8)
≥40 years in urban area	44 (32.8%)	1.9 (0.7 – 5)	2.1 (0.8 – 5.4)	2 (0.8 – 5.4)
Migrants, by lifetime exposure to urban area*	**(n = 559/559)**			
Q1, lowest	30 (21.3%)	1	1	1
Q2	38 (27.3%)	1.6 (0.9 – 2.9)	1.6 (0.9 – 2.8)	1.6 (0.9 – 2.9)
Q3	35 (24.7%)	1.2 (0.7 – 2.2)	1.3 (0.7 – 2.3)	1.2 (0.7 – 2.2)
Q4, highest	29 (27.2%)	1 (0.5 – 1.8)	1 (0.5 – 1.8)	1 (0.5 – 1.8)

Table IX-14. (continued)

Exposure Groups	MS AHA/NHLBI n (%)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)
Migrants, by age at first migration**	**(n = 585/585)**			
≤ 12 yo at first migration	40 (17.8%)	1	1	1
> 12 yo at first migration	103 (28.6%)	1.6 (1 – 2.4)	1.6 (1 – 2.4)	1.6 (1 – 2.5)

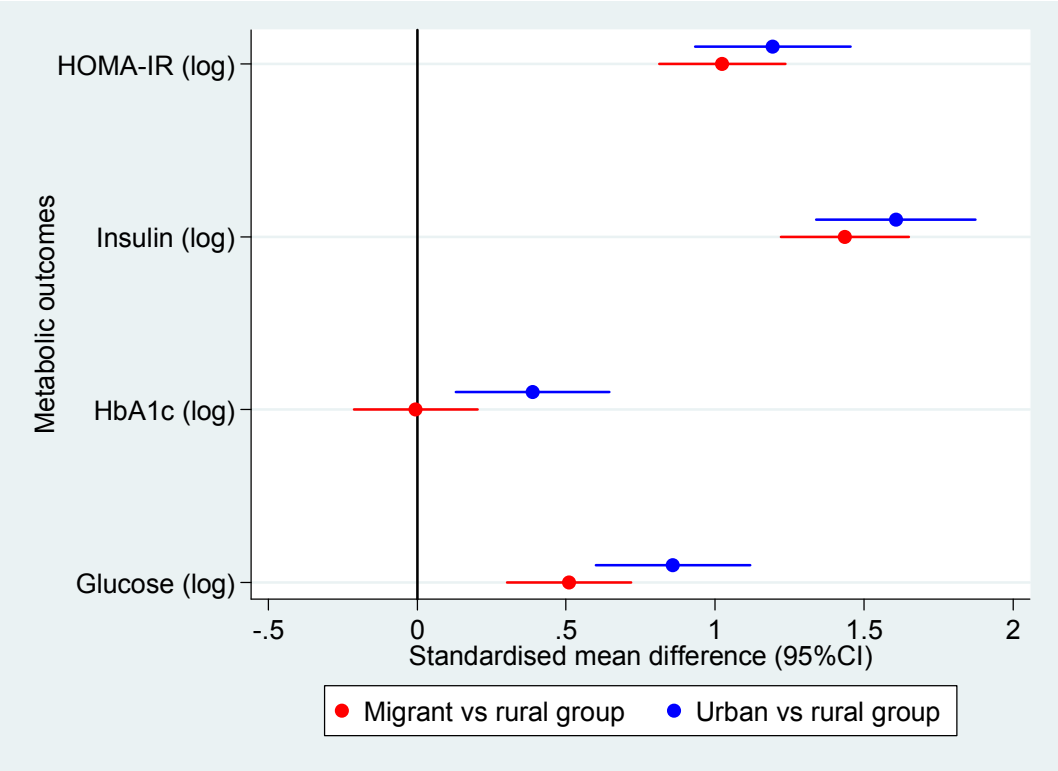
Notes: All values are OR (95% CI)
Model 1: Adjusted for age and sex
Model 2: As model 1 plus individual socioeconomic deprivation
Model 3: As model 2 plus highest parental education
* Quartiles ranges of lifetime exposure to urban area: Q1: 0 – 59.5%; Q2: 59.5% – 69.4%; Q3: 69.6% – 77.8%; Q4: 78% – 100%
** Age 12 years old was used as an indicator of puberty, to differentiate those who migrated away from Ayacucho as children from those who migrated as young adults or adults

9.5. SMD in metabolic risk factors

Figure IX-1 shows the SMD comparisons against rural group of all metabolic-related markers presented in this chapter.

Compared to rural people, the "size" of difference observed were markedly high, with a difference of more than 1 SD unit, for insulin and insulin resistance. Glucose also had an important size of difference, of 0.5 SD for migrants and nearly 1 SD unit for urban people. HbA_{1c} was only higher in the urban group but no difference was observed in migrants.

Figure IX-1. Standardised mean differences in metabolic markers in migrant and urban population compared to rural population



9.6. Discussion of results

9.6.1. Summary

A pattern of gradient of increased risk from rural to migrants to urban people was observed for fasting blood glucose, fasting insulin and insulin resistance. No such differences were observed in HbA_{1c} between migrants and rural people, but HbA_{1c} was higher in urban compared to migrants. IFG or diabetes also showed a gradient pattern and was prevalent in 1.5%, 4% and 8% of rural, migrant and urban population, respectively. In the fully adjusted model compared to the rural group, migrant and urban people were 3.5 (95% CI 0.9 – 13.4) and 8.9 (95% CI 2 –39.1) times more likely to have IFG or diabetes than rural people. A consistent gradient of increasing prevalence of metabolic syndrome from rural to migrant to urban groups was also observed.

9.6.2. Strengths

The management of blood samples that was set up in this study, using standardised protocols and trained personnel, ensure that the quality of samples provided good results. This is one of the very first studies of migration conducted in a low-middle income country that fully addresses cardiovascular risk profile. Up to this chapter, this study has comprehensively measured and reported a wide range of risk factors including behavioural, anthropometric, blood pressure, lipid profile, inflammation and metabolic-related markers. Such extensive profiling provides a wider picture of the complexity of the impact of migration on cardiovascular risk.

The findings of marked differences in nearly all metabolic-related markers including glucose, insulin and insulin resistance —but not HbA1c— provides further insights to the description of the complexity of CVD risk profile in low- and middle-income countries.

9.6.3. Weaknesses

It is very unlikely that the low median glucose observed could be linked to glucose degradation as part of a delay in transport from the rural site to the central laboratory. As recommended by WHO guidelines, all samples taken in the rural site were centrifuged soon after blood was drawn and analysed within 24-36 hours to minimise glucose degradation [159] (see Appendix G, page 441). Other studies in rural Peru have also found low prevalences of diabetes in rural area, thus supporting the observation that in Peruvian rural people mean glucose levels are lower than in people from other parts of the country [108, 109].

9.6.4. Interpretation

9.6.4.1. Glucose-related markers

It has been reported that a value of 6% in HbA1c correlates with a mean plasma glucose level of 135 mg/dL or 7.5 mmol/L of glucose [266]. The median HbA_{1c} found in this study was 5.7% in all groups. The median glucose was 80, 86 and 88 mg/dL in rural, migrant and urban groups, respectively.

In the multivariable analyses, the finding of gradient and significant differences, between migrants and urban people compared to rural population, in glucose, insulin and insulin resistance but not in HbA_{1c} needs an explanation. HbA_{1c} was 3% to 5% higher in urban compared to rural group, but no difference was observed in HbA_{1c} between migrant and rural groups. Such patterns generate the question why, if massive glucose and insulin resistance differences were observed, the same did not occur in HbA_{1c}? The explanation is more likely to be found in the physiology of such molecules, glucose and HbA_{1c}. As Goldstein et al. explain [266]:

> "the rate of formation of HbA_{1c} is directly proportional to the ambient glucose concentration. Because erythrocytes are freely permeable to glucose, the level of HbA_{1c} in a blood sample provides a history of glycaemia of the previous 120 days, the average erythrocyte lifespan… HbA1c is used both as an index of mean glycaemia and as a measure of risk for the development of diabetes complications… Between-subject variation in HbA_{1c} has been shown to be minimal in non-diabetic subjects [267, 268], and to the extent

that differences exist, they may represent differences in mean glycaemia rather than differences in glycation rates [268, 269]".

This explanation is consistent with our findings of no difference between rural and migrant groups in terms of HbA_{1c}. As the prevalence rate of non-diabetics was very low in rural and relatively low in migrants, no major differences were observed in terms of HbA_{1c} between these two study groups due to the large number of non-diabetic subjects. In the same vein, since the urban group carried a much higher proportion of diabetics, and this could explain the differences of higher HbA_{1c} observed in the urban group compared to the rural one.

The prevalence of diabetes in Peru, as informed by local studies presented in Table I-2 in page 21, appears to be around 4% depending on the study. The only discrepant information, a much higher prevalence between 17% to 19% in six Peruvian cities, comes from the Non-Communicable Risk Factors Survey which has been reported in a separate journal paper [114]. These results have not been used or endorsed by Peru's General Directorate of Epidemiology [Luis Suárez, personal communication]. The other relevant comparable study is the ENINBSC Survey [108], which did indeed reported estimates of diabetes of 4% for Lima and between 0.3% to 0.7% for the Andes. However, measurements were made using finger pick assays and diabetes was defined as fasting glucose > 100mg/dL. Compared to the available local literature, results of this study suggest that a clear gradient exists between groups in diabetes or IFG: prevalences double from rural to migrant to urban populations. Also, despite the fact that both migrants and urban population of this study live in the same area of residence in Lima, Peru's capital, the prevalences observed for these groups are quite different. These findings unmask the observation that diabetes may be similar in all groups living in cities, at least in Peru where urban population had a much worse profile for diabetes or IFG than migrants. Also surprisingly, IFG prevalence was at considerable high levels around 4% and 8% in migrants and urban groups, respectively.

The finding that those aged 12 years-old or older when migrated had higher levels of fasting glucose and were more likely to be diabetics or to have IFG than those who migrated younger is an interesting result. Related to this, migrants with longer periods of residence in an urban area, compared to those living in an urban area < 20

years, appeared to have a borderline significant gradient towards increased levels of fasting glucose and HOMA insulin resistance but not fasting insulin. Up to this point, differences in CVD studied by sub-classifications of migrants have only been found for total cholesterol and LDL. Thus, the differential CVD risk observed within migrants by time since migration or age at migration deserves an integrated discussion, including lipid and glucose differences. As such, this is jointly elaborated in Chapter XI, section 11.1.2.

9.6.4.2. Insulin resistance

In this study, the insulin geometric means of migrant and urban groups were 193% (95% CI 135% – 266%) and 251% (95% CI 167% – 362%) greater than the geometric mean of the rural group, respectively. These higher ratios were very similar to those reported in the previous chapter for CRP. After multivariate adjustment and compared to the rural group which had a fasting insulin geometric mean of 2.7 μIU/mL (95% CI 2.3 – 3.2). These levels of fasting insulin in the rural group are much lower than baseline insulin levels reported for British and US women [161, 270]. A small study in Peru, reports the comparison of insulin in 90 people from the area of San Pedro de Cajas, located at 4100 meters above sea level, and 164 people in the district of Rímac in Lima, located at sea level. Baracco et al. found that insulin levels in individuals from high altitude were lower than those from Lima (5.2 vs. 14.5 μIU/mL) [271]. Again, and surprisingly, the median levels of fasting insulin observed in this study are lower than those reported by Baracco in rural population as well as in migrants and urban people established in a different part of Lima.

Insulin resistance was calculated using HOMA calculator (Oxford Centre for Diabetes, Endocrinology & Metabolism, Diabetes Trials Unit, http://www.dtu.ox.ac.uk/) [160] and excluding those with diabetes. An interesting debate surrounds the usefulness of such model [272-274]. However, it is agreed that for epidemiological studies, the HOMA model can provide a simple tool to assess insulin resistance [160]. The HOMA model for insulin resistance has been used in cross-sectional studies [275] and in large cohort studies such as the British Women's Heart and Health Study cohort study [161] and the Women's Health Initiative Observational Study [270]. In relation to its interpretation, it follows that higher

levels of fasting insulin are related to greater insulin resistance and are potentially harmful to cardiovascular health. Lawlor et al. found that fasting insulin levels and insulin resistance, but neither glucose or glycosylated haemoglobin, were important risk factors for coronary heart disease and stroke [161].

9.6.4.3. Metabolic syndrome

Three international guidelines, the most recent and updated ones, were used for the classification of the metabolic syndrome: WHO 1999 definition [159], IDF 2005 definition [162, 163] and the AHA/NHLBI 2005 definition [164]. The last one was selected over the NCEP ATP III, initially published in 2002 [165, 276] and updated in 2004 [277]. However, such update did not make any change on the definition of the metabolic syndrome, which in turn, was later addressed by AHA/NHLBI in 2005 [164, 278].

The use of different thresholds as suggested by each specific guideline did affect the estimates reported. Of note, by lowering the blood glucose threshold from ≥ 110 mg/dL in the WHO definition to ≥ 100 mg/dL in the other two, meant that instead of 40/989 diabetics, 97/989 subjects with IFG could have been classified as having metabolic syndrome. The criteria of lipid-lowering therapy did not have any effect on the classification as none of the participants reported to be on such medications.

This aggregation of risk factors has not been free of criticism and has provoked much debate in the biomedical literature [279-286]. It is suggested that the metabolic syndrome concept can help in the identification of patients at high risk at the clinical individual care level. On the other hand, a Joint statement from the American Diabetes Association and the European Association for the Study of Diabetes raised a number of concerns around the metabolic syndrome [285]. These are related to the ambiguous criteria and thresholds used, the value of including diabetes in the definition and that treatment of syndrome is no different from treatment from each of its components, amongst other concerns [285]. A recent report shows a poor performance of the metabolic syndrome in predicting diabetes, its negligible association with risk of CVD and that the whole is no greater than the sum of the parts [287, 288].

The objective of this study was to address whether or not differences existed in various cardiovascular risk factors in rural, migrant and urban groups. In this study, a strong gradient of increasing prevalences of metabolic syndrome from rural to migrant to urban group was observed with all the definitions used. This observation was sustained within each of the classification systems.

As a separate but related issue, however, important differences in magnitude of prevalences of metabolic syndrome were also noted, particularly compared against the WHO 1999 criteria, suggesting that the performance of different guidelines would need to be explored in detail in this and other similar populations, particularly in rural and urban groups. To assess the performance or agreement between different definitions of the metabolic syndrome was not an objective of this study and thus was not explored in detail.

Similar exercises of comparison of guidelines in the classification of subjects with metabolic syndrome have been conducted for UK [289], African [290], Asian [291] and some Latin American [292] populations. All of these studies, except the one conducted in Asian population [291], describe only modest differences in the classification of individuals using various definitions. However, in Asian countries, age-adjusted prevalences for the various definitions of the metabolic syndrome ranged from 16% to 42% in Australia, 3% to 11% in Japan, 7% to 29% in Korea and 17% to 60% in Samoa [291]. These prevalences were obtained with four definitions of the metabolic syndrome and included all the ones used in this study. Here, substantially wider variations of prevalences estimates were observed in the groups studied using three definitions, ranging from 0.5% to 8% in rural group, 3% to 30% in migrants and 7% to 42% in urban population. As Lee et al. points out, “differences in the prevalence of metabolic syndrome and its components, using the various definitions, both within and between populations, indicate that caution is required when comparing studies from different countries” [291].

Bearing in mind the different classifications used, some studies from Peru report an overall prevalence of metabolic syndrome of the order of 18% (AHA/NHLBI) in Arequipa, an Andean and Peru’s second largest city [112] and of 27.3% (NCEP ATP III) or 30% (IDF) in Lima [292]. These values are within the ranges observed for the

migrant and urban groups. However, the rural group studied had much lower prevalences of metabolic syndrome.

9.7. Summary points

- A pattern of gradient of increased risk from rural to migrants to urban people was observed for fasting blood glucose, fasting insulin and insulin resistance. No such differences were observed in HbA_{1c} between migrants and rural people, but HbA_{1c} was higher in urban compared to migrants. A consistent gradient of increasing prevalence of diabetes, IFG and metabolic syndrome from rural to migrant to urban groups was also observed.

- Compared to the rural group which had a fasting glucose geometric mean of 79.9 mg/dL (95% CI 77 –83), both migrant and urban populations had significantly higher geometric means of fasting glucose. The glucose geometric means of migrant and urban groups were 9% (95% CI 5% – 12%) and 13% (95% CI 8% – 17%) greater than the geometric mean of the rural group, respectively.

- Diabetes and IFG rates double from rural to migrant to urban groups. Diabetes was prevalent in 0.5%, 2.2% and 5% of rural, migrant and urban population, respectively. In the fully adjusted model compared to the rural group, migrant and urban people were 6 (95% CI 0.7 – 51.9) and 15.9 (95% CI 1.6 – 159.2) times more likely to be diabetics than rural people, respectively. IFG or diabetes was prevalent in 1.5%, 4% and 8% of rural, migrant and urban population, respectively. In the fully adjusted model compared to the rural group, migrant and urban people were 3.5 (95% CI 0.9 – 13.4) and 8.9 (95% CI 2 –39.1) times more likely to have IFG than rural people, respectively.

- After multivariate adjustment and compared to the rural group which had a fasting insulin geometric mean of 2.7 μIU/mL (95% CI 2.3 – 3.2), both migrant and urban populations had significantly higher geometric means of insulin. The insulin geometric means of migrant and urban groups were 193% (95% CI 135% – 266%) and 251% (95% CI 167% – 362%) greater than the geometric mean of the rural group, respectively. These high ranges were very similar to those observed for CRP in the previous chapter. In the case of HOMA insulin resistance, the geometric mean of the migrant group was 202% (95% CI 156% –

257%) greater than the geometric mean of rural group. Similarly, the geometric mean of the urban group and 245% (95% CI 181% – 324%) greater than the geometric mean of rural group.

- Using any of the definitions evaluated, a consistent gradient of increasing prevalence of metabolic syndrome was observed from rural to migrant to urban groups. Prevalence estimates of metabolic syndrome, in all main groups and migrant sub-classifications, differed importantly between classifications and were much higher with the IDF or AHA/NHLBI guidelines compared to the WHO definition.

- This study found very high OR for the association between metabolic syndrome and migration in Peru. Using the AHA/NHLBI definition, migrant and urban people were 5.1 (95% CI 2.7 – 9.6) and 8.7 (95% CI 4.2 – 17.9) times more likely to have metabolic syndrome than rural people.

- When metabolic-related outcomes were explored within migrant sub-classifications, those who migrated aged 12 years or older had a 4% (95% CI 1% – 7%) higher ratio of blood glucose compared to those who migrated younger than 12 years-old. They were also more likely to have diabetes (OR 11.9 (1.6 – 90)), IFG (OR 3.2 (1.5 – 7.1)), metabolic syndrome defined by WHO (OR 8.8 (1.1 – 68.2)), by IDF (OR 1.7 (1.1 – 2.6)) and by AHA/NHLBI (OR 1.6 (1 – 2.5)) than those who migrated younger. This is the only significant result, so far, that found out a difference by age at first migration.

- Migrants with longer periods of residence in an urban area, compared to those living in an urban area < 20 years, appeared to have a borderline significant gradient towards increased levels of fasting glucose and HOMA insulin resistance but not fasting insulin. Migrants with longer periods of time in an urban setting had geometric mean ratios ranging 3% to 5% higher for fasting glucose and 7% to 28% higher for HOMA insulin resistance.

- Compared to rural people, the "size" of differences observed for migrant and urban groups were markedly high, with a difference of more than 1 SD unit, for insulin and insulin resistance. Glucose also had an important size of difference,

of 0.5 SD for migrants and nearly 1 SD unit for urban people. HbA_{1c} was only higher in the urban group but no difference was observed in migrants.

Chapter X. Cardiovascular disease risk

Following an extensive presentation of various outcomes related to cardiovascular disease (CVD), this final result chapter deals with aggregation of risk factors as a single outcome.

Cardiovascular epidemiology, particularly observational epidemiology, has made important progress in the field of risk prediction. Making use of widely available data from the population distribution of risk factors and strong outcomes such as mortality, myocardial infarction or stroke, various risk-prediction instruments have been developed and are widely used, particularly in the assessment of individuals at the clinical end of health-care delivery. Furthermore, some guidelines recommend the assessment of these risk factors following these risk-prediction instruments as an informative tool during the decision-making process of providing or prescribing treatment to individuals. This is part of the rationale to present this chapter.

This chapter will enable the assessment of the last specific research question: what are the specific CVD risk burdens on each of the study populations? Although the study of burden of disease requires a much larger venture including a long-term follow-up, this study is capable of addressing such issue partially through the assessment of a related question: to what extend are the groups studied on risk? For the latter, various risk-scoring instruments are available that provide, by taking into account an array of common CVD risk factors, simple and quick ascertainment of whether or not an individual falls into a high-risk for CVD category.

Although none of them is a perfect instrument, they can be informative. This chapter thus presents the aggregation of risk factors using three different scoring systems: the risk-scoring instruments developed by the World Health Organization (WHO) / International Hypertension Society (ISH), the *Lancet*'s Chronic Disease Action Group and Framingham Heart Study. The first two risk-scoring charts were selected because they were specifically designed for the Latin American region or for low- and middle-income countries (LMIC) [18, 166, 167]. The Framingham risk scoring

was also used for comparative purposes, and because of its geographical vicinity, it is widely used in the Latin American clinical setting. The risk scoring systems are described in more detail below.

This chapter will only address the extent of people at risk following these risk-scoring instruments using descriptive statistics. It does not present multivariable analyses on each of the scoring aggregation systems as they were never designed for such use. This chapter does not compare the performance of one scoring system against each other, and does not present a disaggregated evaluation of risk-scoring within migrants as in previous chapters.

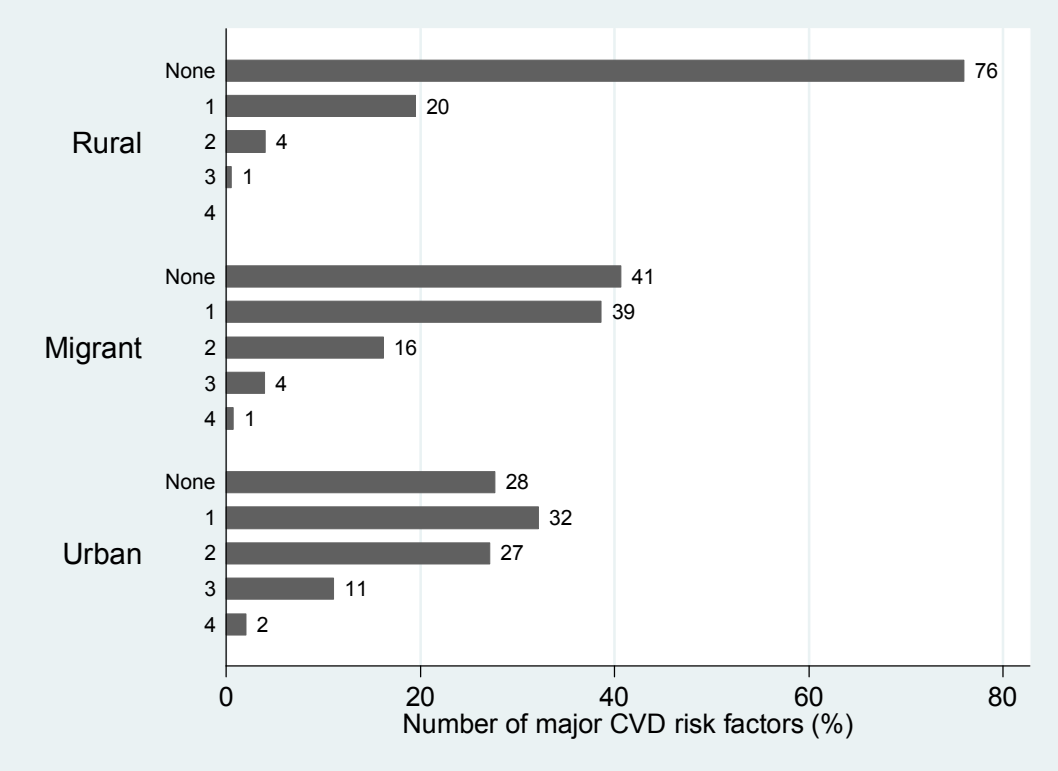
10.1. Aggregation of major cardiovascular risk factors

Before presenting the distributions of people at high-risk of having a CVD event, it is important to present the simple aggregation of major risk factors by study groups.

Figure X-1 shows the percentage distribution of individuals with a number of risk factors. Major risk factors counted in this aggregation were smoking, hypertension, diabetes and obesity as previously defined (Section 2.3.3, page 39) as well as hypercholesterolaemia (defined as total cholesterol ≥ 200 mg/dL or ≥ 5.2 mmol/L). The aggregations shown in Figure X-1 correspond to the sum of "Yes" of each individual risk factors and it ranges from zero, no risk factors, up to 4, presence of four risk factors concomitantly.

Remarkably, 77% of the rural population sampled did not have any of the major CVD risk factors and that proportion halves in the other groups. The clustering or aggregation of a number of risk factors is more apparent in migrant and urban groups and follows a gradient. Amongst migrant and urban groups, the proportion of people with two risk factors was 16% versus 28% while those with three or more risk factors were 4% and 10%, respectively.

Figure X-1. Number of major cardiovascular risk factors by study group



Note: Major risk factors counted were smoking, hypertension, diabetes, obesity and hypercholesterolaemia.

10.2. Classification of high-risk individuals

10.2.1. Risk-scoring instruments

The definitions of high risk have been presented in Table II-13 (page 63). The WHO/IHS have recently published CVD risk prediction chart for LMIC [166, 167]. These charts have been developed for each WHO geographic region, and the chart AMR-D (Americas region, mortality strata D: high child mortality and high adult mortality) was used as it corresponded for Peru [166]. WHO/ISH risk chart predicts 10-year risk of a fatal or non-fatal cardiovascular event by gender, age, systolic blood pressure (SBP), total blood cholesterol, smoking status and presence or absence of diabetes mellitus. It classifies individuals according to the following risk levels: <10%, 10% to <20%, 20% to <30%, 30% to <40% and ≥ 40% 10-year risk. High-risk was defined as those with a 10-year risk > 20% based on WHO/IHS risk prediction charts for low- and middle-income countries [166, 167].

In 2007, the *Lancet*'s Chronic Disease Action Group published a series of papers on the impact of chronic non-communicable diseases in LMIC [15-21]. One of these, a paper by Lim et al. evaluated specifically the number of deaths between 2006 and 2015 that could be averted and the financial cost of scaling up a multidrug regimen for prevention of CVD in a selection of LMIC [18]. For this, Lim et al. developed country-specific risk charts, hereafter the *Lancet*'s risk-chart, that relied on easily measurable risk factors, including age, sex, body mass index (BMI), SBP and smoking status. Specifically, a country risk-chart was available for Mexico, *inter alia*, it made its application to the present study more realistic. In this, high-risk of fatal ischaemic heart disease or cardiovascular event was defined as 10-year risk ≥ 15% [18].

For comparison purposes, the Framingham high risk of coronary heart disease, defined as a 10-year risk ≥ 20% and derived using the Framingham equation [168], was also included in the analyses. Average risk estimates are based on typical Framingham subjects, a community sample of white subjects drawn from a suburb

west of Boston, and estimates of idealized risk are based on optimal blood pressure, total cholesterol or LDL, HDL, no diabetes, and no smoking [168, 293].

Both, the WHO/ISH and the *Lancet*'s charts calculate CVD risk for individuals aged 40 to 79 years-old, whilst Framingham does so for people aged 30 to 74 years-old. All of the system use age, sex, BP and smoking status. The *Lancet*'s risk-chart does not use lipid information but takes advantage of BMI. Both WHO/IHS and Framingham risk scoring also take into account the presence or absence of a diagnose of diabetes mellitus.

10.2.2. Data completeness in risk classification

Table X-1 shows the breakdown of subjects included in risk estimation according to the scoring system used. 695/989 (70.3%) and 696/989 (70.4%) individuals were classified either as having a high or non-high risk using the WHO/IHS and the *Lancet*'s charts, respectively. In both, 292/989 (29.5%) subjects were excluded because of the age range: 282 were younger than 40 years-old and 10 were 80 or older. Additionally 1/989 individual had no information on BP and 1/989 individual did not have information on blood glucose or diabetes mellitus, the latter only affecting the WHO/IHS classification. In the case of Framingham, risk status was calculated in 955/989 (96.6%) subjects, excluding 32/989 (3.2%) because they were aged 75 or older, one had no data on diabetes and one no data on BP.

Risk status was codified following aggregation indicated by the WHO/IHS or *Lancet*'s graphic charts that use colours for risk assessment. For Framingham, the equation of risk calculation was used as it enables the calculation of risk as a continuous variable [168] that later on was recoded into a categorical one.

Table X-1. Number of subjects included and excluded from risk estimation for cardiovascular disease

	WHO/HIS n	*Lancet*'s n	Framingham n
Total n	989	989	989
Total included	695/989	696/989	955/989
Total excluded	294/989	293/989	34/989
Reasons for exclusion			
Outside age range	292/294	292/293	32/34
No data on BP	1/294	1/293	1/34
No data on diabetes	1/294	–	1/34

10.2.3. Proportion of individuals at high-risk

The proportion of individuals classified as being at high-risk of developing a CVD event varies importantly depending on the risk scoring system used (Table X-2). A pattern of increased gradient of high-risk status from rural to migrant to urban was only observed with the Framingham scoring.

The other two scoring systems provided discordant patterns of proportions of high-risk groups. On one hand, the WHO/IHS charts yielded similar proportions for migrants and urban groups, around 6%, both of them higher than rural people (1%). On the contrary, the *Lancet*'s chart estimated very low proportions, 1 to 3% of high-risk individuals in both rural and migrants, and a higher proportion of high-risk individuals in the urban group (7%).

The overall proportion of individuals at high-risk to develop a CVD event for all study groups was much higher with the Framingham risk score, approximately 13%, compared to 5% to 3% with the WHO/IHS or *Lancet*'s risk prediction charts, respectively.

Table X-2. Distribution of high-risk categories for cardiovascular disease by study group

	WHO/HIS n (%)	***Lancet*'s n (%)**	**Framingham n (%)**
n	695/989	696/989	955/989
Rural	1 (0.7%)	4 (2.9%)	12 (6.3%)
Migrant	26 (6.2%)	4 (1%)	72 (12.6%)
Urban	9 (6.4%)	10 (7.1%)	37 (19.3%)
p for trend	0.03	<0.001	0.001
Total	**36 (5.2%)**	**18 (2.6%)**	**121 (12.7%)**

10.3. Discussion of results

10.3.1. Summary

Seventy-seven percent of the rural population sampled did not have any of the major CVD risk factors —smoking, hypertension, diabetes, obesity or hypercholesterolaemia— studied. A pattern of gradient of increased risk from rural to migrants to urban people was observed for CVD high-risk status. The proportion of individuals classified as being at high-risk of developing a CVD event varies importantly depending on the risk scoring system used. The overall proportion of individuals at high-risk for all study groups was much higher with the Framingham risk score, approximately 13% overall, compared to 5% to 3% with the WHO/IHS or *Lancet*'s risk prediction charts, respectively.

10.3.2. Validity of risk scoring systems

In this document, several separate specific CVD risk factors have been presented, explored and discussed. The main purpose of such detailed and individualistic assessment has been to answer the study's main question, oriented to find out if there is a difference in specific CVD risk factors in the rural-to-urban migrant group compared to those who did not migrate. The rich information generated from migrant and non-migrant groups provide a unique scenario to address that question. But also, that information can be in turn aggregated using standard tools for the assessment of overall cardiovascular risk. It is acknowledged that all risk-scoring strategies available are far from being criticism-free. Despite criticisms and limitations, however, their use for internal comparison may be appropriate. It can be argued that the instrument selected carry internal validity, that is, the errors are unlikely to be differential by comparison group.

Many techniques for assessing the CVD risk status of individual patients have been described [168, 262, 293-302]. The European Guidelines on CVD prevention [235] uses a different model for total risk estimation based on the SCORE (Systematic

Coronary Risk Evaluation) system, derived from a large dataset of prospective European studies [262]. It is argued that

> "the great strength of the risk scoring approach is that it provides a rational means of making decisions about intervening in a targeted way, thereby making best use of resources available to reduce cardiovascular risk… Risk scoring moves the focus of treatment from the management of individual risk factors to the best means of reducing an individual's overall risk of disease. It enables the intensity of interventions to be matched to the degree of total risk" [303].

Risk scores are usually developed using a modelling approach. For this, a set of individual-level CVD risk factor profiles (age, sex, systolic blood pressure, total cholesterol and the presence or absence of type 2 diabetes) are generated using information on the population distribution of these risk factors from various large sources. For example, the WHO/ISH as well as the *Lancet*'s charts were developed based on information available from WHO Comparative Risk Assessment study [304] as well as WHO InfoBase [305]. On the contrary, Framingham takes advantage of a relatively large follow-up over decades of the same group of individuals [168, 293].

The additional advantage of risk charts released recently by WHO/IHS in 2007, is that they were developed specifically at the regional level [306, 307]. In this study, the chart relevant for the Latin America region was used. The *Lancet*'s chart was developed "through a microsimulation model used to create for each country a series of 10 000 individual life histories for each 5-year age-group and sex-group over the period 2006 to 2015. This simulation was done using information on the population distribution of risk factors, correlations between risk factor levels, associations between risk factors and disease, and population-level estimates of ischaemic heart disease, cerebrovascular events, and other mortality" [18]. The additional advantage of this work is that it developed country-specific charts and a table for Mexico was available [18]. Thus, instead of using a regional risk-assessment tool, the use of CVD risk chart from another Latin American country such as Mexico may be more relevant. In the same vein, Framingham risk equations were also selected for comparison purposes, in order to have a clearer perspective of the performance of the newly developed risk charts compared to one widely used.

Thus, in this chapter, newly developed risk-assessment tools have been used. This analysis allows the evaluation of categories of high-risk for CVD in the groups studied. This exercise enables the combined used of data gathered to, certain extent, evaluate this study's last specific research question: what are the specific CVD risk burdens on each of the study populations?

10.3.3. Limitations of risk scoring systems

It is widely known that Framingham risk prediction has limitations and its performance is not as accurate in certain groups, as demonstrated by several publications including British men [216], British older women [308] and British black and ethnic minorities [301]. Risk scores using the Framingham equations have been widely tested in North American and European populations of European origin [297, 309], and validated in a Chinese population [310], but not in other populations. Risk scores have different accuracy in different populations, tending to overpredict in low-risk populations and underpredict in high-risk populations [303].

Under these circumstances, the Framingham risk score system —and even the other risk charts used in this chapter— may be deemed not applicable to this study. Such concern arises primarily because the groups studied were not represented in the elaboration of such risk prediction models. Most certainly, an intrinsic weakness of the risk-assessment approach is linked to the fact that the tools used may not perform well in the populations studied, particularly acknowledging the low risk profile observed in the rural group in this study. As suggested by Mendis et al., the accuracy and predictive value of current risk prediction charts could be improved as more epidemiological data become available from individual countries [167]. Also, as recently noted, all attempts to make risk tables more accurate are necessary but the use of risk tables treatment in decision-making processes in practice remains a key problem [311].

A wide debate exists between the population-wide versus the high-risk approach to health. The population-wide shifting of risk factors to a much healthier level, proposed by Geoffrey Rose [312, 313], is recommended as an important public health goal as much as a prevention strategy. Conflicting evidence is available on the

effectiveness of such an approach [314, 315], which makes it difficult to ignore the high-risk approach despite its limitations. Indeed, WHO, as an international technical agency, has debated the pros and cons of a population-wide versus the high-risk approach in various technical documents [167, 303, 316]. Despite the limitations of the high-risk stratification, it is worth noticing that the "total risk approach acknowledges that many cardiovascular risk factors tend to appear in clusters, [thus] combining risk factors to predict total cardiovascular risk is consequently a logical approach to deciding who should receive treatment" [303].

10.3.4. Interpretation

It was noticeable that 77% of the rural population did not have any of the major risk factors usually assessed in risk scoring systems. In contrast to this, the proportion of people with two risk factors was 16% and 28% while those with three or more risk factors were 4% and 10%, in migrant and urban groups, respectively. The marked differences in the distribution of risk factors in the groups studied will certainly have an impact in the development of CVD events and may differ from those estimations based on risk scores. Of note, and as shown in previous results chapters, major differences in BMI between groups were observed while BP did not show a clear gradient of difference between migrants and non-migrants. Also, overall, smoking prevalence was quite low and diabetes prevalences averaged 3%. This profile distribution may challenge the performance of selected risk estimation systems used in this chapter.

The overall proportion of individuals at high-risk of a CVD event for all study groups was much higher with the Framingham risk score, approximately 13%, compared to 5% and 3% with the WHO/IHS or *Lancet*'s risk prediction charts, respectively. Although it is not an objective of the study to evaluate the agreement between risk prediction systems, this study clearly demonstrates that the performance of various scoring systems differ substantially.

PREVENCION, the other available population-based study in Peru in a different urban area located in the Andes [111, 112, 123], has also reported CVD risk estimations based on Framingham risk. They found an overall prevalence of 84%,

10% and 6% for 10-year risk <10%, 10 to 20% and > 20 %, respectively [317]. Corresponding figures of Framingham risk categories in this study were 60%, 27% and 13% for all study participants and 75%, 19% and 6% in the rural group only. In the present study, both migrant and urban groups had worse classification than individuals in the PREVENCION study. Data from these two studies from Peru could suggest that Framingham risk equations may not be applicable for a country such as Peru.

Due to the lack of prospective data in Peru, it is difficult to ascertain whether Framingham, and indeed the other two scoring systems, overpredict or underpredict risk in this population. Due to the varied pattern of distribution of risk factors between groups demonstrated in this study, it remains a challenge to develop better specific risk-prediction tools for these settings. The challenge also persists as to the applicability of risk prediction charts in developing countries. As stated by WHO, "further research is required to validate existing sub-regional risk prediction charts for individual populations at national and local levels, and to confirm that the use of risk stratification methods in LMIC countries results in benefits for both patients and the health care system [303]".

10.4. Summary points

- Seventy-six percent of the rural population sampled did not have any of the major CVD risk factors —smoking, hypertension, diabetes, obesity or hypercholesterolaemia— studied.

- The clustering or aggregation of a number of risk factors is more apparent in migrant and urban groups and follows a gradient. Amongst migrant and urban groups, the proportion of people with two risk factors was 16% versus 28% while those with three or more risk factors were 4% and 10%, respectively.

- The proportion of individuals classified as being on a high-risk status of developing a CVD event varies importantly depending on the risk scoring system used.

- A pattern of increased gradient of high-risk status from rural to migrant to urban was only observed with the Framingham scoring. WHO/IHS and *Lancet*'s risk scoring systems provided discordant patterns of proportions of high-risk groups.

- The overall proportion of individuals at high-risk to develop a CVD event for all study groups was much higher with the Framingham risk score, approximately 13%, compared to 5% to 3% with the WHO/IHS or *Lancet*'s risk prediction charts, respectively.

Chapter XI. Discussion of findings

The introduction of this thesis began with a straightforward message: “migration is one example of social and cultural change”. Urbanisation, aided by migration, is a major feature of today’s world [48]. In this context, this and other migration studies become relevant and crucial tools to address the impact of such complex processes on health. This background served as the basis to address to what extent migration may have had an impact on cardiovascular disease (CVD) risk factors in Peruvian population. The exciting findings deriving from this study thus provides a much clearer panorama, previously unknown, of the health profile of rural-to-urban migrants within Peru in relation to CVD risk factors. This information in turn, will prove useful for the understanding of CVD in Peru, and to certain extent in other low- and middle-income countries (LMIC) undergoing similar process. It is hoped that, later on, the information generated in this study can inform key policy- and decision-makers in the design and implementation of preventative strategies.

This study was designed to address a major overall and few specific research questions (Section 2.1.4). Herein, in order to maintain an organised structure in the reporting, the same research questions are outlined in this chapter and overall answers elaborated upon. In so doing, the answers provided aim to bring together the data presented in each specific result chapters in a single response framework. The evaluation of individual prevalence rates found in this study compared to other published data has already been approached in each specific chapter and will not be discussed in this chapter.

Later on, a discussion on the overall strengths and limitations of the study are presented as well as the overall implications derived from this study.

11.1. Overall summary of main findings

This section is organised according to each of the research questions (Section 2.1.4) that guided the present study. These were:

Overall research question

i) is there a difference in specific CVD risk factors in the rural-to-urban migrant group compared to those who did not migrate?

Specific research questions

Does the pattern of CVD risk factors in the migrant population vary by

ii) length of residence in urban environment?

iii) lifetime exposure to urban environment?

iv) age at first migration?

And, additionally

v) what are the specific CVD risk burdens on each of the study populations?

11.1.1. Answer to research question 1

Is there a difference in specific CVD risk factors in the rural-to-urban migrant group compared to those who did not migrate?

The single major overall research question of this study was oriented to find out if there was a difference in specific CVD risk factors in the rural-to-urban migrant group compared to those who did not migrate. It was hypothesized that the risk of CVD increases following migration from rural to urban areas in Peru. To address this question, analyses of categorical and continuous outcomes were carried out using the rural group as baseline group. Comparisons were made between migrant-to-rural

group and urban-to-rural groups using, in most cases, information from all study participants (n = 989).

The answer to this research question is "Yes, mostly". Using the rural group as baseline, this study demonstrates that differences in migrants risk profile exist in all CVD risk factors studied except for blood pressure, HDL cholesterol, triglycerides and HbA_{1c}. Furthermore, and perhaps most interestingly and relevant, the pattern of differences observed was not uniform. As postulated in the study's hypothesis, it was expected that the migrants will have increased levels of risks factors compared to rural counterparts. The selection a priori of two comparison groups, rural and urban people, was made with the intention of having two different populations serving as "margins" or "extremes". These two groups established the ranges where risk factors of migrants could fit. This decision proved useful because it paved the way for a comprehensive assessment and interpretation of the panorama of CVD risk factors in migrants. Such profiling is summarised in the next pages. Figure XI-1 (page 334) and Table XI-1 (page 335) present summary results of prevalences and OR for all categorical variables. Figure XI-2 (page 336) presents standardised mean differences (SMD) of all continuous outcomes evaluated in this study.

In the case of categorical risk factors, two patterns of risk profile were clear. First, a pattern of gradients, where increases were observed from rural to migrant to urban groups. Second, a pattern of similarities or no difference between two groups but at the same time being higher or lower, in striking contrast, with the third group. As shown in Figure XI-1, smoking, diabetes, impaired fasting glycaemia (IFG) and metabolic syndrome using WHO definition show a gradient of doubling of prevalences from rural to migrant to urban groups. On the contrary and as examples of the second pattern of risk profile, prevalence rates of overweight and obesity combined together, obesity and hypercholesterolaemia were very similar between migrant and urban populations and much higher in than rural group. Surprisingly, hypertension also had the latter profiling but in a different arrangement, i.e. hypertension rates were very similar in rural and migrant populations but lower than in urban group. Table XI-1 provides the strength of the association (OR and 95% CI) between these categorical risk factors and study groups. In some circumstances, due to the lower number of cases in the rural reference group, CI were wide or

overlapped the value of one. Nonetheless, except for smoking (OR 1.1, 95% CI 0.4 – 2.7) and hypertension (OR 1.5, 95% CI 0.8 – 2.9) that had a non-significant OR below two, migrants consistently were more likely to have or being overweight, obesity, IFG and diabetes and metabolic syndrome using WHO, IDF and AHA/NHLBI definitions, all of them with OR higher than three. Of these, only diabetes, the combination of IFG or diabetes and metabolic syndrome using WHO definition were not significant. In the case of urban compared to rural people and with the exception of smoking, urban population was consistently more likely to have all CVD risk factors with high or very high OR.

The assessment of continuous outcomes in CVD risk factors further contributed to answer the main research question of this study. Again these analyses were consistent with the study's hypothesis and provided firmer evidence of differences, without a uniform pattern, in risk profile in migrants in some but not all risk factors studied. More importantly, the use of SMD, using the rural group as baseline for comparisons, allowed a simple and clear way to present magnitude of the difference or 'difference sizes' across risk factors. In so doing, the profiling of risk factors analysed as continuous traits enabled the observation of mirroring or very similar patterns, gradient patterns or no difference in the groups studied together with their difference sizes.

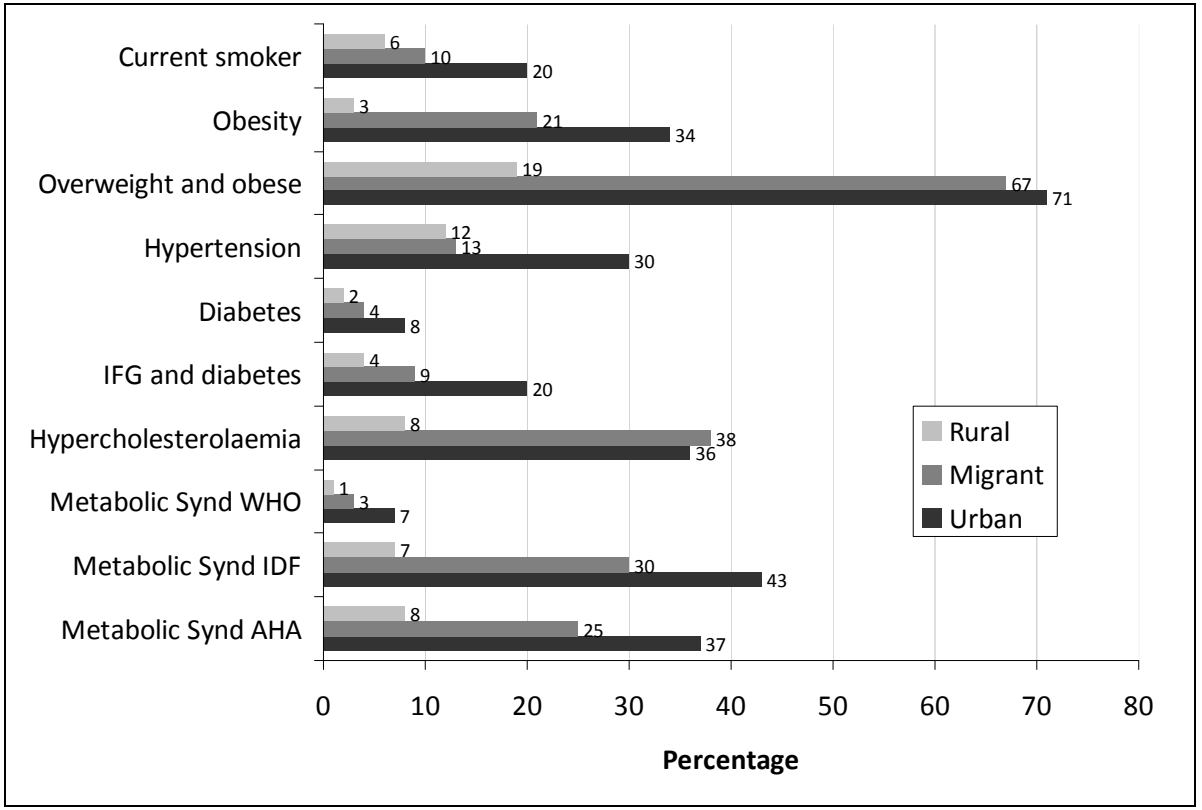
As shown in Figure XI-2, the distribution of lipid profile and inflammation markers in migrants mirrored almost exactly the urban one. Anthropometric risk factors as well as glucose-related and insulin-related traits showed a trend towards a gradient pattern, with migrants having lower means than urban group, although CI overlapped in all of them. Surprisingly, the same gradient —migrant lower than urban— was also observed in systolic blood pressure (SBP) and diastolic blood pressure (DBP). However, when compared to rural group, the migrant group had similar SBP but lower DBP than rural group while the urban group had higher SBP but similar DBP than rural group.

Cochrane recommendations indicate that SMD higher than 0.5 provide evidence of a moderate effect and 0.8 indicates a large effect [175]. Thus, and as presented in Figure XI-2, difference sizes were in the range of moderate to high for some risk

factors, e.g. differences between 0.5 to 0.8 SD for body mass index (BMI), waist-to-hip ratio, total cholesterol, CRP, fibrinogen and glucose. Surprisingly, some risk factors presented SMD higher than 1 SD, thus indicating that their difference sizes were considerably large as in the case of waist circumference, skinfolds, total cholesterol/HDL cholesterol ratio, LDL cholesterol, insulin and insulin resistance. Aside from the very specific pattern of differences described earlier for BP in both migrant and urban population compared to the rural group, HDL did not show any difference in all three groups studied. Triglycerides were slightly higher in both migrant and urban population than in rural group but their CI overlapped or were near zero. No difference in HbA_{1c} was found between migrant and rural populations but urban people had higher mean HbA_{1c} levels.

In summary, differences were found in most but not all CVD risk factors studied and they did not follow a uniform pattern. While in some cases gradients were observed, in others migrants mirrored their urban counterparts.

Figure XI-1. Prevalence of cardiovascular disease risk factors by study groups



Notes: Hypercholesterolaemia (total cholesterol ≥ 200 mg/dL or ≥ 5.2 mmol/L) was not a major categorical outcome in this study, and thus OR were not calculated. However, for informative purposes, it is presented in this figure alongside other major categorical risk factors. *p* for trend in all cases <0.01.

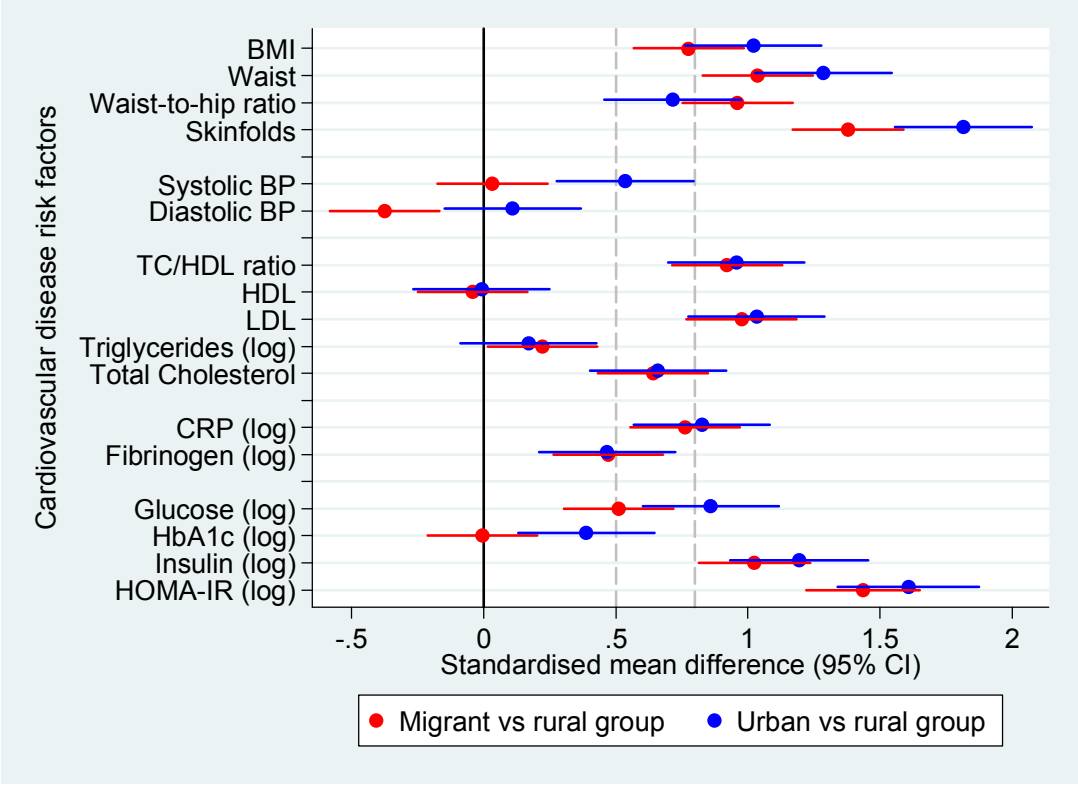
Table XI-1. Association between cardiovascular risk factors and migration

Risk factor	Rural (baseline)	Migrant OR (95% CI)	Urban OR (95% CI)
Smoking	1	1.1 (0.4 – 2.7)	2.5 (1 – 6.7)
Obesity	1	9.5 (3.8 – 23.4)	20.1 (7.6 – 53.3)
Overweight and obese	1	5.9 (3.7 – 9.4)	5.7 (3.2 – 10.2)
Hypertension	1	1.5 (0.8 – 2.9)	5 (2.3 – 10.6)
Diabetes	1	6 (0.7 – 51.9)	15.9 (1.6 – 159.2)
IFG and diabetes	1	3.5 (0.9 – 13.4)	8.9 (2 – 39.1)
MS WHO	1	5.4 (0.6 – 47.1)	13.8 (1.4 – 134.2)
MS IDF	1	7.2 (3.7 – 14)	12 (5.7 – 25.3)
MS AHA/NHLBI	1	5.1 (2.7 – 9.6)	8.7 (4.2 – 17.9)

Notes: The following definitions were used:

- Smoking: Has smoked more than 100 cigarettes in lifetime and last cigarette was less than 6 months ago
- Obesity: BMI $\geq$ 30 Kg/m^2
- Overweight or obese: BMI $\geq$ 25 Kg/m^2
- Hypertension: SBP $\geq$ 140 mm Hg or DBP $\geq$ 90 mm Hg, or self report of physician diagnosis and currently receiving antihypertensive medication
- Diabetes: Fasting plasma glucose $\geq$ 126 mg/dL (or $\geq$ 7 mmol/L) or self report of physician diagnosis and currently receiving antidiabetic medication
- IFG (Impaired fasting glucose) or diabetes: Fasting glucose $\geq$ 110 mg/dL (or $\geq$ 6.1 mmol/L)
- MS WHO: Metabolic syndrome using WHO 1999 definition [159]
- MS IDF: Metabolic syndrome using IDF 2005 definition [162, 163]
- MS AHA/NHLBI: Metabolic syndrome using AHA/NHLBI 2005 definition [164]

Figure XI-2. Standardised mean differences in all cardiovascular disease risk factors in migrant and urban population compared to rural population



Notes: The solid line at zero indicate no difference compared to rural group. Additional dashed lines at 0.5 and 0.8 correspond to thresholds for moderate and large differences, respectively

11.1.2. Answer to research question 2, 3 and 4

Does the pattern of CVD risk factors in the migrant population vary by length of residence in urban environment?

Does the pattern of CVD risk factors in the migrant population vary by lifetime exposure to urban environment?

Does the pattern of CVD risk factors in the migrant population vary by age at first migration?

Three separate specific research questions were also part of this research. These explored whether the pattern of CVD risk factors amongst migrants vary by ii) length of residence in urban environment, iii) lifetime exposure to urban environment, and iv) age at first migration.

The distribution of migrants according to length of residence in urban area —either as absolute number of years or proportionally as lifetime exposure to urban area— and age at first migration was presented in Table V-6, Table V-7 and Table V-8, respectively. All these three specific research questions used information only from the migrant group. Comparisons were made using as baseline the lowest category of exposure created after the sub-classification of the migrant group. Research questions ii) and iii) were evaluated with information available from 559/589 (95%) of migrants, whereas research question iv) was evaluated using information from 585/589 (99%) of migrants. Such losses of information occurred because not all migrants provided information on either number of years living on urban area or age at first migration needed to proceed to migrant's group sub-classification.

The answers to these questions are not as straightforward as the answer to the overall research question. No consistent pattern of variation of CVD risk factors was observed using migrants' sub-classifications. Three exceptions, however, involving separate sub-classifications, were observed with lipid profile and metabolic-related markers.

First, migrant with 20 or more years living in an urban area had in average 15 mg/dL units of total cholesterol (Table VIII-8, page 238) and around 12 mg/dL units of LDL (Table VIII-10, page 240) higher than those with less than 20 years in urban area. No clear pattern of difference was observed in these traits using other migrant sub-classifications or in other lipid markers.

Second, in the case of metabolic-related outcomes and compared to those who migrated younger than 12 years-old that served as the baseline group, those who migrated aged 12 years or older had or were more likely to have:

- 4% (95% CI 1% – 7%) higher geometric mean ratio of blood glucose, Table IX-2, page 272.
- Diabetes, OR 7.5 (95% CI 0.9 – 58.7), Table IX-5, page 277.
- IFG or diabetes, OR 6.4 (95% CI 1.5 – 27.8), Table IX-6, page 280.
- Metabolic syndrome defined by WHO, OR 8.8 (95% CI 1.1 – 68.2), Table IX-12, page 294.
- Metabolic syndrome defined by IDF, OR 1.7 (95% CI 1.1 – 2.6), Table IX-13, page 296.
- Metabolic syndrome defined by AHA/NHLBI, OR 1.6 (95% CI 1 – 2.5), Table IX-14, page 298.

Third, migrants with longer periods of residence in an urban area, compared to those living in an urban area < 20 years, appeared to have a borderline significant gradient towards increased levels of fasting glucose and HOMA insulin resistance but not fasting insulin. Migrants with longer periods of time in an urban setting had geometric mean ratios ranging 3% to 5% higher for fasting glucose (Table IX-2, page 272) and 7% to 28% higher for HOMA insulin resistance (Table IX-9, page 288).

In the case of lipids, the observations raise an interesting possibility that with increasing length of migration lipids also increase. Such observation is plausible, if taken together with results presented before that lipid profile of migrants tend to

"mirror" those of the urban people. By analysing all migrants together, a gradient by length of migration could have been masked, at least for lipid profile. A likely explanation for this involves the potential role of diet. How migrant's diet becomes more similar to urban groups by length of migration has not been addressed in this study and remains unanswered.

In the case of differences observed in metabolic-related outcomes it is clear that those are primarily driven by glucose which, in turn, have a direct contribution to the classification of diabetes, IFG and metabolic syndrome. An age of 12 years-old was selected a priori as a cut-off to separate younger migrants from those about to enter adolescence. The findings in relation to age at migration benefit from concepts more familiar to life-course epidemiology: programming and (des)adaptation. The literature is quite prominent on this case, particularly in establishing the long term effects of early exposures, ranging from intrauterine stages to childhood, on the development of CVD later in life [318-332]. It could be postulated that a greater exposure to a healthier rural environment for most of childhood could indeed exert a long-term programming towards a healthier state. This issue was partially addressed by sub-classifying migrants according to lifetime-exposure to urban environment and no differences were observed in any of the outcomes studied. In the same vein, it can also be postulated that those with a longer exposure to a rural environment can have higher degrees of adaptation to that specific environment and are less likely to re-adapt when exposed to a different environment. This is coherent with the findings in relation to glucose trait and age at migration. Those who migrated younger were more likely to physiologically adapt to a new urban setting and to sustain healthier levels of glycaemia, expressed by the low mean glucose level they presented as a group and lower rates of diabetes or IFG. On the contrary, and plausibly interesting, those who migrated at later ages were less likely to re-adapt to the newer urban environment and thus established higher mean levels of glycaemia as baseline that in turn explained higher levels of diabetes and related outcomes. If the borderline differences observed for fasting glucose and HOMA insulin resistance by length of residence in urban area mentioned above were not to be chance findings, a similar explanation related to a lack of complete adaptation to an urban environment could also be postulated.

Although inconclusive, these findings by specific features of migration opens and raises further question on the role and long-term impact of migration on the development of chronic conditions.

11.1.3. Answer to research question 5

What are the specific CVD risk burdens on each of the study populations?

The last research question proposed for this study was v) what are the specific CVD risk burdens on each of the study populations? This question was addressed in Chapter X. Risk-scoring instruments developed by the World Health Organization (WHO) / International Hypertension Society (ISH), the *Lancet*'s Chronic Disease Action Group and Framingham Heart Study were used.

The answer to this question would be that a pattern of gradient of increased risk from rural to migrants to urban people exists for CVD high-risk status. CVD risk burden is high in urban group and low in rural group. Migrant population overall, depending on the instrument used to assess burden, have a risk profile in between the reference groups. As presented and discussed before, migrants have a lipid profile that mirrors the urban group, higher anthropometric, inflammation and metabolic traits when compared to rural people but show no difference in terms of BP. Aggregating all these patterns for a "comprehensive" assessment of migrant's burden in terms of CVD risk will depend on how much weight is provided to individual traits. The distribution of individual CVD risk evaluated in migrants and non-migrants in the population studied does indeed challenge the performance of any risk estimation systems to be used.

Before any aggregation into any scoring system and just by counting individual number of risk factors, it was noticeable that 96% of the rural population did not have any (77%) or had one (19%) of the major risk factors usually included in the risk scoring systems and only 5% had two risk factors. In stark contrast, the proportion of people with two risk factors was 16% and 28% while those with three or more risk factors were 5% and 12%, in migrant and urban groups, respectively.

When individual traits were aggregated using risk scoring systems, the overall proportion of individuals categorised as "at high-risk" to develop a CVD event for all study groups was much higher with the Framingham risk score, approximately 13%, compared to 5% to 3% with the WHO/IHS or *Lancet*'s risk prediction charts, respectively.

These observations put together, that migrants do not necessarily resemble on very specific high- or low-risk group, place a further challenge on the epidemiology of CVD in LMIC. Much long-term work needs to be done, particularly in today's LMIC societies, to understand the intricacies of the impact of migration on CVD.

11.2. Contextualising findings

Findings obtained from this study need to be interpreted in a contextualised manner with particular reference to other published literature on cardiovascular risk in LMIC. A brief section about other studies has been presented previously in the introductory chapter, e.g. Section 1.3, page 8, and some relevant studies have been presented and discussed within each individual results chapter.

However, this section includes a critical consideration of how findings in general compare to those from other migration studies, and to literature on rural-urban differences in cardiovascular risk. Of note, migrant studies do not necessarily constitute a rural-urban study as migrants may not originate from a rural setting. This specific study reported in this document does indeed constitute a migrant and, at the same time, a rural-urban study as migrants originated in a rural area, were evaluated in an urban area and there were two comparison groups, namely rural and urban.

In this section, as well as throughout this document, the focus is on internal or within-country migration. Thus literature related to international migration was not considered. Two of the most representative migrant studies, especially the Chinese Yi Migrant and the Kenyan Luo migration studies, are referred in this section. The literature from other LMIC settings is much more prominent in the case of rural and urban differences, and they have been included in the respective sections.

11.2.1. Blood pressure

The results for blood pressure showed no clear consistent patterns comparing the three groups. Reported findings in the international literature are similarly inconsistent. For example, SBP in rural and migrant Chinese in the Yi People Study was 105 and 112 mm Hg, respectively [84-86]. Remote indigenous populations from Brazil, Papua New Guinea and Kenya that took part on the INTERSALT study had a combined SBP, average of four centres, of 103 mm Hg compared to 120 mm Hg in the remaining 48 INTERSALT centres [242, 243]. On the contrary, the Kenyan Luo migration study found a mean SBP of 121 and 112 mm Hg in rural males and

females, contrasted with 129 and 119 mm Hg in urban males and females, respectively [78, 244, 245]. In Cameroon, mean SBP was 125 and 119 mm Hg for rural male and female groups, and 136 and 127 mm Hg for urban males and females, respectively [83]. A rural-to-urban migration study in Iran found much higher values of mean SBP with figures of 126 mm Hg for rural people contrasted with 138 mm Hg in migrants from Azerbaijan and urban non-migrants [80].

Although it is difficult to establish the comparability between rural and migrant populations of cited examples with the respective groups of this study except for their common condition of being born in a rural place in their country of origin, the common pattern is that generally rural populations have a lower SBP than migrants or urban counterparts. Of interest, it is clear that the range of rural baseline mean SBP varies widely in each example, having the Chinese rural population and other indigenous groups with lower blood pressures overall than other groups.

In the case of DBP, the Chinese Yi Migrant Study reports a mean DBP of 66 mm Hg for rural people and 71 mm Hg for migrants and urban groups [84]. The Kenyan Luo migration study reports a mean DBP of 60 and 62 mm Hg in the rural and urban groups, respectively [78]. The INTERSALT study reports a mean DBP of 57-67 mm Hg in four remote areas compared to an average of 74 mm Hg in the 48 remaining sites [242]. On the contrary, the Iran rural-to-urban migration report a higher mean DBP of 72 versus 84 mm Hg in rural and migrant population, respectively [80]. In Cameroon, even higher mean DBP are reported for both rural (female/male 76/81 mm Hg) and urban dwellers (female/male 80/86 mm Hg) [83]. In Costa Rica, however, in a survey conducted in 1986, no significant differences in DBP were found between urban and rural people, perhaps due to small sample size (n ~ 230 in each group) [333].

The comparison with previous published studies indicates that the blood pressure pattern in rural group is the one of a low blood pressures overall and can serve as a good “healthy” baseline group for the comparisons of interest, similar to other risk factors already explored in this study. Thus, a surprising finding is that the migrant group also shows a healthy BP profile, as good as the rural group in the case of SBP or even a better profile in the case of DBP.

The observation from this study that migrants from a wide age range, after a sustained process of migration and establishment into an urban environment for a number of years, have similar SBP to their rural counterparts, to the best of our knowledge, has not been previously reported. However, similar findings of surprising decrease in BP following migration in a much younger cohort and within 6 months of migration have recently been reported in Tanzania [82]. As suggested by Unwin et al. in relation to these findings on BP, this "suggest that the pattern of change on rural to urban migration may be more complex than commonly thought and is worthy of further study" [82].

11.2.2. Hypertension

Hypertension was prevalent in 12%, 13% and 30% of rural, migrant and urban population, respectively. Such prevalences fit within the wide range of reported hypertension in Peru for various contexts (see Table I-2) with a clear pattern of lower hypertension rates in rural areas [109, 117, 118]. Again, surprisingly, migrants have a similar rate of hypertension to rural counterparts despite their considerable exposition to an urban environment. Migrants appear to have increased odds of being hypertensive than rural population but CI overlap the value 1 (OR 1.5 (95% CI 0.8 – 2.9)). The urban group, in contrast, was 5 (95% CI 2.3 – 10.6) times more likely to be hypertensive than the rural group.

Rural-to-urban migrants and non-migrants comparison groups in China, in the 1986 Yi People Study, found a very low prevalence of hypertension [86]. Hypertension, defined as SBP ≥ 140 or DBP ≥ 90 mm Hg, in those days termed "borderline hypertension", was <1% amongst rural farmers. Yi Migrants had a prevalence of 2.2% that did not differ much from county Han people with 2.5% [86]. Unfortunately, reported data from the Kenyan Luo migration study concentrated mostly on SBP and DBP and did not report rates of high blood pressure other than indicating that it was higher in urban visits [78, 244, 245]. These results from Chinese migrants differ from those found in this study, perhaps reflecting population and cohort differences between China and Peru.

In contrast to these settings, among African population higher levels of hypertension have been found in both in rural and urban sites. For example, in Cameroon, estimates of hypertension for women and men were 19% and 27% in rural dwellers and 25% and 42% in urban dwellers, respectively [83]. In the rural, poor black community in Limpopo, South Africa, hypertension was found in 26% and 22% of women and men, respectively [334]. Surprisingly, one study from Tanzania found no urban-rural differences in the prevalence of hypertension, with both sites reporting hypertension rates around 30% in both sexes [335]. Similar findings of no rural-urban difference in has been previously described in one of South Africa's states [336].

However, recent findings from South Africa's first national Demographic and Health Survey do indeed report hypertension rates by socioeconomic groups, using quintiles derived from an asset index, ranging from 20% to 30% and increasing with higher socioeconomic status [337]. In addition to this and concentrating only on the geographical setting, the South African data reports a non-significant —OR did overlap value 1— tendency towards an urban-rural difference suggesting that rural groups had lower chances of hypertension than urban groups [337]. When analysed by specific self-reported ethnic groups, this data indicates that rural blacks had a significantly lower risk of hypertension than urban black, coloured and white participants. The authors interpret these results as suggestive of a clear urban-rural difference rather than inherent differences among the population groups regarding the risk of developing hypertension. As the authors point out, "the data show that observed differences in the prevalence of hypertension are not ethnically based and can be accounted for by other socio-demographic parameter differences" [337].

Moving one step further in the assessment of the impact of urbanisation, the THUSA study [338], recruited mainly Setswana speaking people from 37 sites from the four geographical quarters of the North West Province of South Africa. Participants were divided into different levels or strata of urbanisation, from stratum 1 (rural) to stratum 5 (urbanized). Interestingly, males and females from stratum 3 showed the highest rate of hypertension and stratum 5 the lowest. The authors suggest that factors associated with urbanisation, driven to a certain extent by migration, are related to the manifestation of hypertension given the highest mean blood pressure in

people living in informal settlements, where most newcomers to the urban areas live [338].

Although the prevalence of hypertension among the urban population found in the present study is similar to the rates reported for African population, the marked urban-rural difference observed in Peruvian and Chinese populations, at least in relation to hypertension, seems, to certain extent with the exception of South Africa, to be absent in the African setting.

11.2.3. Lipid profile

In relation to lipid profile, the very first observation that comes from this study is that, overall, lipid markers in all three groups studied fit into a healthy lipid profile when compared to Western populations in developed countries of North America and Europe. All groups had a desirable level (below 200 mg/dL) for mean total cholesterol, above 40 mg/dL for HDL and optimal (below 100 mg/dL) or near or above optimal (100-129 mg/dL) for LDL [165]. European guidelines suggest that "in general, total plasma cholesterol should be below 5mmol/L (190mg/dL), and LDL cholesterol should be below 3mmol/L (115mg/dL)" [235]. Following European cut-offs, the rural and migrant groups population means are the ones within ranges proposed by European guidelines. The US National Health and Nutrition Examination Survey 2003-2004 reported the following overall mean values for lipid markers: total cholesterol 200.3 mg/dL, LDL 118.7 mg/dL, HDL 54.3 mg/dL, and triglycerides 129.5 mg/dL [246]. In contrast with these US results, the results from this study indicate that rural, migrant and urban populations have not only lower mean total cholesterol levels but also lower mean HDL levels. Migrant and urban but not rural people's mean LDL (116, 120 and 86 mg/dL, respectively) were similar to US levels (119 mg/dL). Interestingly, migrant and urban groups but not rural showed a higher mean triglyceride level (around 133 mg/dL) compared to the US population (129.5 mg/dL).

Chinese individuals have also been described as a low lipid profile population [247]. The Yi People Study, showed better profile than this study for total cholesterol, HDL and LDL but not for triglycerides (range 136 – 152 mg/dL) [84]. In the case of Costa

Rica, rural women did not show differential lipid profile from urban women, but urban men had higher total cholesterol and higher LDL than their rural men counterparts [333]. Actually, no significant differences between rural and urban women were found for any of the cardiovascular risk factors [333]. In the rural, poor black community in Limpopo, South Africa, 42% of women and 29% of men had LDL levels of 3 mmol/L (116 mg/dL) or more [334]. The South African THUSA study, conducted in black people of the North West Province aiming at disentangle the impact of urbanisation, reported lower serum lipid levels in the less urbanised strata [339].

Lipid profile is very strongly related to diet which varies between groups from different settings. Indeed, in this study it was found that the lipid profile of migrants almost mirrored the one of their urban counterparts. However, dietary intake was not addressed in this study, limiting exploration of this issue.

11.2.4. Diabetes

Diabetes in this study, defined using WHO 1999 guidelines as fasting plasma glucose ≥ 126 mg/dL (or ≥ 7 mmol/L), was prevalent in 0.5%, 2.2% and 5% of rural, migrant and urban population, respectively. The literature on urban-rural differences on diabetes shows interesting contrasts in recent years, particularly in relation to LMIC.

For example, using the same criteria, the Prevalence Of Diabetes in India Study (POSIS) [340] reports a standardized prevalence rate in India's rural and urban populations of 2.7% and 5.9%, respectively. Also from India, the Coronary Risk of Insulin Sensitivity in Indian Subjects (CRISIS) Study [341], studied rural, urban slum and urban middle-class males aged 30-50 years. They found that adiposity, waist circumference, HOMA-IR, index and both fasting and 120 min plasma glucose concentrations increased progressively from rural through to urban slum and urban middle-class men [341], which is consistent with the increased gradient from rural to migrants to urban found for glucose, insulin and HOMA-IR in this study. However, compared to this study and using the same WHO criteria, the increased gradients or "doubling" of diabetes prevalences reported in the CRISIS study were of higher magnitudes, from 0% to 6% to 10% [341].

In Tanzania, the Essential Non-Communicable Disease Health Intervention Project, found a difference of 3.8% in the prevalence of diabetes between urban and rural areas [335]. In Limpopo, a rural part of South Africa, diabetes was diagnosed in 8.8 and 8.5% of women and men, respectively [334]. In Cameroon, the prevalence of IFG and diabetes together was 15% and similar in both men and women [83], which is higher than the reported in this study and in the CRISIS study, 8% and 11%, respectively.

11.3. Overall strengths of this study

The discussion section of each of the results chapter has already elaborated a summary of the strengths of the study in relation to the specific outcomes presented. In this section, a summary of strengths is presented in relation the whole study as a single entity. As with any epidemiological study, the strengths of this study can be classified of strengths related to its design, execution and performance and approach to data analysis.

11.3.1. Study design

- Advantage taken from one unique context of long-term established migration in a LMIC.
- A hypothesis-driven study, well designed to address and clearly answer each of research questions.
- Study groups well defined to enable the evaluation of one group of interest, i.e. migrants, in relation to rural and urban counterparts providing a complete panorama of CVD risk profile.
- Classification of migrants in subgroups according to length of migration, lifetime exposure to urban environment and age at first migration —usually not reported in migrant studies— that expanded the evaluation of the impact of migration.
- Evaluation of a wide number of established CVD markers, including behavioural risk factors, anthropometrics, blood pressure, lipid profile, inflammation and metabolic traits.
- Standard definitions used that enables comparability of data with other resources from LMIC.

11.3.2. Study execution and performance

- Conduction of a pilot phase that provided practical lessons for the larger study.

- Quality assurance procedures placed during fieldwork including training of personnel for clinical measurements and application of questionnaires as well as for management of laboratory samples.

- Study conducted within time framework and resources available. Monitoring processes and strategies for the assessment of progress and achievement of targets were established.

- Targets for enrolment of participants into the study were achieved with a recruitment of 98.9% of target sample size.

- Study accomplished with good response rates: overall response rate at enrolment was 73.2% and overall response rate at completion of the study was 61.6%.

- Study carried out within frameworks of ethical conduct of research involving human subjects.

11.3.3. Data analysis

- Good quality of data generated and doubly entered. Very low number of missing values and inconsistencies.

- Analysis strategy followed standard recommended statistical procedures for categorical and continuous outcomes, including the censoring techniques for those receiving medication.

- SMD used for visual comparison of continuous traits without units across risk factors.

- Prevalence estimates provided as simple proportions, and because of the study's age-matched design they did not differ to prevalence estimates derived from direct or indirect standardisation techniques.

11.4. Overall limitations of this study

11.4.1. Selection bias

Selection bias remains an important concern in migrant studies [58, 59, 62] and this study was not free from this limitation. This is basically a concern with population denominators whereby migrants studied as a proportion of those who remain in their place of origin or as a proportion of total migrants are not generally known.

Such concern dates back to 1938 [58]. Dorothy Thomas reviewed some studies evaluating the conflicting results of rural to urban migration studies which provided

> "apparently conflicting hypotheses as to the direction of this selection…: (1) cityward migrants are selected from the superior elements of the parent population; (2) cityward migrants are selected from the inferior elements; (3) cityward migrants are selected from the extremes, i.e., both the superior and the inferior elements; and (4) cityward migrants represent a random selection of the parent population" [58].

The latter scenario of random selection of migrants was expected for this study because of the specific Peruvian context, a few decades ago, of intense internal conflict in the area of study. However, being a process that occurred two decades ago and without population censuses of that period to provide denominators, it is a difficult task to confirm such assumptions. The study design was not able to completely eliminate migrant's selection bias but gave a less extreme type of selection amongst the migrants. A full chapter, Chapter V, was elaborated and discussed in detail on this matter. The ascertainment of migration status, key factors to identify an individual as a migrant, was confirmed by self-reported place of origin (rural versus urban), by self-reported ability of language proficiency (mother tongue) and an established migration pattern from Ayacucho to Lima. As concluded in Chapter V, these observations point towards a case where misclassification of exposure was minimal, but, socioeconomic-driving migration factors were not fully removed from the population study.

Being this a rural-to-urban migration study, an additional concern related to selection bias is the fact that because all migrants were not born in the same rural village selected as the rural group. Most of migration studies carry this weakness, particularly those that evaluate international migration. Basically, the unit of selection for place of origin is reduced to wider geographical spaces such as country of origin. In this study, the unit of selection for place of origin was department of origin. The rural group selected was truly a rural environment and it was selected because of that criteria. As inclusion criteria it was set out that both rural and migrant population should have been born in the department of Ayacucho, Peru. Overall, 73% of the country live in urban areas [88, 94]. While historically Lima has been mostly urban, Ayacucho has seen a progressive increase in its proportion of urban inhabitants: 25%, 33%, 37%, 48% and 58% in 1961, 1972, 1981, 1993 and 2007, respectively [88]. 70% of Ayacucho population live in poverty conditions [96]. Ideally, all migrants evaluated should have been born in that same village but such endeavour —without wider population censuses or historical records at the district or village levels— was not feasible. By design, this was not a before-after migration study as the classical Kenyan Luo migration study [78], but one more similar to the Yi Migrant Study where clearly identified groups were evaluated in a cross-sectional design [85]. Thus, the potential risk that some migrants may have been born in a non-rural part of Ayacucho, specially its capital, was a possibility that would have severely affected the selection of migrants included in the study. This possibility was refuted by the observation that nearly 80% in both rural and migrant group reported being born in a rural area (Table V-5, page 132).

11.4.2. Confounding

Section 2.7.3, on page 72, explored the role of potential confounding variables to be considered in this study including socioeconomic status, mental health, acculturation and high altitude. In addition to this, two major determinants of CVD risk, physical activity and diet, have not been included in this study. Physical activity has an impact on a number of specific traits analysed in this study, mainly on obesity-related risk factors. Diet also plays a role in the same factors but in addition, diet is also closely liked to lipid markers. These two factors are important ones for any study of CVD

risk. However, this study was set out to find out whether or not differences exist in a number of risk factors and is not capable, limited by its design, to explain but to postulate why these differences occur.

Insofar, much of epidemiology in the real world, can be contended, is about quantification of known risk factors —which are likely to combine together in ways that are not seen in other settings. A quicker proposition that favours this statement is that much of CVD epidemiology in the developed world arises from populations with important levels of smoking and hypertension. Yet, they were very low in the rural and migrant populations studied, but regardless of this, migrants had very high levels of various other risk factors. This study, by defining that complex patterns of differences in CVD risk exist in migrant and non-migrant populations, opens the venue for the need for further assessments to identify key determinants for such differences. One option, that is feasible using the data already generated by this study, is to select single specific outcomes and to evaluate the association using various other variables. For example, if blood pressure were to be selected as a single outcome, this study could well use behavioural risk factors such as alcohol and tobacco consumption as well as obesity-related markers in further multivariable analyses. Another option, in the medium-term future, is that this study generates the challenge to fully address the role of factors not measured. Physical activity and diet become important requirements of further evaluations, and sufficient financial and technical resources should be made available to ensure the feasibility and quality of such measurements.

11.4.3. Generalisation of findings

By its design, a cross-sectional study, this study cannot address causality and therefore caution has been placed in interpreting its findings. Due to the lack of baseline information on the traits studied, the claim that migration produces an increase in various CVD risk factors does not stand. Careful wording has been placed throughout in the interpretation of results, referring to them as associations.

Following the answering of all research questions posed in this study indicating that differences do exists on CVD risk profile following migration, it then becomes

relevant to ask what does this tell us about the impact of urbanisation on health more generally?

Currently, Lima hosts 50% of all internal migrants within Peru and, on the other hand, five Andean departments, including Ayacucho and similar ones, contribute to the largest "production" of out-migrants totalling 37% [88]. Therefore, it can be argued that migrants sampled in this study can be considered, to a large extent, representative of other rural-to-urban Andean migrants into Lima.

However, migrants sampled in this study are not necessarily representative of all rural-to-urban migrants in other LMIC or of international migrants. Although approximations can be made, it is important to remind that migrants of this study had a specific profile: born in a rural setting, relocated to periurban Lima, mean (SD) age at first migration was 14.7 (9 years-old), mean (SD) length of residence in an urban area was 32 (10.5 years) and mean (SD) lifetime exposure to urban environment was 67.7 (15.1%).

In relation to the whole population studied, it is also noticeable that they belong to the least rich sectors of Peruvian society. As presented before, the urban study site is a shantytown in periurban Lima. In 2003, Peru had a Gini coefficient of 52 (100 equals perfect inequality) [225], indicating a high level of economic inequality. Overall, 81% of all participants had a monthly income of less than USD $250, while Peru's gross domestic product per capita was USD $2,300 in 2004 [92, 93]. Thus, participants studied were largely rural and urban poor. As such this is one of the very first attempts to address a complete CVD risk profile in the poorest sectors of Peru. While rural poverty is well-acknowledged worldwide, migration and increasing urbanisation in today's LMIC goes along with increased levels urban poverty [54]. As recently reported, a vivid example of this is the societal change is currently being experienced by China with largest level of out-rural migration and increased urban poverty [47]. Also South Africa is facing a mortality transition with important levels of chronic disease-related mortality despite de high burden of HIV [342]. In these circumstances, findings of this study can be relevant to other similar LMIC settings.

11.5. Implications derived from the study

Together with PREVENCION study —conducted solely in a middle-class urban city in another Andean area [111, 112, 123, 317]— this is one of largest comprehensive CVD studies conducted in Peru to date. Unlike PREVENCION study, this study took advantage of rural and urban residents.

Findings from this study contribute to fill the massive knowledge gap on NCD and CVD in LMIC as currently advocated [43, 44]. As Yusuf et al. pointed out, there is urgent need to better document current rates —incidence and prevalence— of CVD mortality and morbidity in LMIC in order to properly assess burdens and future projections [44]. In the same vein, Unwin et al. argue for improved surveillance of all diseases in order to place non-communicable diseases properly within the context of the overall burden of disease [45].

Although this study answers a clear research question, it also opens various implications for research and policy. This study provides clear evidence that substantial differences exist on CVD risk profile in migrant and non-migrant populations in a LMIC setting. These findings can inform other similar LMIC but also, notoriously, challenges the adoption and incorporation of research findings from developed countries, particularly in CVD epidemiology [41], to other LMIC settings without prior knowledge of the risk profile of such populations. Therefore, potential areas of interest or implications, relevant to a larger and wider scientific and policy-making community are listed below.

11.5.1.1. Implications for further areas of research

- What is the role of physical activity and diet in explaining differential CVD risk in migrant and non-migrant groups?
- Are the difference observed in variation of CVD risk by time since migration or age at first migration replicable in other settings?

- Are the differences observed in CVD risks following migration also present in other CVD-related areas, e.g. lung function?

- Acknowledging potential limitations of physical activity questionnaires in developing countries, what is the quantifiable impact of migration on physical activity?

- Does the profile of differential CVD risk found in this study contribute to explain CVD-related and overall mortality patterns in general in LMIC?

- What is the risk of developing a CVD event in populations with low smoking and low blood pressure but with high lipid profile and metabolic markers levels?

- How can CVD risk prediction be improved in LMIC?

- Does the role of inflammation and insulin resistance in CVD disease differ in LMIC compared to developed societies?

11.5.1.2. Implications for health policy

- In the wider context of continuing migration and greater and faster urbanisation in LMIC, is there a need to study internal migrants as separate groups at high-risk?

- What is the role of the urban host environment on explaining differential CVD risk?

- Which features of the rural host environment explain healthier CVD profiles?

11.6. Further research

Potential areas for subsequent analyses with the data generated are also listed below in format of research questions. Although not specified in each of the questions, it is expected that the areas listed below include the evaluation of migrants and non-migrant groups.

11.6.1.1. Factors associated with differential CVD risk

- Why blood pressure levels and hypertension rates do not change in the direction and magnitude as other risk factors?
- What are the contributions of various risk factors on single major risk factors, e.g. alcohol, smoking and obesity on blood pressure?
- Which factor or factors, amongst all the ones studied, can better explain the higher levels of insulin resistance in the migrant and urban population studied?

11.6.1.2. Control and treatment

- What are the proportions of subjects with hypertension, diabetes or hypercholesterolemia currently on medication and with controlled levels?
- What are the risk factors associated with poor management and control of these conditions?

11.6.1.3. Performance of definitions

- What is the agreement of available metabolic syndrome definitions in classifying subjects with the condition in the population studied?
- What is the agreement of CVD risk scoring systems in identifying individuals at high-risk in the population studied?

11.7. Conclusion

Much of epidemiology in today's world is about quantification of known risk factors —which are likely to combine together in ways that are not seen in other settings. Findings of this study on the impact of rural-to-urban migration, at least within Peru, challenges common views and suspicions that following migration all risk factors amongst migrants will mirror the urban population. These findings suggest that the impact of migration on cardiovascular risk profile is not uniform across risk factors and this study add to the understanding of the complexity of migration and urbanisation.

References

1. Marmot M. Changing places changing risks: the study of migrants. *Public Health Rev* 1993;21(3-4):185-95.

2. McKay L, Macintyre S, Ellaway A. *Migration and health : a review of the international literature. Occasional paper No. 12*. Glasgow: Medical Research Council, Social & Public Health Sciences Unit, University of Glasgow; 2003.

3. Gaziano TA, Reddy KS, Paccaud F, Horton S, Chaturvedi V. Cardiovascular Disease. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al., editors. *Disease control priorities in developing countries*. 2nd ed. New York: Oxford University Press; 2006. p. 645-62.

4. Ollila E. Global health priorities - priorities of the wealthy? *Global Health* 2005;1(1):6.

5. Miranda JJ, Kinra S, Casas JP, Davey Smith G, Ebrahim S. Non-communicable diseases in low- and middle-income countries: context, determinants and health policy. *Trop Med Int Health* 2008;13(10):1225-34.

6. Miranda JJ, Patel V. Achieving the Millennium Development Goals: Does Mental Health Play a Role? *PLoS Medicine* 2005;2(10):e291.

7. Miranda JJ, Patel V. Mental Health in the Millennium Development Goals: Authors' Reply. *PLoS Medicine* 2007;4(1):e57.

8. World Health Organization. *Preventing chronic diseases: A vital investment. WHO Global Report*. Geneva: World Health Organization; 2005.

9. Adeyi O, Smith O, Robles S, World Bank. *Public policy and the challenge of chronic noncommunicable diseases*. Washington, D.C.: World Bank; 2007.

10. Horton R. The neglected epidemic of chronic disease. *Lancet* 2005;366(9496):1514.

11. Strong K, Mathers C, Leeder S, Beaglehole R. Preventing chronic diseases: how many lives can we save? *Lancet* 2005;366(9496):1578-82.

12. Epping-Jordan JE, Galea G, Tukuitonga C, Beaglehole R. Preventing chronic diseases: taking stepwise action. *Lancet* 2005;366(9497):1667-71.

13. Reddy SK, Shah B, Varghese C, Ramadoss A. Responding to the threat of chronic diseases in India. *Lancet* 2005;366(9498):1744-9.

14. Wang L, Kong L, Wu F, Bai Y, Burton R. Preventing chronic diseases in China. *Lancet* 2005;366(9499):1821-4.

15. Horton R. Chronic diseases: the case for urgent global action. *Lancet* 2007;370(9603):1881-2.

16. Abegunde DO, Mathers CD, Adam T, Ortegon M, Strong K. The burden and costs of chronic diseases in low-income and middle-income countries. *Lancet* 2007;370(9603):1929-38.

17. Gaziano TA, Galea G, Reddy KS. Scaling up interventions for chronic disease prevention: the evidence. *Lancet* 2007;370(9603):1939-46.

18. Lim SS, Gaziano TA, Gakidou E, Reddy KS, Farzadfar F, Lozano R, et al. Prevention of cardiovascular disease in high-risk individuals in low-income and middle-income countries: health effects and costs. *Lancet* 2007;370(9604):2054-62.

19. Banatvala N, Donaldson L. Chronic diseases in developing countries. *Lancet* 2007;370(9605):2076-8.

20. Adshead F, Stachenko S. International cooperation to combat chronic diseases. *Lancet* 2007;370(9605):2078-80.

21. Beaglehole R, Ebrahim S, Reddy S, Voute J, Leeder S. Prevention of chronic diseases: a call to action. *Lancet* 2007;370(9605):2152-7.

22. Anderson GF, Chu E. Expanding Priorities -- Confronting Chronic Disease in Countries with Low Income. *N Engl J Med* 2007;356(3):209-11.

23. Perel P, Casas JP, Ortiz Z, Miranda JJ. Noncommunicable Diseases and Injuries in Latin America and the Caribbean: Time for Action. *PLoS Medicine* 2006;3(9):e344.

24. Ebrahim S, Smeeth L. Non-communicable diseases in low and middle-income countries: a priority or a distraction? *Int J Epidemiol* 2005;34(5):961-6.

25. Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: Part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation* 2001;104(22):2746-53.

26. Suhrcke M, Nugent RA, Stuckler D, Rocco L. *Chronic disease: an economic perspective*. London: Oxford Health Alliance; 2006.

27. Leeder S, Raymond S, Greenberg H, Liu H, Esson K. *A Race Against Time: The Challenge of Cardiovascular Disease in Developing Economies*. New York: The Earth Institute, Columbia University; 2004.

28. Daar AS, Singer PA, Leah Persad D, Pramming SK, Matthews DR, Beaglehole R, et al. Grand challenges in chronic non-communicable diseases. *Nature* 2007;450(7169):494-6.

29. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006;367(9524):1747-57.

30. Baingana FK, Bos ER. Changing Patterns of Disease and Mortality in Sub-Saharan Africa: An Overview. In: Jamison DT, Feachem RG, Makgoba MW, Bos ER, Baingana FK, Hofman KJ, et al., editors. *Disease and Mortality in Sub-Saharan Africa*. 2nd ed. Washington, D.C.: World Bank; 2006. p. 1-9.

31. Lopez AD, Disease Control Priorities Project. *Global burden of disease and risk factors*. New York and Washington, D.C.: Oxford University Press and World Bank; 2006.

32. World Health Organization. *The World Health Report 2002 - Reducing risks, promoting healthy life*. Geneva: World Health Organization; 2002.

33. Murray CJL, Lopez AD, Harvard School of Public Health, World Health Organization, World Bank. *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020*. Cambridge, MA: Published by the Harvard School of Public Health on behalf of the World Health Organization and the World Bank, Distributed by Harvard University Press; 1996.

34. Omran AR. The epidemiologic transition. A theory of the epidemiology of population change. *Milbank Mem Fund Q* 1971;49(4):509-38.

35. Ounpuu S, Anand S, Yusuf S. The impending global epidemic of cardiovascular diseases. *Eur Heart J* 2000;21(11):880-3.

36. Reddy KS. Cardiovascular disease in non-Western countries. *N Engl J Med* 2004;350(24):2438-40.

37. Reddy KS. Cardiovascular diseases in the developing countries: dimensions, determinants, dynamics and directions for public health action. *Public Health Nutr* 2002;5(1A):231-7.

38. Yusuf S, Ounpuu S. Tackling the growing global burden of atherosclerotic cardiovascular diseases. *Eur J Cardiovasc Prev Rehabil* 2003;10(4):236-9.

39. Ezzati M, Lopez AD. Regional, disease specific patterns of smoking-attributable mortality in 2000. *Tob Control* 2004;13(4):388-95.

40. Mendis S, Yach D, Bengoa R, Narvaez D, Zhang X. Research gap in cardiovascular disease in developing countries. *Lancet* 2003;361(9376):2246-7.

41. Ebrahim S, Davey Smith G. Exporting failure? Coronary heart disease and stroke in developing countries. *Int J Epidemiol* 2001;30(2):201-5.

42. Victora CG, Schellenberg JA, Huicho L, Amaral J, El Arifeen S, Pariyo G, et al. Context matters: interpreting impact findings in child survival evaluations. *Health Policy Plan* 2005;20 Suppl 1:i18-i31.

43. Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al., editors. *Disease control priorities in developing countries*. 2nd ed. New York: Oxford University Press; 2006.

44. Yusuf S, Vaz M, Pais P. Tackling the challenge of cardiovascular disease burden in developing countries. *Am Heart J* 2004;148(1):1-4.

45. Unwin N, Setel P, Rashid S, Mugusi F, Mbanya JC, Kitange H, et al. Noncommunicable diseases in sub-Saharan Africa: where do they feature in the health research agenda? *Bull World Health Organ* 2001;79(10):947-53.

46. United Nations Centre for Human Settlements. *Cities in a globalizing world: Global report on human settlements 2001*. London; Sterling, VA: Earthscan Publications; 2001.

47. Hu X, Cook S, Salazar MA. Internal migration and health in China. *Lancet* 2008;372(9651):1717-9.

48. United Nations Population Fund. *The State of World Population 2007: Unleashing the Potential of Urban Growth*. New York: UNFPA; 2007.

49. Hinrichsen D, Salem R, Blackburn R. *Meeting the Urban Challenge*. Series M, No.16. Baltimore: The Johns Hopkins Bloomberg School of Public Health, Population Information Program; 2002.

50. Leon DA. Cities, urbanization and health. *Int J Epidemiol* 2008;37(1):4-8.

51. Whiting D, Unwin N. Cities, urbanization and health. *Int J Epidemiol* 2008;doi:10.1093/ije/dyn152

52. Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: Part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. *Circulation* 2001;104(23):2855-64.

53. Galea S, Freudenberg N, Vlahov D. Cities and population health. *Soc Sci Med* 2005;60(5):1017-33.

54. Fay M. *The urban poor in Latin America*. Washington, D.C.: World Bank; 2005.

55. Last M. *A Dictionary of Epidemiology*. 4th ed. New York: Oxford University Press; 2001.

56. Bogin B. *Patterns of Human Growth*. 2nd ed. Cambridge: Cambridge University Press; 1999.

57. Boas F. Changes in the bodily form of descendents of immigrants. *Am Anthropol* 1912;14:530-63.

58. Thomas DS. Review: Selective Migration. *Milbank Mem Fund Q* 1938;16(4):403-7.

59. Hill AB. *Internal Migration and its Effects upon the Death-Rates: With Special Reference to the County of Essex*. London: His Majesty's Stationery Office, Medical Research Council Special Report Series, No. 95; 1925.

60. Doll R. Austin Bradford Hill. 8 July 1897-18 April 1991. *Biogr Mem Fellows R Soc* 1994;40:129-40.

61. Doll R. Sir Austin Bradford Hill: A Personal View of His Contribution to Epidemiology. *J R Stat Soc Ser A Stat Soc* 1995;158(1):155-63.

62. Razum O. Commentary: Of salmon and time travellers—musing on the mystery of migrant mortality. *Int J Epidemiol* 2006;35(4):919-21.

63. Dufour DL, Piperata BA. Rural-to-urban migration in Latin America: an update and thoughts on the model. *Am J Hum Biol* 2004;16(4):395-404.

64. Jacoby E. The obesity epidemic in the Americas: making healthy choices the easiest choices. *Rev Panam Salud Publica* 2004;15(4):278-84.

65. Jacoby E, Bull F, Neiman A. Rapid changes in lifestyle make increased physical activity a priority for the Americas. *Rev Panam Salud Publica* 2003;14(4):223-5, 6-8.

66. Fraser B. Latin America's urbanisation is boosting obesity. *Lancet* 2005;365(9476):1995-6.

67. Marmot MG, Syme SL, Kagan A, Kato H, Cohen JB, Belsky J. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: prevalence of coronary and hypertensive heart disease and associated risk factors. *Am J Epidemiol* 1975;102(6):514-25.

68. Kelleher CC, Lynch JW, Daly L, Harper S, Fitz-simon N, Bimpeh Y, et al. The "Americanisation" of migrants: Evidence for the contribution of ethnicity, social deprivation, lifestyle and life-course processes to the mid-20th century Coronary Heart Disease epidemic in the US. *Social Science & Medicine* 2006;63(2):465-84.

69. Hawkins SS, Lamb K, Cole TJ, Law C, the Millennium Cohort Study Child Health G. Influence of moving to the UK on maternal health behaviours: prospective cohort study. *BMJ* 2008;336(7652):1052-5.

70. Patel JV, Vyas A, Cruickshank JK, Prabhakaran D, Hughes E, Reddy KS, et al. Impact of migration on coronary heart disease risk factors: comparison of Gujaratis in Britain and their contemporaries in villages of origin in India. *Atherosclerosis* 2006;185(2):297-306.

71. Wild SH, Fischbacher C, Brock A, Griffiths C, Bhopal R. Mortality from all causes and circulatory disease by country of birth in England and Wales 2001-2003. *J Public Health (Oxf)* 2007;29(2):191-8.

72. Abraido-Lanza AF, Dohrenwend BP, Ng-Mak DS, Turner JB. The Latino mortality paradox: a test of the "salmon bias" and healthy migrant hypotheses. *Am J Public Health* 1999;89(10):1543-8.

73. Lizarzaburu JL, Palinkas LA. Immigration, acculturation, and risk factors for obesity and cardiovascular disease: a comparison between Latinos of Peruvian descent in Peru and in the United States. *Ethn Dis* 2002;12(3):342-52.

74. Sharma RD, Michalowski M, Verma RB. Mortality differentials among immigrant populations in Canada. *Int Migr* 1990;28(4):443-50.

75. Razum O, Twardella D. Time travel with Oliver Twist--towards an explanation foa a paradoxically low mortality among recent immigrants. *Trop Med Int Health* 2002;7(1):4-10.

76. Marmot MG, Syme SL. Acculturation and coronary heart disease in Japanese-Americans. *Am J Epidemiol* 1976;104(3):225-47.

77. Serrano-Rios M, Goday A, Martinez Larrad T. Migrant populations and the incidence of type 1 diabetes mellitus: an overview of the literature with a focus on the Spanish-heritage countries in Latin America. *Diabetes Metab Res Rev* 1999;15(2):113-32.

78. Poulter NR, Khaw KT, Hopwood BE, Mugambi M, Peart WS, Rose G, et al. The Kenyan Luo migration study: observations on the initiation of a rise in blood pressure. *BMJ* 1990;300(6730):967-72.

79. Poulter NR, Khaw KT, Sever PS. Higher blood pressures of urban migrants from an African low-blood pressure population are not due to selective migration. *Am J Hypertens* 1988;1(3 Pt 3):143S-5S.

80. Nadim A, Amini H, Malek-Afzali H. Blood Pressure and Rural-Urban Migration in Iran. *Int J Epidemiol* 1978;7(2):131-8.

81. Diamond J. The double puzzle of diabetes. *Nature* 2003;423(6940):599-602.

82. Unwin N, McLarty D, Machibya H, Aspray T, Tamin B, Carlin L, et al. Changes in blood pressure and lipids associated with rural to urban migration in Tanzania. *J Hum Hypertens* 2006;20(9):704-6.

83. Sobngwi E, Mbanya JC, Unwin NC, Porcher R, Kengne AP, Fezeu L, et al. Exposure over the life course to an urban environment and its relation with obesity, diabetes, and hypertension in rural and urban Cameroon. *Int J Epidemiol* 2004;33(4):769-76.

84. He J, Klag MJ, Wu Z, Qian MC, Chen JY, Mo PS, et al. Effect of migration and related environmental changes on serum lipid levels in southwestern Chinese men. *Am J Epidemiol* 1996;144(9):839-48.

85. He J, Klag MJ, Whelton PK, Chen JY, Mo JP, Qian MC, et al. Migration, blood pressure pattern, and hypertension: the Yi Migrant Study. *Am J Epidemiol* 1991;134(10):1085-101.

86. He J, Tell GS, Tang YC, Mo PS, He GQ. Effect of migration on blood pressure: the Yi People Study. *Epidemiology* 1991;2(2):88-97.

87. Torun B, Stein AD, Schroeder D, Grajeda R, Conlisk A, Rodriguez M, et al. Rural-to-urban migration and cardiovascular disease risk factors in young Guatemalan adults. *Int J Epidemiol* 2002;31(1):218-26.

88. Instituto Nacional de Estadística e Informática. *Perfil Sociodemográfico del Perú 2007 [Peru's 2007 Sociodemographic Profile]*. Lima, Perú: INEI. Available at: http://www1.inei.gob.pe/Anexos/libro.pdf; 2008.

89. Instituto Nacional de Estadística e Informática. *Censo de Población y Vivienda 2005 [2005 Peru's Population and Household's Census]*. Lima, Perú: INEI; 2006.

90. Instituto Nacional de Estadística e Informática. Perú: Estimaciones y proyecciones de población 1950–2050 [Peru: population estimations and projections 1950-2050]. *Boletín de Análisis Demográfico* 2001;35.

91. Instituto Cuanto. *Perú en números 2007: anuario estadístico [Peru in numbers 2007: annual statistics]*. Lima: Cuanto S.A.; 2007.

92. Pan American Health Organization. Peru. *Health in the Americas - 2007 Edition*. Washington, D.C.: PAHO; 2007. p. 576-95.

93. EarthTrends. *Economic Indicators -- Peru*. Washington, DC: World Resources Institute. Available at: http://earthtrends.wri.org/pdf_library/country_profiles/eco_cou_604.pdf; 2003.

94. World Health Organization. *WHO Statistical Information System, Core Health Indicators*. Geneva: World Health Organization. Available at: http://www.who.int/whosis/database/core/core_select.cfm.

95. Instituto Nacional de Estadística e Informática. *La pobreza en el Perú 2003–2004 [Poverty in Peru 2003–2004]*. Lima, Perú: INEI; 2005.

96. Instituto Nacional de Estadística e Informática. *Condiciones de vida en los departamentos del Perú: 2003–2004 [Living conditions of departments of Peru: 2003–2004]*. Lima, Perú: INEI; 2005.

97. Ministerio de Trabajo y Promoción de Empleo. *Informe Estadístico Mensual [Monthly Statistical Report], Año 9, N° 106 (March)*. Lima: Ministerio de Trabajo y Promoción de Empleo del Perú; 2005.

98. Instituto Nacional de Estadística e Informática. *Encuesta Nacional de Hogares 2007 [2007 Peru's National Household's Survey]*. Lima, Perú: INEI; 2007.

99. Instituto Cuanto. *Perú en números 2002: anuario estadístico [Peru in numbers 2002: annual statistics]*. Lima: Cuanto S.A.; 2002.

100. Instituto Nacional de Estadística e Informática and Macro International. *Encuesta Demográfica y de Salud Familiar 2000 [Demographic and Health Survey 2000]*. Lima: INEI and Macro International. Available at: http://www.measuredhs.com/pubs/pub_details.cfm?ID=334; 2001.

101. Subaiya L. *Internal Migration and the Use of Reproductive and Child Health Services in Peru*. Calverton, MD: Macro International. Available at: http://www.measuredhs.com/pubs/pdf/WP38/WP38.pdf; 2007.

102. Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bull World Health Organ* 2005;83(3):171-7.

103. Pan American Health Organization. Annex Tables. *Health statistics from the Americas*. 2006 ed. Washington, D.C.: Pan American Health Organization, Pan American Sanitary Bureau, Regional Office of the World Health Organization. Available at: http://www.paho.org/English/DD/AIS/HSA2006_AT.pdf (Accessed July 9, 2007); 2006.

104. Pan American Health Organization. Chapter I. Average deaths and estimated rates per 100,000 population from selected groups of causes, by age group, sex, and country, around 1980 and last three years available. *Health Statistics from the Americas, 2003 Edition*. Washington, D.C.: Pan American Health Organization; 2003. p. 1-154.

105. International Diabetes Federation. *Diabetes atlas*. 2nd ed. Brussels: International Diabetes Federation; 2003.

106. Seclen-Palacin JA, Jacoby ER. [Sociodemographic and environmental factors associated with sports physical activity in the urban population of Peru]. *Rev Panam Salud Publica* 2003;14(4):255-64.

107. Mispireta ML, Rosas ÁM, Velásquez JE, Lescano AG, Lanata CF. Transición nutricional en el Perú, 1991 - 2005. *Rev Perú Med Exp Salud Publica* 2007;24(2):129-35.

108. Cárdenas De Jurado HG, Monterrey Gutiérrez PA, Roldán Arbieto L, Mendoza Tasayco F. *Encuesta Nacional de Indicadores Nutricionales, Bioquímicos, Socioeconómicos y Culturales relacionados con las Enfermedades Crónicas Degenerativas [National Survey of Nutritional, Bioquemical, Socioeconomic and Cultural indicators related to Chronic Degenerative Diseases]*. Lima: Centro Nacional de Alimentación y Nutricion, Instituto Nacional de Salud; 2006.

109. Seclén S, Leey J, Villena A, Herrera B, Menacho J, Carrasco A, et al. Prevalencia de obesidad, diabetes mellitus, hipertensión, arterial e hipercolesterolemia como factores de riesgo coronario y cerebrovascular en población adulta de la costa, sierra y selva del Peru [Prevalence of obesity, diabetes mellitus, hypertension and hypercholesterolemia as risk factors for coronary disease and stroke in adult population from Peruvian coast, sierra and jungle]. *Acta Med Per* 1999;XVII(1):8-12.

110. Schargrodsky H, Hernandez-Hernandez R, Champagne BM, Silva H, Vinueza R, Silva Aycaguer LC, et al. CARMELA: assessment of cardiovascular risk in seven Latin American cities. *Am J Med* 2008;121(1):58-65.

111. Medina-Lezama J, Zea-Diaz H, Morey-Vargas O, Bolanos-Salazar J, Postigo-Macdowall M, Paredes-Diaz S, et al. Prevalence and patterns of hypertension in Peruvian Andean Hispanics: The PREVENCION study. *J Am Soc Hypertens* 2007;1(3):216-25.

112. Medina-Lezama J, Zea-Diaz H, Morey-Vargas OL, Bolanos-Salazar JF, Munoz-Atahualpa E, Postigo-Macdowall M, et al. Prevalence of the metabolic syndrome in Peruvian Andean hispanics: The PREVENCION study. *Diabetes Res Clin Pract* 2007;78(2):270-81.

113. Jacoby E, Goldstein J, Lopez A, Nunez E, Lopez T. Social class, family, and life-style factors associated with overweight and obesity among adults in Peruvian cities. *Prev Med* 2003;37(5):396-405.

114. Goldstein J, Jacoby E, Del Aguila R, Lopez A. Poverty is a predictor of non-communicable disease among adults in Peruvian cities. *Prev Med* 2005;41(3-4):800-6):800-6.

115. Davies AR, Miranda JJ, Gilman RH, Smeeth L. Hypertension among adults in a deprived urban area of Peru - undiagnosed and uncontrolled? *BMC Research Notes* 2008;1:2.

116. Lindgarde F, Benavente Ercilla M, Retamozo Correa L, Ahren B. Body adiposity, insulin, and leptin in subgroups of Peruvian Amerindians. *High Alt Med Biol* 2004;5(1):27-31.

117. Baracco R, Mohanna S, Seclen S. A comparison of the prevalence of metabolic syndrome and its components in high and low altitude populations in peru. *Metab Syndr Relat Disord* 2007;5(1):55-62.

118. Seclen S, Villena A, Larrad MT, Gamarra D, Herrera B, Perez CF, et al. Prevalence of the metabolic syndrome in the mestizo population of peru. *Metab Syndr Relat Disord* 2006;4(1):1-6.

119. Mohanna S, Baracco R, Seclen S. Lipid profile, waist circumference, and body mass index in a high altitude population. *High Alt Med Biol* 2006;7(3):245-55.

120. Lindgarde F, Soderberg S, Olsson T, Ercilla MB, Correa LR, Ahren B. Overweight is associated with lower serum leptin in Peruvian Indian than in Caucasian women: A dissociation contributing to low blood pressure? *Metabolism* 2001;50(3):325-9.

121. Llanos-Zavalaga F, Nájar Trujillo NE, Mayca Pérez J, Rosas A. Prevalencia de obesidad e hipercolesterolemia en la Facultad de Medicina de la Universidad Peruana Cayetano Heredia - 1998. *Rev Med Hered* 2001;12(3):78-84.

122. Rosas A, Lama G, Llanos-Zavalaga F, Dunstan J. Prevalencia de obesidad e hipercolesterolemia en trabajadores de una institución estatal de Lima - Perú. *Rev Perú Med Exp Salud Publica* 2002;19(2):87-92.

123. Medina-Lezama J, Chirinos JA, Zea Diaz H, Morey O, Bolanos JF, Munoz-Atahualpa E, et al. Design of PREVENCION: a population-based study of cardiovascular disease in Peru. *Int J Cardiol* 2005;105(2):198-202.

124. Golte J. *Cultura, racionalidad y migración andina*. Lima: Instituto de Estudios Peruanos; 2001.

125. Comisión de la Verdad y Reconciliación. *Informe Final de la Comisión de la Verdad y Reconciliación [Peruvian Truth and Reconciliation Comision's Final Report]*. Lima: Comisión de la Verdad y Reconciliación. Available at: http://www.cverdad.org.pe/ifinal/index.php; 2003.

126. Stern SJ. *Shining and other paths: war and society in Peru, 1980-1995*. Durham, NC: Duke Univeristy Press; 1998.

127. Coral I. *Desplazamiento por violencia politica en el Peru, 1980-1992*. Lima: Instituto de Estudios Peruanos; 1994.

128. Rojas RM. El Peru despues de 15 anos de violencia (1980-1995). *Estudos Avançados* 1997;11:287-308.

129. Pedersen D, Tremblay J, Errázuriz C, Gamarra J. The sequelae of political violence: Assessing trauma, suffering and dislocation in the Peruvian highlands. *Soc Sci Med* 2008;67(2):205-17.

130. Instituto Nacional de Estadística e Informática. *Migraciones internas en el Perú [Internal migration in Peru]*. Lima, Perú: INEI. Available at: http://www1.inei.gob.pe/biblioineipub/bancopub/Est/Lib0018/n00.htm; 1995.

131. Rothhammer F, Llop E, Carvallo P, Moraga M. Origin and evolutionary relationships of native Andean populations. *High Alt Med Biol* 2001;2(2):227-33.

132. Shinoda K, Adachi N, Guillen S, Shimada I. Mitochondrial DNA analysis of ancient Peruvian highlanders. *Am J Phys Anthropol* 2006;131(1):98-107.

133. Brutsaert TD, Parra E, Shriver M, Gamboa A, Palacios JA, Rivera M, et al. Effects of birthplace and individual genetic admixture on lung volume and exercise phenotypes of Peruvian Quechua. *Am J Phys Anthropol* 2004;123(4):390-8.

134. Modiano G, Bernini L, Carter ND, Santachiara B, Detter JC, Baur EW, et al. Survey of several red cell and serum genetic markers in a Peruvian population. *Am J Hum Genet* 1972;24(2):111-23.

135. Rodriguez-Delfin LA, Rubin-de-Celis VE, Zago MA. Genetic diversity in an Andean population from Peru and regional migration patterns of Amerindians in South America: data from Y chromosome and mitochondrial DNA. *Hum Hered* 2001;51(1-2):97-106.

136. Brutsaert TD, Parra EJ, Shriver MD, Gamboa A, Palacios JA, Rivera M, et al. Spanish genetic admixture is associated with larger V(O2) max decrement from sea level to 4338 m in Peruvian Quechua. *J Appl Physiol* 2003;95(2):519-28.

137. de Pablo R, Beraun Y, Nieto A, Calzada JE, Rementeria MC, Sanz L, et al. HLA class I and class II allele distribution in the Peruvian population. *Tissue Antigens* 2000;56(6):507-14.

138. Sapsford R. *Survey Research*. London: SAGE Publications Ltd; 1999.

139. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health* 2006;60(1):7-12.

140. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 2). *J Epidemiol Community Health* 2006;60(2):95-101.

141. Galobardes B, Lynch J, Davey Smith G. Measuring socioeconomic position in health research. *Br Med Bull* 2007;81-82(1):21-37.

142. Howe LD, Hargreaves JR, Huttly SR. Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries. *Emerg Themes Epidemiol* 2008;5:3.

143. Expert Group on Poverty Statistics (Rio Group). *Compendium of best practices in poverty measurement*. Rio de Janeiro: United Nations Economic Commission for Latin America and the Caribbean (ECLAC). Available at: http://www.eclac.org/publicaciones/xml/3/26593/rio_group_compendium.pdf; 2006.

144. Davey Smith G, Gordon D, Kelly M, Nandy S, Subramanian SV. *Inequalities in health in India: the methodological construction of indices and measures. Report for the UK Department for International Development*. Bristol: Townsend Centre for International Poverty Research, University of Bristol. Available at: http://www.bristol.ac.uk/poverty/health%20inequalities_files/Methodology%20report.doc; 2003.

145. Gordon D, Pantazis C. *Breadline Britain in the 1990s*. Aldershot: Ashgate; 1997.

146. Townsend P. *Poverty in the United Kingdom: A survey of household resources and standards of living*. Harmondsworth: Penguin; 1979.

147. Gordon D. Census based deprivation indices: their weighting and validation. *J Epidemiol Community Health* 1995;49 Suppl 2:S39-44.

148. Marsh A, Gordon D, Pantazis C, Heslop P. *Home sweet home? The impact of poor housing on health*. Bristol: Policy Press; 1999.

149. Shaw M, Dorling D, Gordon D, Davey Smith G. *The widening gap: Health inequalities and policy in Britain*. Bristol: Policy Press; 1999.

150. Townsend P, Gordon D. *World poverty: New policies to defeat an old enemy*. Bristol: Policy Press; 2002.

151. Gordon D, Nandy S, Pantazis C, Pemberton S, Townsend P. *Child poverty in the developing world*. Bristol: Policy Press; 2003.

152. Spicker P, Álvarez Leguizamón S, Gordon D. *Poverty: An international glossary*. 2nd ed. London: Zed Books; 2007.

153. Nandy S. "Misunderestimating" Chronic Poverty?: Exploring Chronic Poverty in Developing Countries Using Cross-Sectional Demographic and Health Data. *Global Social Policy* 2008;8(1):45-79.

154. O'Brien E, Waeber B, Parati G, Staessen J, Myers MG. Blood pressure measuring devices: recommendations of the European Society of Hypertension. *BMJ* 2001;322(7285):531-6.

155. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA* 2003;289:2560-72.

156. Mancia G, De Backer G, Dominiczak A, Cifkova R, Fagard R, Germano G, et al. 2007 Guidelines for the management of arterial hypertension: The Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *Eur Heart J* 2007;28(12):1462-536.

157. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the Concentration of Low-Density Lipoprotein Cholesterol in Plasma, Without Use of the Preparative Ultracentrifuge. *Clin Chem* 1972;18(6):499-502.

158. Warnick GR, Knopp RH, Fitzpatrick V, Branson L. Estimating low-density lipoprotein cholesterol by the Friedewald equation is adequate for classifying patients on the basis of nationally recommended cutpoints. *Clin Chem* 1990;36(1):15-9.

159. World Health Organization. *Definition, diagnosis and classification of diabetes mellitus and its complications*. Geneva: World Health Organization; 1999.

160. Wallace TM, Levy JC, Matthews DR. Use and Abuse of HOMA Modeling. *Diabetes Care* 2004;27(6):1487-95.

161. Lawlor DA, Fraser A, Ebrahim S, Smith GD. Independent associations of fasting insulin, glucose, and glycated haemoglobin with stroke and coronary heart disease in older women. *PLoS Med* 2007;4(8):e263.

162. Alberti KGMM, Zimmet P, Shaw J. The metabolic syndrome—a new worldwide definition. *Lancet* 2005;366(9491):1059-62.

163. Alberti KGMM, Zimmet P, Shaw J. Metabolic syndrome—a new world-wide definition. A Consensus Statement from the International Diabetes Federation. *Diabetic Medicine* 2006;23(5):469-80.

164. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, et al. Diagnosis and Management of the Metabolic Syndrome: An American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement: Executive Summary. *Circulation* 2005;112(17):e285-90.

165. Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults. Executive Summary of the 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). *JAMA* 2001;285(19):2486-97.

166. World Health Organization. *WHO/ISH risk prediction charts: Pocket guidelines for assessment and management of cardiovascular risk*. Geneva: World Health Organization. Available at: http://www.who.int/cardiovascular_diseases/guidelines/Pocket_GL_information/en/index.html; 2007.

167. Mendis S, Lindholm LH, Mancia G, Whitworth J, Alderman M, Lim S, et al. World Health Organization (WHO) and International Society of Hypertension (ISH) risk prediction charts: assessment of cardiovascular risk for prevention and control of cardiovascular disease in low and middle-income countries. *J Hypertens* 2007;25(8):1578-82.

168. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97(18):1837-47.

169. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;370(9596):1453-7.

170. Newson R. Confidence intervals and p-values for delivery to the end user. *The Stata Journal* 2003;3(3):245-69.

171. Newson R. Stata tip 1: The eform() option of `regress`. *The Stata Journal* 2003;3(4):445.

172. Statistical Consulting Group. *FAQ: How do I interpret a regression model when some variables are log transformed?* Los Angeles, CA: UCLA, Academic Technology Services. Available at: http://www.ats.ucla.edu/stat/mult_pkg/faq/general/log_transformed_regression.htm.

173. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: antihypertensive therapy and systolic blood pressure. *Stat Med* 2005;24(19):2911-35.

174. Statistical Consulting Group. *Annotated Stata Output: Multiple Regression Analysis*. Los Angeles, CA: UCLA, Academic Technology Services. Available at: http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter1/statareg_annotated2.htm.

175. Schünemann HJ, Oxman AD, Vist GE, Higgins JP, Deeks JJ, Glasziou P, et al. Interpreting results and drawing conclusions. In: Higgins J, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions, Version 500 [updated February 2008]*: The Cochrane Collaboration. Available at www.cochrane-handbook.org; 2008.

176. Cohen J. *Statistical Power Analysis in the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1988.

177. Nuffield Council on Bioethics. *The ethics of clinical research in developing countries*. London: Nuffield Council on Bioethics; 1999.

178. Nuffield Council on Bioethics. *The ethics of research related to healthcare in developing countries*. London: Nuffield Council on Bioethics; 2002.

179. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health* 2004;94(3):400-5.

180. Department of Epidemiology and Population Health. *Extended Epidemiology. Course Manual*. London: London School of Hygiene and Tropical Medicine; 2004.

181. Acheson D. *Independent inquiry into inequalities in health report*. London: Stationery Office; 1997.

182. Macleod J, Davey Smith G, Metcalfe C, Hart C. Is subjective social status a more important determinant of health than objective social status? Evidence from a prospective observational study of Scottish men. *Soc Sci Med* 2005;65(9):1916-29.

183. Davey Smith G, Neaton JD, Wentworth D, Stamler R, Stamler J. Socioeconomic differentials in mortality risk among men screened for the Multiple Risk Factor Intervention Trial: I. White men. *Am J Public Health* 1996;86(4):486-96.

184. Marmot MG, Rose G, Shipley M, Hamilton PJ. Employment grade and coronary heart disease in British civil servants. *J Epidemiol Community Health* 1978;32(4):244-9.

185. Bovet P, Ross AG, Gervasoni JP, Mkamba M, Mtasiwa DM, Lengeler C, et al. Distribution of blood pressure, body mass index and smoking habits in the urban population of Dar es Salaam, Tanzania, and associations with socioeconomic status. *Int J Epidemiol* 2002;31(1):240-7.

186. Monteiro CA, Conde WL, Lu B, Popkin BM. Obesity and inequities in health in the developing world. *Int J Obes Relat Metab Disord* 2004;28(9):1181-6.

187. Monteiro CA, Moura EC, Conde WL, Popkin BM. Socioeconomic status and obesity in adult populations of developing countries: a review. *Bull World Health Organ* 2004;82(12):940-6.

188. Davey Smith G, Shaw M, Dorling D. Shrinking areas and mortality. *Lancet* 1998;352(9138):1439-40.

189. Harpham T. Urbanization and mental health in developing countries: a research role for social scientists, public health professionals and social psychiatrists. *Soc Sci Med* 1994;39(2):233-45.

190. Rosengren A, Hawken S, Ounpuu S, Sliwa K, Zubaid M, Almahmeed WA, et al. Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* 2004;364(9438):953-62.

191. Macleod J, Davey Smith G. Psychosocial factors and public health: a suitable case for treatment? *J Epidemiol Community Health* 2003;57(8):565-70.

192. Lara M, Gamboa C, Kahramanian MI, Morales LS, Hayes Bautista DE. Acculturation and Latino health in the United States: a review of the literature and its sociopolitical context. *Annu Rev Public Health* 2005;26(1):367-97.

193. Abraido-Lanza AF, Armbrister AN, Florez KR, Aguirre AN. Toward a Theory-Driven Model of Acculturation in Public Health Research. *Am J Public Health* 2006;96(8):1342-6.

194. Deyo RA, Diehl AK, Hazuda H, Stern MP. A simple language-based acculturation scale for Mexican Americans: validation and application to health care research. *Am J Public Health* 1985;75(1):51-5.

195. Himmelgreen DA, Perez-Escamilla R, Martinez D, Bretnall A, Eells B, Peng Y, et al. The longer you stay, the bigger you get: length of time and language use in the U.S. are associated with obesity in Puerto Rican women. *Am J Phys Anthropol* 2004;125(1):90-6.

196. Moran A, Roux AV, Jackson SA, Kramer H, Manolio TA, Shrager S, et al. Acculturation is associated with hypertension in a multiethnic sample. *Am J Hypertens* 2007;20(4):354-63.

197. Koya DL, Egede LE. Association between length of residence and cardiovascular disease risk factors among an ethnically diverse group of United States immigrants. *J Gen Intern Med* 2007;22(6):841-6.

198. Viruell-Fuentes EA. Beyond acculturation: immigration, discrimination, and health research among Mexicans in the United States. *Soc Sci Med* 2007;65(7):1524-35.

199. León-Velarde F, Arregui A. *Desadaptación a la vida en las grandes alturas [Desadaptation to living in high altitude]*. Lima: l'Institut Français d'Etudes Andines and Universidad Peruana Cayetano Heredia; 1994.

200. Brito J, Siques P, Leon-Velarde F, De La Cruz JJ, Lopez V, Herruzo R. Chronic intermittent hypoxia at high altitude exposure for over 12 years: assessment of hematological, cardiovascular, and renal effects. *High Alt Med Biol* 2007;8(3):236-44.

201. World Health Organization. *WHO STEPwise approach to Surveillance (STEPS). STEPS Manual*. Geneva: World Health Organization. Available at: http://www.who.int/chp/steps/manual/en/index.html.

202. Adler NE. *The MacArthur Scale of Subjective Social Status*. San Francisco, CA: John D. and Catherine T. MacArthur Research Network on Socioeconomic Status and Health. Available at: http://www.macses.ucsf.edu/Research/Psychosocial/notebook/subjective.html; 2000.

203. Adler NE, Epel ES, Castellazzo G, Ickovics JR. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. *Health Psychol* 2000;19(6):586-92.

204. Hu P, Adler NE, Goldman N, Weinstein M, Seeman TE. Relationship between subjective social status and measures of health in older Taiwanese persons. *J Am Geriatr Soc* 2005;53(3):483-8.

205. Subramanian SV, Nandy S, Kelly M, Gordon D, Davey Smith G. Patterns and distribution of tobacco consumption in India: cross sectional multilevel evidence from the 1998-9 national family health survey. *BMJ* 2004;328(7443):801-6.

206. Centers for Disease Control and Prevention (CDC). *Behavioral Risk Factor Surveillance System Survey Questionnaire*. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2004.

207. Maconochie N, Doyle P, Prior S. The National Women's Health Study: assembly and description of a population-based reproductive cohort. *BMC Public Health* 2004;4(1):35.

208. Goldberg D, Williams P. *A user's guide to the General Health Questionnaire*. Windsor, Berkshire: NFER-Nelson; 1988.

209. de Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull* 2004;25(1 Suppl):S27-36.

210. Habicht JP. [Standardization of quantitative epidemiological methods in the field]. *Bol Oficina Sanit Panam* 1974;76(5):375-84.

211. Roche. *Serum Work Area: cobas® modular platform*: Available at: http://labsystems.roche.com/content/cobas_modular_platform.html.

212. Bio-Rad. *Bio-Rad Diabetes Monitoring Products*: Available at: http://diabetes.bio-rad.com/html/products.html.

213. Bowling A. *Research methods in health: Investigating health and health services*. 2nd ed. Maidenhead: Open University Press; 2002.

214. Walker M, Whincup PH, Shaper AG. The British Regional Heart Study 1975-2004. *Int J Epidemiol* 2004;33(6):1185-92.

215. Shaper AG, Pocock SJ, Walker M, Cohen NM, Wale CJ, Thomson AG. British Regional Heart Study: cardiovascular risk factors in middle-aged men in 24 towns. *Br Med J (Clin Res Ed)* 1981;283(6285):179-86.

216. Brindle P, Emberson J, Lampe F, Walker M, Whincup P, Fahey T, et al. Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. *BMJ* 2003;327(7426):1267.

217. Lawlor DA, Whincup P, Emberson JR, Rees K, Walker M, Ebrahim S. The challenge of secondary prevention for coronary heart disease in older patients: findings from the British Women's Heart and Health Study and the British Regional Heart Study. *Fam Pract* 2004;21(5):582-6.

218. Lawlor DA, Emberson JR, Ebrahim S, Whincup PH, Wannamethee SG, Walker M, et al. Is the Association Between Parity and Coronary Heart Disease Due to Biological Effects of Pregnancy or Adverse Lifestyle Risk Factors Associated With Child-Rearing?: Findings From the British Women's Heart and Health Study and the British Regional Heart Study. *Circulation* 2003;107(9):1260-4.

219. Jackson R, Chambless LE, Yang K, Byrne T, Watson R, Folsom A, et al. Differences between respondents and nonrespondents in a multicenter community-based study vary by gender ethnicity. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. *J Clin Epidemiol* 1996;49(12):1441-46.

220. Elliot D. *Weighting for non-response: A survey researcher's guide*. London: Office of Population Censuses and Surveys; 1991.

221. Gelman A, Carlin JB. Poststratification and weighting adjustments. In: Groves R, Eltinge J, Little RJA, editors. *Survey Nonresponse*. New York: Wiley; 2001.

222. Mishra V, Hong R, Khan S, Gu Y, Liu L. *Evaluating HIV Estimates from National Population-Based Surveys for Bias Resulting from Non-Response. DHS Analytical Studies No. 12*. Calverton, MA: Macro International Inc.; 2008.

223. Altamirano T, Copestake J, Figueroa A, Wright K. *Poverty Studies in Peru: Towards a More Inclusive Study of Exclusion*. Bath: University of Bath. WeD Working Paper 05; 2003.

224. United Nations Development Programme. *Human Development Report, Peru 2002: Maximizing opportunities*. Lima: UNDP; 2002.

225. World Bank. *World development Indicators 2007*. Washington, DC: World Bank; 2007.

226. Booysen F. An Overview and Evaluation of Composite Indices of Development. *Social Indicators Research* 2002;59(2):115-51.

227. UNICEF. *The State of the World's Children 2005: Childhood under threat*. New York, NY: UNICEF; 2005.

228. Nunnally JC. *Psychometric Theory*. New Delhi: Tate McGraw-Hill Publishing Company Ltd; 1981.

229. World Health Organization. *The World Health Report 2002: Reducing risks, promoting healthy life*. Geneva: World Health Organization; 2002.

230. Peto R, Lopez AD, Boreham J, Thun M, Heath C, Jr. Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *Lancet* 1992;339(8804):1268-78.

231. Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet* 2003;362(9387):847-52.

232. Organización Panamericana de la Salud. *Alcohol, género, cultura y daños en las Américas: reporte final del estudio multicéntrico OPS*. Washington, D.C.: OPS/PAHO; 2007.

233. Basham P, Luik J. Is the obesity epidemic exaggerated? Yes. *BMJ* 2008;336(7638):244-.

234. Jeffery RW, Sherwood NE. Is the obesity epidemic exaggerated? No. *BMJ* 2008;336(7638):245-.

235. De Backer G, Ambrosioni E, Borch-Johnsen K, Brotons C, Cifkova R, Dallongeville J, et al. European guidelines on cardiovascular disease prevention in clinical practice: Third Joint Task Force of European and other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of eight societies and by invited experts). *Eur Heart J* 2003;24(17):1601-10.

236. Barter PJ, Ballantyne CM, Carmena R, Castro Cabezas M, Chapman MJ, Couture P, et al. Apo B versus cholesterol in estimating cardiovascular risk and in guiding therapy: report of the thirty-person/ten-country panel. *J Intern Med* 2006;259(3):247-58.

237. Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52

countries (the INTERHEART study): case-control study. *Lancet* 2004;364(9438):937-52.

238. McQueen MJ, Hawken S, Wang X, Ounpuu S, Sniderman A, Probstfield J, et al. Lipids, lipoproteins, and apolipoproteins as risk markers of myocardial infarction in 52 countries (the INTERHEART study): a case-control study. *Lancet* 2008;372(9634):224-33.

239. Ingelsson E, Schaefer EJ, Contois JH, McNamara JR, Sullivan L, Keyes MJ, et al. Clinical Utility of Different Lipid Measures for Prediction of Coronary Heart Disease in Men and Women. *JAMA* 2007;298(7):776-85.

240. Silventoinen K, Magnusson PK, Neovius M, Sundstrom J, Batty GD, Tynelius P, et al. Does obesity modify the effect of blood pressure on the risk of cardiovascular disease? A population-based cohort study of more than one million Swedish men. *Circulation* 2008;118(16):1637-42.

241. Dyer AR, Elliott P, Shipley M. Body mass index versus height and weight in relation to blood pressure. Findings for the 10,079 persons in the INTERSALT Study. *Am J Epidemiol* 1990;131(4):589-96.

242. Mancilha Carvalho JJ, Baruzzi RG, Howard PF, Poulter N, Alpers MP, Franco LJ, et al. Blood pressure in four remote populations in the INTERSALT Study. *Hypertension* 1989;14(3):238-46.

243. Intersalt Cooperative Research Group. INTERSALT: an international study of electrolyte excretion and blood pressure. Results for 24 hour urinary sodium and potassium excretion. *BMJ* 1988;297(6644):319-28.

244. Poulter NR, Khaw KT, Mugambi M, Peart WS, Rose G, Sever P. Blood pressure patterns in relation to age, weight and urinary electrolytes in three Kenyan communities. *Trans R Soc Trop Med Hyg* 1985;79(3):389-92.

245. Poulter NR, Khaw KT, Mugambi M, Peart WS, Sever PS. Migration-induced changes in blood pressure: a controlled longitudinal study. *Clin Exp Pharmacol Physiol* 1985;12(3):211-6.

246. Ghandehari H, Kamal-Bahl S, Wong ND. Prevalence and extent of dyslipidemia and recommended lipid levels in US adults with and without cardiovascular comorbidities: the National Health and Nutrition Examination Survey 2003-2004. *Am Heart J* 2008;156(1):112-9.

247. Chen Z, Peto R, Collins R, MacMahon S, Lu J, Li W. Serum cholesterol concentration and coronary heart disease in population with low cholesterol concentrations. *BMJ* 1991;303(6797):276-82.

248. Masson LF, McNeill G, Avenell A. Genetic variation and the lipid response to dietary intervention: a systematic review. *Am J Clin Nutr* 2003;77(5):1098-111.

249. Stamler J, Neaton JD. The Multiple Risk Factor Intervention Trial (MRFIT)--Importance Then and Now. *JAMA* 2008;300(11):1343-5.

250. Clarke R, Frost C, Collins R, Appleby P, Peto R. Dietary lipids and blood cholesterol: quantitative meta-analysis of metabolic ward studies. *BMJ* 1997;314(7074):112-.

251. Hegsted DM, Ausman LM, Johnson JA, Dallal GE. Dietary fat and serum lipids: an evaluation of the experimental data [published erratum appears in Am J Clin Nutr 1993 Aug;58(2):245]. *Am J Clin Nutr* 1993;57(6):875-83.

252. Weggemans RM, Zock PL, Katan MB. Dietary cholesterol from eggs increases the ratio of total cholesterol to high-density lipoprotein cholesterol in humans: a meta-analysis. *Am J Clin Nutr* 2001;73(5):885-91.

253. Howell WH, McNamara DJ, Tosca MA, Smith BT, Gaines JA. Plasma lipid and lipoprotein responses to dietary fat and cholesterol: a meta-analysis. *Am J Clin Nutr* 1997;65(6):1747-64.

254. Jacobs DR, Jr., Anderson JT, Hannan P, Keys A, Blackburn H. Variability in individual serum cholesterol response to change in diet. *Arteriosclerosis* 1983;3(4):349-56.

255. Katan MB, Beynen AC, de Vries JH, Nobels A. Existence of consistent hypo- and hyperresponders to dietary cholesterol in man. *Am J Epidemiol* 1986;123(2):221-34.

256. O'Hanesian MA, Rosner B, Bishop LM, Sacks FM. Effects of inherent responsiveness to diet and day-to-day diet variation on plasma lipoprotein concentrations. *Am J Clin Nutr* 1996;64(1):53-9.

257. Hooper L, Summerbell CD, Higgins JPT, Thompson RL, Capps NE, Davey Smith G, et al. Dietary fat intake and prevention of cardiovascular disease: systematic review. *BMJ* 2001;322(7289):757-63.

258. Durrington PN, Prais H. Methods for the prediction of coronary heart disease risk. *Heart* 2001;85(5):489-90.

259. Durrington PN, Prais H, Bhatnagar D, France M, Crowley V, Khan J, et al. Indications for cholesterol-lowering medication: comparison of risk-assessment methods. *Lancet* 1999;353(9149):278-81.

260. Jones AF, Walker J, Jewkes C, Game FL, Bartlett WA, Marshall T, et al. Comparative accuracy of cardiovascular risk prediction methods in primary care patients. *Heart* 2001;85(1):37-43.

261. Stevens RJ, Coleman RL, Shine BL, Holman RR. Could non-HDL cholesterol replace total/HDL cholesterol ratio to estimate coronary heart disease risk in the UKPDS risk engine? *Diabetologia* 2004;47(Suppl 1):A61.

262. Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;24(11):987-1003.

263. Casas JP, Shah T, Hingorani AD, Danesh J, Pepys MB. C-reactive protein and coronary heart disease: a critical review. *J Intern Med* 2008;264(4):295-314.

264. Danesh J, Lewington S, Thompson SG, Lowe GD, Collins R, Kostis JB, et al. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *JAMA* 2005;294(14):1799-809.

265. Iverson C, Christiansen S, Flanagin A, et al. *AMA manual of style: A guide for authors and editors*. 10th ed. New York, NY: Oxford University Press; 2007.

266. Goldstein DE, Little RR, Lorenz RA, Malone JI, Nathan D, Peterson CM, et al. Tests of Glycemia in Diabetes. *Diabetes Care* 2004;27(7):1761-73.

267. Rohlfing C, Wiedmeyer HM, Little R, Grotz VL, Tennill A, England J, et al. Biological variation of glycohemoglobin. *Clin Chem* 2002;48(7):1116-8.

268. Kilpatrick ES, Maylor PW, Keevil BG. Biological variation of glycated hemoglobin. Implications for diabetes screening and monitoring. *Diabetes Care* 1998;21(2):261-4.

269. Meigs JB, Nathan DM, Wilson PW, Cupples LA, Singer DE. Metabolic risk factors worsen continuously across the spectrum of nondiabetic glucose tolerance. The Framingham Offspring Study. *Ann Intern Med* 1998;128(7):524-33.

270. Song Y, Manson JE, Tinker L, Howard BV, Kuller LH, Nathan L, et al. Insulin Sensitivity and Insulin Secretion Determined by Homeostasis Model Assessment and Risk of Diabetes in a Multiethnic Cohort of Women: The Women's Health Initiative Observational Study. *Diabetes Care* 2007;30(7):1747-52.

271. Baracco R, Mohanna S, Seclen S. Determinación de la sensibilidad a la insulina usando el método HOMA en poblaciones adultas habitantes de grandes alturas y a nivel del mar. *Rev Med Hered* 2006;17(4):206-11.

272. McAuley KA, Mann JI, Chase JG, Lotz TF, Shaw GM. Point: HOMA Satisfactory for the Time Being: HOMA: The best bet for the simple determination of insulin sensitivity, until something better comes along. *Diabetes Care* 2007;30(9):2411-3.

273. Boyko EJ, Jensen CC. Do We Know What Homeostasis Model Assessment Measures?: If not, does it matter? *Diabetes Care* 2007;30(10):2725-8.

274. Hockaday D, Sayyad M, Yajnik C. Counterpoint: Appreciating Homeostasis Model Assessment: More useful earlier rather than later. *Diabetes Care* 2007;30(9):2414-8.

275. Haffner SM, Kennedy E, Gonzalez C, Stern MP, Miettinen H. A prospective analysis of the HOMA model. The Mexico City Diabetes Study. *Diabetes Care* 1996;19(10):1138-41.

276. Expert Panel on Detection Evaluation and Treatment of High Blood Cholesterol in Adults. 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* 2002;106(25):3143-.

277. Grundy SM, Cleeman JI, Merz CN, Brewer HB, Jr., Clark LT, Hunninghake DB, et al. Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III guidelines. *Circulation* 2004;110(2):227-39.

278. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, et al. Diagnosis and Management of the Metabolic Syndrome: An American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation* 2005;112(17):2735-52.

279. Gale EAM. Should we dump the metabolic syndrome?: Yes. *BMJ* 2008;336(7645):640-.

280. Alberti KGMM, Zimmet PZ. Should we dump the metabolic syndrome? No. *BMJ* 2008;336(7645):641-.

281. Yudkin JS. Insulin resistance and the metabolic syndrome--or the pitfalls of epidemiology. *Diabetologia* 2007;50(8):1576-86.

282. Reaven GM. The metabolic syndrome: is this diagnosis necessary? *Am J Clin Nutr* 2006;83(6):1237-47.

283. Grundy SM. Metabolic Syndrome: Connecting and Reconciling Cardiovascular and Diabetes Worlds. *J Am Coll Cardiol* 2006;47(6):1093-100.

284. Reaven GM. The Metabolic Syndrome: Requiescat in Pace. *Clin Chem* 2005;51(6):931-8.

285. Kahn R, Buse J, Ferrannini E, Stern M. The metabolic syndrome: time for a critical appraisal. Joint statement from the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetologia* 2005;48(9):1684-99.

286. Meigs JB. Invited commentary: insulin resistance syndrome? Syndrome X? Multiple metabolic syndrome? A syndrome at all? Factor analysis reveals patterns in the fabric of correlated metabolic risk factors. *Am J Epidemiol* 2000;152(10):908-11; discussion 12.

287. Kahn R. Metabolic syndrome--what is the clinical usefulness? *Lancet* 2008;371(9628):1892-3.

288. Sattar N, McConnachie A, Shaper AG, Blauw GJ, Buckley BM, de Craen AJ, et al. Can metabolic syndrome usefully predict cardiovascular disease and diabetes? Outcome data from two prospective studies. *Lancet* 2008;371(9628):1927-35.

289. Lawlor DA, Smith GD, Ebrahim S. Does the new International Diabetes Federation definition of the metabolic syndrome predict CHD any more strongly than older definitions? Findings from the British Women's Heart and Health Study. *Diabetologia* 2006;49(1):41-8.

290. Kelliny C, William J, Riesen W, Paccaud F, Bovet P. Metabolic syndrome according to different definitions in a rapidly developing country of the African region. *Cardiovascular Diabetology* 2008;7(1):27.

291. Lee CM, Huxley RR, Woodward M, Zimmet P, Shaw J, Cho NH, et al. Comparisons of metabolic syndrome definitions in four populations of the Asia-Pacific region. *Metab Syndr Relat Disord* 2008;6(1):37-46.

292. Lorenzo C, Serrano-Rios M, Martinez-Larrad MT, Gonzalez-Sanchez JL, Seclen S, Villena A, et al. Geographic Variations of the International Diabetes Federation and the National Cholesterol Education Program-Adult Treatment Panel

III Definitions of the Metabolic Syndrome in Nondiabetic Subjects. *Diabetes Care* 2006;29(3):685-91.

293. D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* 2008;117(6):743-53.

294. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336(7659):1475-82.

295. Pocock SJ, McCormack V, Gueyffier F, Boutitie F, Fagard RH, Boissel J-P. A score for predicting risk of death from cardiovascular disease in adults with raised blood pressure, based on individual patient data from randomised controlled trials. *BMJ* 2001;323(7304):75-81.

296. Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. *Circulation* 2002;105(3):310-5.

297. Ferrario M, Chiodini P, Chambless LE, Cesana G, Vanuzzo D, Panico S, et al. Prediction of coronary events in a low incidence population. Assessing accuracy of the CUORE Cohort Study prediction equation. *Int J Epidemiol* 2005;34(2):413-21.

298. Wallis EJ, Ramsay LE, Ul Haq I, Ghahramani P, Jackson PR, Rowland-Yeo K, et al. Coronary and cardiovascular risk estimation for primary prevention: validation of a new Sheffield table in the 1995 Scottish health survey population. *BMJ* 2000;320(7236):671-6.

299. Williams B, Poulter NR, Brown MJ, Davis M, McInnes GT, Potter JF, et al. Guidelines for management of hypertension: report of the fourth working party of the British Hypertension Society, 2004-BHS IV. *J Hum Hypertens* 2004;18(3):139-85.

300. Williams B, Poulter NR, Brown MJ, Davis M, McInnes GT, Potter JF, et al. British Hypertension Society guidelines for hypertension management 2004 (BHS-IV): summary. *BMJ* 2004;328(7440):634-40.

301. Brindle P, May M, Gill P, Cappuccio F, D'Agostino R, Sr., Fischbacher C, et al. Primary prevention of cardiovascular disease: a web-based risk score for seven British black and minority ethnic groups. *Heart* 2006;92(11):1595-602.

302. Gaziano TA, Young CR, Fitzmaurice G, Atwood S, Gaziano JM. Laboratory-based versus non-laboratory-based method for assessment of cardiovascular disease risk: the NHANES I Follow-up Study cohort. *Lancet* 2008;371(9616):923-31.

303. World Health Organization. *Prevention of cardiovascular disease. Guidelines for assessment and management of total cardiovascular risk*. Geneva: World Health Organization; 2007.

304. Ezzati M, Hoorn SV, Rodgers A, Lopez AD, Mathers CD, Murray CJ. Estimates of global and regional potential health gains from reducing multiple major risk factors. *Lancet* 2003;362(9380):271-80.

305. World Health Organization. *WHO Global InfoBase*. Geneva: World Health Organization. Available at: http://www.who.int/ncd_surveillance/infobase/web/InfoBaseCommon/.

306. Lindholm LH, Mendis S. Prevention of cardiovascular disease in developing countries. *Lancet* 2007;370(9589):720-2.

307. Mendis S, Mohan V. Non-laboratory-based prediction of cardiovascular risk. *Lancet* 2008;371(9616):878-9.

308. May M, Lawlor DA, Brindle P, Patel R, Ebrahim S. Cardiovascular disease risk assessment in older women: can we improve on Framingham? British Women's Heart and Health prospective cohort study. *Heart* 2006;92(10):1396-401.

309. D'Agostino RB, Sr., Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA* 2001;286(2):180-7.

310. Liu J, Hong Y, D'Agostino RB, Sr., Wu Z, Wang W, Sun J, et al. Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-Provincial Cohort Study. *JAMA* 2004;291(21):2591-9.

311. Christiaens T. Cardiovascular risk tables. *BMJ* 2008;336(7659):1445-6.

312. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;14(1):32-8.

313. Rose G, Day S. The population mean predicts the number of deviant individuals. *BMJ* 1990;301(6759):1031-4.

314. Emberson J, Whincup P, Morris R, Walker M, Ebrahim S. Evaluating the impact of population and high-risk strategies for the primary prevention of cardiovascular disease. *Eur Heart J* 2004;25(6):484-91.

315. Manuel DG, Lim J, Tanuseputro P, Anderson GM, Alter DA, Laupacis A, et al. Revisiting Rose: strategies for reducing coronary heart disease. *BMJ* 2006;332(7542):659-62.

316. World Health Organization. *WHO CVD-risk management package for low – and medium-resource settings*. Geneva: World Health Organization; 2002.

317. Medina-Lezama J, Chirinos JA, Zea-Diaz H, Morey-Vargas O, Bolanos-Salazar J, Corrales-Medina F, et al. Estimaciones del riesgo cardiovascular global en la poblacion adulta de Arequipa metropolitana: Resultados del Estudio PREVENCION. *Rev Per Cardiologia* 2006;XXXII(2):129-44.

318. Barker DJP. *Fetal and Infant Origins of Adult Disease*. London: BMJ Publishing; 1992.

319. Barker DJP. *Mothers, babies and health in later life*. Edinburgh: Churchill Livingstone; 1998.

320. Barker DJ, Osmond C, Forsen TJ, Kajantie E, Eriksson JG. Trajectories of growth among children who have coronary events as adults. *N Engl J Med* 2005;353(17):1802-9.

321. Yajnik CS, Lubree HG, Rege SS, Naik SS, Deshpande JA, Deshpande SS, et al. Adiposity and hyperinsulinemia in Indians are present at birth. *J Clin Endocrinol Metab* 2002;87(12):5575-80.

322. Yajnik CS. Early life origins of insulin resistance and type 2 diabetes in India and other Asian countries. *J Nutr* 2004;134(1):205-10.

323. Yajnik CS. The insulin resistance epidemic in India: fetal origins, later lifestyle, or both? *Nutr Rev* 2001;59(1 Pt 1):1-9.

324. Armitage JA, Khan IY, Taylor PD, Nathanielsz PW, Poston L. Developmental programming of the metabolic syndrome by maternal nutritional

imbalance: how strong is the evidence from experimental models in mammals? *J Physiol* 2004;561(Pt 2):355-77.

325. Gale CR, Martyn CN, Kellingray S, Eastell R, Cooper C. Intrauterine programming of adult body composition. *J Clin Endocrinol Metab* 2001;86(1):267-72.

326. Kensara OA, Wootton SA, Phillips DI, Patel M, Jackson AA, Elia M. Fetal programming of body composition: relation between birth weight and body composition measured with dual-energy X-ray absorptiometry and anthropometric methods in older Englishmen. *Am J Clin Nutr* 2005;82(5):980-7.

327. Langley-Evans SC. Critical differences between two low protein diet protocols in the programming of hypertension in the rat. *Int J Food Sci Nutr* 2000;51(1):11-7.

328. Lucas A. Programming by early nutrition: an experimental approach. *J Nutr* 1998;128(2 Suppl):401S-6S.

329. Phillips DI, Jones A. Fetal programming of autonomic and HPA function: do people who were small babies have enhanced stress responses? *J Physiol* 2006;572(Pt 1):45-50.

330. Singhal A, Wells J, Cole TJ, Fewtrell M, Lucas A. Programming of lean body mass: a link between birth weight, obesity, and cardiovascular disease? *Am J Clin Nutr* 2003;77(3):726-30.

331. Stanner SA, Yudkin JS. Fetal programming and the Leningrad Siege study. *Twin Res* 2001;4(5):287-92.

332. Vehaskari VM, Woods LL. Prenatal programming of hypertension: lessons from experimental models. *J Am Soc Nephrol* 2005;16(9):2545-56.

333. Campos H, Mata L, Siles X, Vives M, Ordovas JM, Schaefer EJ. Prevalence of cardiovascular risk factors in rural and urban Costa Rica. *Circulation* 1992;85(2):648-58.

334. Alberts M, Urdal P, Steyn K, Stensvold I, Tverdal A, Nel JH, et al. Prevalence of cardiovascular diseases and associated risk factors in a rural black population of South Africa. *Eur J Cardiovasc Prev Rehabil* 2005;12(4):347-54.

335. Aspray TJ, Mugusi F, Rashid S, Whiting D, Edwards R, Alberti KG, et al. Rural and urban differences in diabetes prevalence in Tanzania: the role of obesity, physical inactivity and urban living. *Trans R Soc Trop Med Hyg* 2000;94(6):637-44.

336. Mollentze WF, Moore AJ, Steyn AF, Joubert G, Steyn K, Oosthuizen GM, et al. Coronary heart disease risk factors in a rural and urban Orange Free State black population. *S Afr Med J* 1995;85(2):90-6.

337. Steyn K, Bradshaw D, Norman R, Laubscher R. Determinants and treatment of hypertension in South Africans: the first Demographic and Health Survey. *S Afr Med J* 2008;98(5):376-80.

338. van Rooyen JM, Kruger HS, Huisman HW, Wissing MP, Margetts BM, Venter CS, et al. An epidemiological study of hypertension and its determinants in a population in transition: the THUSA study. *J Hum Hypertens* 2000;14(12):779-87.

339. Oosthuizen W, Vorster HH, Kruger A, Venter CS, Kruger HS, de Ridder JH. Impact of urbanisation on serum lipid profiles--the THUSA survey. *S Afr Med J* 2002;92(9):723-8.

340. Sadikot SM, Nigam A, Das S, Bajaj S, Zargar AH, Prasannakumar KM, et al. The burden of diabetes and impaired glucose tolerance in India using the WHO 1999 criteria: prevalence of diabetes in India study (PODIS). *Diabetes Res Clin Pract* 2004;66(3):301-7.

341. Yajnik CS, Joglekar CV, Lubree HG, Rege SS, Naik SS, Bhat DS, et al. Adiposity, inflammation and hyperglycaemia in rural and urban Indian men: Coronary Risk of Insulin Sensitivity in Indian Subjects (CRISIS) Study. *Diabetologia* 2008;51(1):39-46.

342. Tollman SM, Kahn K, Sartorius B, Collinson MA, Clark SJ, Garenne ML. Implications of mortality transition for primary health care in rural South Africa: a population-based surveillance study. *Lancet* 2008;372(9642):893-901.

Appendix A.

Informed consent form, English version

INFORMATION SHEET AND CONSENT FORM FOR A STUDY

Study:	The effect on cardiovascular risk factors of migration from rural to urban areas in Peru
Institutions:	London School of Hygiene and Tropical Medicine (London, United Kingdom) Universidad Peruana Cayetano Heredia (Lima, Peru)
Investigators:	Dr. Jaime Miranda, Dr. Liam Smeeth, Dr. Carlos Cáceres, Dr. Juan Lema

Purpose of the Study.

We would like to invite you to participate in this study aimed to determine those conditions related to diseases of the circulatory system such as obesity, high blood pressure, diabetes and high cholesterol levels among the population.

We are inviting people who have been born in the department of Ayacucho and moved to Lima during the years of political violence. Also, we are inviting people who have not migrated – people born and currently living in Lima and Ayacucho, respectively – in order to compare the conditions mentioned above between these population groups. We hope to carry out this study in a total of 1000 people, both in Ayacucho and Lima. Before making a decision about whether you wish to participate in this study, please read this document and discuss it with the interviewer.

It is now known that there is an increased risk of heart attack (myocardial infarction) or "strokes" (medically known as cerebrovascular accidents) in people who have been diagnosed with one or more of the following diseases: diabetes, high blood pressure, high levels of serum lipids (such as LDL cholesterol and triglycerides) and obesity.

In this study, we would like to make a brief interview to gather general information related to these diseases, take measurements of height, weight and waist, and take a blood sample to determine, through a laboratory study, the conditions we want to study. We will store small quantities of each blood sample in a freezer, to verify any error in the future, such as in the event that we find unexpected results during the study. Also, we will store a small part of your blood samples in order to be able, in the future, to continue with the studies, by comparing our findings with other populations.

This study is financed by a charity called The Wellcome Trust.

Procedures.

If you decide to participate in this study, we will:

- Ask you to sign the consent form enclosed.
- Ask questions about you and your health, and measure your height, weight and waist.
- Measure your blood pressure up to three times
- Ask you to allow us to take, only once, a blood sample equivalent to three teaspoons.
- Carry out a blood analysis, at no cost to you.
- Freeze and store a small portion of each sample (with an identification label that will not include your name)
- Give you the results of your blood tests, for your records.

Reasons for not being included in the study.

You may not participate in the study if you are under 25 years old.

Benefits.

The main benefit of this study is that many people do not know if they suffer from high blood pressure, diabetes or elevated serum lipid levels. Should this be your case, you will know if you have any of these conditions and be able to seek appropriate care. You will not be given any money for participating in this study.

Risks and Discomforts

Nothing in this study represents a great risk to your health. Most questions we will ask in this visit are questions that would be normally asked in a health centre. It will only take a few minutes of your time to answer all the questions. Additional tests shall be made using the blood samples. Blood samples are obtained with minimum

risk, and using the same technique used by professionals in health centres and hospitals. We just need your permission to carry out these tests on your blood and to ask questions about yourself and your health.

Privacy.
For the purposes of this study, all the information about yourself and the results of the tests carried out will be collected and stored in a database in password-protected computers. Even so, you will not be identifiable, as numbers will be used instead of names in the database. However, in order to be able to report the results to participants, your samples will be identifiable by the laboratory staff, as is customary with laboratory tests. Your test results will only be known by the laboratory staff, the main researcher and yourself.

Participation.
You do not have to participate in this study if you do not want to. If you decide not to participate in the study, care provided to you and your family in the health centre or by any other physician will not be affected in any way. If you decide not to participate, no questions about your health will be asked; no weight and height measurements will be made; and no blood samples will be taken.

Questions.
You may ask any questions about anything that is not clear, either now or in the future. You may ask the staff involved in the study or call the phone numbers shown at the beginning of this form. You may also ask questions about the study ethics to the Institutional Ethics Committee of the Universidad Peruana Cayetano Heredia (telephone 319-0005, ext. 2271).

You may contact the investigators responsible for this study and people taking part of it at any time, and for any reason related to this study. For such purpose, please call the following numbers:

Contact:	Dr. Jaime Miranda, Principal Investigator	292 0999

By signing below, you agree that you have understood the contents of this document and decided to participate in this study.

Participant's signature: ……..……..…….
Participant's name: ……

If the Participant is illiterate, then the signature of a Witness is also required:
Witness' signature: ……
Witness' name: ……

Interviewer's signature: ……
Interviewer's name: ……

Researcher's signature: ……
The whole form must be completed on the same date, which must be specified below:

Date: ……..Time:……..

You will be provided with a copy of this document for your records.

Appendix B.

Informed consent form, Spanish version

DOCUMENTO DE INFORMACIÓN Y FICHA DE CONSENTIMIENTO PARA UN ESTUDIO

Estudio:	Efecto de la migración rural-urbana en los factores de riesgo para enfermedad cardiovascular en Perú
Instituciones:	London School of Hygiene and Tropical Medicine (Londres, Reino Unido) Universidad Peruana Cayetano Heredia (Lima, Perú)
Investigadores:	Dr. Jaime Miranda, Dr. Liam Smeeth, Dr. Carlos Cáceres, Dr. Juan Lema

Propósito del Estudio.

Nos gustaría invitarlo a participar en este estudio cuyo propósito es determinar en la población condiciones que están relacionadas con enfermedades del aparato circulatorio, tales como obesidad, presión arterial elevada, presencia de diabetes y colesterol elevado.

Estamos invitando a personas que hayan nacido en el departamento de Ayacucho y que se mudaron a Lima durante los años de violencia política. Así mismo, estamos invitando a personas que no han migrado –personas nacidas y que viven actualmente en Lima y Ayacucho, respectivamente–, a fin de poder comparar las condiciones mencionadas arriba entre estos grupos de poblaciones. Esperamos hacer este estudio en un total de 1000 personas, entre Ayacucho y Lima. Antes de decidir si desea participar en este estudio, por favor lea este documento y discútalo con el entrevistador.

Actualmente se sabe que existe un riesgo incrementado de sufrir ataques al corazón (infartos de miocardio) o de "derrames cerebrales" (conocido médicamente como accidentes cerebro vasculares) en aquellas personas que tienen uno o mas de los siguiente diagnósticos: diabetes, de presión alta, de tener ciertos lípidos elevados (como el LDL colesterol y los triglicéridos) y obesidad.

En este estudio nos gustaría hacerle una entrevista breve para obtener información general relacionada a estas enfermedades, tomar medidas de talla, peso y cintura, y tomar una muestra de sangre para poder determinar mediante un estudio de laboratorio las condiciones que queremos estudiar. Nosotros guardaremos pequeñas cantidades de cada muestra de sangre en un congelador, para verificar en el futuro cualquier error en caso encontremos resultados inesperados durante el estudio. Así mismo, guardaremos una pequeña parte de sus muestras de sangre, para en el futuro, poder continuar los estudios comparando lo encontrado con otras poblaciones.

Este estudio es financiado por una organización caritativa llamada The Wellcome Trust.

Procedimientos.

Si usted decide participar en este estudio, nosotros:

- Le pediremos que firme la ficha de consentimiento adjunta.
- Le haremos preguntas acerca de usted y su salud, y mediremos su talla, peso y cintura.
- Le mediremos la presión arterial hasta en tres oportunidades
- Le pediremos que nos permita sacar una muestra de sangre, solamente una vez, equivalente a tres cucharaditas de te
- Haremos análisis en sangre, sin ningún costo para usted.
- Congelaremos y guardaremos una pequeña porción de cada muestra (con una etiqueta de identificación que no contendrá su nombre).
- Le devolveremos los resultados de sus estudios de sangre, para que usted cuente con una copia personal.

Razones para no ser incluidos en el estudio.

Usted no podrá unirse al estudio si es menor de 25 años.

Beneficios.

El principal beneficio de este estudio es que mucha gente no sabe si sufre de presión alta, diabetes o lípidos elevados. De suceder esto en su caso, usted podrá conocer si tiene algunas de estas condiciones y buscar ayuda especializada respectiva. Usted no recibirá ningún dinero por participar en este estudio.

Riesgos e Incomodidades.

Nada en este estudio representa un gran riesgo a su salud. La mayoría de preguntas que le haremos en esta visita son preguntas que le harían normalmente en un centro de salud. Responder a todas nuestras preguntas le tomará solamente unos cuantos minutos de su tiempo. Las pruebas adicionales serán realizadas usando las muestras de sangre. La obtención de la muestra de sangre tiene riesgos mínimos y se hace utilizando la misma técnica que usan los profesionales en los centros de salud y hospitales. Simplemente necesitamos su permiso para hacer estas pruebas en su sangre, y para hacerle preguntas acerca de usted y su salud.

Privacidad.
Para el propósito de este estudio, toda la información acerca de usted y los resultados de las pruebas que realicemos será recogida y guardada en una base de datos en computadoras protegidas con contraseñas. Incluso así, a usted no se le podrá identificar ya que se usarán números en vez de nombres en la base de datos. Sin embargo, con el propósito de poder reportar los resultados a los participantes, sus muestras podrán ser identificadas por el personal de laboratorio como es rutina con la pruebas de laboratorio. Los resultados de sus pruebas sólo serán de conocimiento del personal se laboratorio, del investigador principal, y de usted.

Participación.
Usted no tiene que participar en este estudio si no lo desea. Si usted decide no participar en el estudio, el cuidado que a usted y su familia se les brinda en el centro de salud o por cualquier otro doctor no será afectado de ninguna manera. Si decide no participar, no se harán las preguntas sobre su salud, ni las mediciones de peso y talla, ni se tomará la muestra de sangre.

Preguntas.
Puede preguntar acerca de cualquier cosa que no entienda, ahora o en el futuro. Usted puede preguntarle al personal del estudio, o puede llamar a los números telefónicos que aparecen al comienzo de esta ficha. Puede también hacer preguntas acerca de la ética del estudio al Comité Institucional de Ética de la Universidad Peruana Cayetano Heredia (teléfono 319-0005, anexo 2271).

Usted puede comunicarse con el investigador responsable de este estudio y las personas que participan en el momento que desee por cualquier motivo relacionado al mismo. Para ello, por favor sírvase llamar a los teléfonos indicados:

Contactos: Dr. Jaime Miranda, Investigador Principal 292 0999

Si firma su nombre abajo, significa que usted ha entendido el contenido de este documento y ha decidido participar de este estudio.

Firma del Participante: ……
Nombre del Participante: ……

Si el Participante no sabe leer, entonces necesitamos también la firma de un Testigo:
Firma del Testigo: ……
Nombre del Testigo: ……

Firma del Encuestador: ……
Nombre del Encuestador: ……

Firma del Investigador: ……
Toda esta ficha debe ser completada en la misma fecha, la cual debe estar especificada aquí:

Fecha: ……..Hora:……..

Nosotros le entregaremos una copia de este documento para sus registros.

Appendix C.
Survey questionnaires: Full questionnaire for participants

Efecto de la migración rural-urbana en los factores de riesgo para enfermedad cardiovascular en Perú

Información sobre la encuesta

DNI del entrevistado └┴┴┴┴┴┴┴┴┘

Código del entrevistado └┴┴┴┴┴┴┘

Sitio y fecha		Respuesta		Código
*1	Código del grupo (Marcar con un círculo)	Rural	1	I1a
		Rural-Urbano	2	I1b
		Rural - Urb - Rural	3	I1c
		Urbano	4	I1d
*2	Nombre del centro/ pueblo donde Nació	Departamento		I2a
		Provincia		I2b
		Distrito		I2c
		Localidad		I2d
*3	Nombre del centro/ pueblo donde se hace la entrevista	Departamento		I3a
		Provincia		I3b
		Distrito		I3c
		Localidad		I3d
4	Identificación del entrevistador (Colocar iniciales)	└┴┴┴┴┘		I4
5	Fecha en que fue rellenado el instrumento	└┴┘ └┴┘ └┴┴┴┘ Día Mes Año		I5

Consentimiento, Entrevista, Idioma y Nombre		Respuesta		Código
6	Se ha leído el consentimiento al entrevistado	Si	1	I6
		No	2 Si NO, leer el consentimiento	
7	Se ha obtenido el consentimiento (verbal o escrito)	Si	1	I7
		No	2 Si NO, Terminar la entrevista	
8	Idioma de la entrevista	Español	1	I8
		Quechua	2	
		Ambos	3	
9	Hora de la entrevista (0-24 horas)	└┴┘ : └┴┘ horas minutos		I9
10	Apellidos completos			I10
11	Nombres completos			I11

Código del entrevistado └┴┴┴┴┴┘ DNI del entrevistado └┴┴┴┴┴┴┴┴┘

	Información adicional	Respuesta		Código
12	Número de teléfono de contacto **(cuando sea posible)**			I12
13	Especificar de qué teléfono se trata	Trabajo 1 Casa 2 Vecino 3 Otro (Especificar) 4		I13
14	Dirección exacta	Localidad		I14a
		Av. / Ca. / Jr.		I14b
		Mz		I14c
		Lote		I14d
		No.		I14e
		Otro, especificar		I14f

	Exámenes de Laboratorio	Respuesta		Código
15	Ha tenido alguna enfermedad la última semana?	Si No	1 2	I15
16	Si respondió Sí, especificar que enfermedad			I16
17	¿Esta enfermedad le provocó la disminución del apetito durante la última semana?	No Levemente Mucho	1 2 3	I17
18	Fecha de la última comida ingerida	Hoy día Ayer	1 2	I18
19	Hora de la última comida	└┴┘ : └┴┘ horas minutos		I19
20	Hora de la toma de muestra de sangre	└┴┘ : └┴┘ horas minutos		I20
21	¿Éxito en la obtención de muestra de sangre en TODOS los tubos?	No Parcialmente Completa	1 2 3	I21
22	¿Hubo necesidad de tomar muestra en tubo tapa ploma?	Si No	1 2	I22
23	**¿Faltó completar algún tubo?**	Si No	1 2	I23
24	Especificar que tubo(s) no fueron completados			I24

	Exámenes de Laboratorio	Respuesta		Código
25	¿Ha sido pesado y tallado?	Si No	1 2	I25
26	¿Se tomó la presión arterial?	Si No	1 2	I26
27	¿Se tomaron las medidas de circunferencias y pliegues corporales?	Si No	1 2	I27

La información contenida en ésta sección debe guardarse separada del cuestionario, ya que contiene información confidencial.

Código Paciente: ☐☐☐☐☐ **DNI del Entrevistado:** ☐☐☐☐☐☐☐☐

Apellidos y Nombres: _____ **Código Laboratorio:** ☐

Mediciones de Antropometría y Presión Arterial			
Peso y Talla			
Peso	☐☐☐.☐ [kg]		
Número de la balanza	☐		
Talla parado	☐☐☐.☐ [cm]		
Largo de pierna	☐☐☐.☐ [cm]		
Altura de la banca	☐☐☐.☐ [cm]		
Talla sentado	☐☐☐.☐ [cm]		
Pliegues corporales	**Medición 1**	**Medición 2**	**Medición 3**
Biceps	☐☐☐[mm]	☐☐[mm]	☐☐mm
Triceps	☐☐☐[mm]	☐☐[mm]	☐☐mm
Subescapular	☐☐☐[mm]	☐☐[mm]	☐☐mm
Suprailiaco	☐☐☐[mm]	☐☐[mm]	☐☐mm
Número de Caliper (Equipo)	☐		
Circunferencia	**Medición 1**	**Medición 2**	**Medición 3**
Cintura	☐☐☐[cm]	☐☐☐[cm]	☐☐☐cm
Cadera	☐☐☐[cm]	☐☐☐[cm]	☐☐☐cm
Número del centímetro	☐		
Información General: Medidas de Antropometría			
¿Medidas en el Lado Izquierdo?	☐ **[1=Si 2=No]**		
Si marcó No, especificar			
¿Todas las medidas adecuadas?	☐ **[1=Si 2=No]**		
Si marcó No, especificar			
Presión Arterial			
	Medición 1	**Medición 2**	**Medición 3**
PA Sistólica	☐☐☐[mmHg]	☐☐[mmHg]	☐☐☐[mmHg]
PA Diastólica	☐☐☐[mmHg]	☐☐[mmHg]	☐☐☐[mmHg]
Pulso	☐☐☐[mmHg]	☐☐[mmHg]	☐☐☐[mmHg]
Cuff usado	☐ **[1=Pequeño; 2=Mediano; 3=Grande]**		
Número del aparato	☐		
¿Medidas en el brazo derecho?	☐ **[1=Si 2=No]**		
¿Todas las medidas adecuadas?	☐ **[1=Si 2=No]**		
Si marcó No, especificar			
Observación:			

MODULO: SOCIOECONÓMICO Código del entrevistado: P --

Apellidos y nombres del entrevistado:_____ **Fecha:**____/____/____

Información del Entrevistador		Respuesta		
1	Identificador del Entrevistador			
2	Fecha de llenado el instrumento (DD/MM/AAAA)	//	//	2007
3	Idioma de la entrevista	Español	1	
		Quechua	2	
		Ambos	3	
4	Hora de la entrevista (00-24 horas)	:		

Información demográfica y socioeconómico		Respuesta		
5	Sexo (Indique hombre o mujer según observe)	Hombre	1	
		Mujer	2	
6	Fecha de nacimiento (DD/MM/AAAA)	//	//	19____
7	Su fecha de nacimiento es	Exacta	1	
		Aproximada	2	
		No sabe/No responde	98	
8	Años cumplidos a la fecha	AÑOS		
9	Su edad es…	Exacta	1	
		Aproximada	2	
		No sabe/No responde	98	
10	¿Cuál es el nivel de educación más alto que ha alcanzado? (LEER LAS OPCIONES)	No fue a la escuela / Analfabeto	1	
		Lee y/o escribe, pero no fue a escuela	2	
		Primaria incompleta	3	
		Primaria completa	4	
		Secundaria incompleta	5	
		Secundaria completa	6	
		Estudios técnicos o superiores incompletos	7	
		Estudios técnicos o superiores completos	8	
		Rehúsa responder	99	
11	¿Cuántos años de estudio ha completado en total, empezando desde la primaria? (sin considerar etapa pre-escolar)	AÑOS		
12	Esta información es:	Exacta	1	
		Aproximada	2	
13	¿Edad que tuvo el ultimo año asistencia a un centro educativo? (SEA PRIMARIA, SECUNDARIA O SUPERIOR)	AÑOS		

14	¿A qué se dedica? Entrevistador deberá indagar y colocar el trabajo del entrevistado (MARCA CON UNA "X" HASTA 3 OPCIONES CONSIDERANDO EL ORDEN DE PRIORIDAD: ESTABILIDAD, TIEMPO, BENEFICIO ECONÓMICO)			1ra Opción	2da Opción	3ra Opción
		Empleada(o) del gobierno				
		Empleada(o) del sector privado				
		Independiente	Comerciante			
			Manual (obrero)			
			Campesino			
		No remunerado				
		Estudiante				
		Ama(o) de casa				
		Jubilado (a)				
		Desempleado (que puede trabajar)				
		Desempleado (que NO puede trabajar)				
		Otro (Especificar) _____				
		_____				
		Rehúsa responder		99		

15	¿La semana pasada estuvo trabajando?	Si	1 (Pase a la pregunta 18)
		No	2
16	La semana pasada...: (LEER LAS OPCIONES Y MARCAR SOLO UNO)	No trabajó, pero tenia trabajo	1 (Pase a la pregunta 18)
		Aunque no trabajó, tiene algún negocio propio	2 (Pase a la pregunta 18)
		Realizó algún cachuelo por un pago en dinero o especie	3 (Pase a la pregunta 18)
		Estuvo ayudando en la chacra, tienda o negocio de algún familiar sin pago alguno	4 (Pase a la pregunta 18)
		No trabajó	5

MODULO: SOCIOECONÓMICO Código del entrevistado: P --

17	La semana pasada estuvo: (LEER LAS OPCIONES Y MARCAR SOLO UNO)	Buscando trabajo, habiendo trabajado antes	1
		Buscando trabajo por primera vez	2
		Estudiando y no trabajó	3
		Viviendo de sus pensión o jubilación y no trabajó	4
		Viviendo de sus rentas y no trabajó	5
		Al cuidado de sus hogar y no trabajó	6
		Otra _____	7
		Rehúsa responder	99
18	¿Cuántas personas en total, incluyéndolo a usted, duernen en su casa?		PERSONAS
19	De estas personas ¿Cuántas son personas mayores de 18 años o tienen 18 años? (Incluirse Ud en la cuenta)		PERSONAS
20	Tomando como referencia el año pasado: ¿Cuál fue el ingreso familiar mensual incluyendo apoyo de los todos los familiares?	1	<= 160 soles (<= $50 dólares americanos)
		2	Entre 161 - 480 soles ($51 - 150)
		3	Entre 481 - 800 soles ($151 - 250)
		4	Entre 801 - 1120 soles ($251 - 350)
		5	Entre 1121 - 1440 soles ($351 - 450)
		6	>= 1441 soles (>= $450)
		9	No sabe
		99	Rehúsa responder
21	¿Cuántas familias que cocinan sus propios alimentos viven en su vivienda?		Número de familias
22	¿Cuántos ambientes de su vivienda se usan para dormir?		Número de ambientes
23	¿Cuál es la fuente principal de abastecimiento de agua que utilizan en su hogar? (LEER LAS OPCIONES)	Caño dentro de la vivienda	1
		Pozo en la casa o lote	2
		Caño o pilón de uso público	3
		Pozo público	4
		Manantial	5
		Río/acequia	6
		Camión, tanque o aguatero	7
		_____Otro	8
		Rehúsa responder	99
24	¿Qué tipo de servicio higiénico tiene su hogar? (LEER LAS OPCIONES)	Conectado a red pública dentro de la vivienda	1
		Conectado a red pública fuera de la vivienda	2
		Letrina o pozo ciego propia	3
		Letrina o pozo ciego común	4
		Río / acequia / canal	5
		No hay servicio / matorral / campo	6
		Rehúsa responder	99
25	¿Tiene en su hogar: (LEER LAS OPCIONES Y MARQUE LAS OPCIONES QUE APLIQUEN)	Cocina a gas ?	1
		Radio?	2
		Televisor blanco y negro?	3
		Televisor a color?	4
		Refrigerador?	5
		Computador?	6
		Teléfono fijo?	7
		Celular?	8
		Cable?	9
		Internet?	10
		Bicicleta?	11
		Motocicleta?	12
		Carro?	13
		Rehúsa responder	99
26	Mayormente, ¿qué tipo de combustible utiliza para cocinar? (LEER LAS OPCIONES)	Electricidad	1
		Gas	2
		Kerosene	3
		Carbón	4
		Leña	5
		Bosta/Carca/Tusa (Coronta)	6
		Rehúsa responder	99

MODULO: SOCIOECONÓMICO Código del entrevistado: | P | | -- | | | |

27	¿Cuál es el material predominante de los pisos de su vivienda? (LEER LAS OPCIONES)	Piso natural: Tierra / Arena	1
		Piso rústico: Entablado	2
		Piso cemento no acabado	3
		Piso acabado: Parquet / Vinílicos / Losetas / Cemento acabado	4
		Rehúsa responder	99
28	¿Cuál es el material predominante de las paredes exteriores de su vivienda? (LEER LAS OPCIONES)	Ladrillo o bloque de cemento	1
		Adobe	2
		Madera	3
		Triplay	4
		Estera	5
		_____Otro	6
		Rehúsa responder	99
29	¿Cuál es el material predominante en los techos de su vivienda? (LEER LAS OPCIONES)	Concreto	1
		Madera	2
		Tejas	3
		Calamina	4
		Eternit	5
		Caña con torta de barro	6
		Ichu	7
		_____Otro	8
		Rehúsa responder	99

Circunstancias del hogar durante la niñez (cuando tenía 10-12 años)

Instrucciones: En el caso de que la persona haya estado en más de un lugar en ese periodo, la pregunta se refiere a un solo lugar, aquel lugar donde el entrevistado pasó la mayor cantidad de tiempo cuando tenía 10 a 12 años

30	¿Cuál de las siguientes opciones describe mejor la actividad laboral principal de su PADRE o TUTOR cuando tenía 10 a 12 años? (LEER LAS OPCIONES)		Empleada(o) del gobierno	1
			Empleada(o) del sector privado	2
		Independiente	Comerciante	3
			Manual (obrero)	4
			Campesino	5
			No remunerado	6
			Estudiante	7
			Ama(o) de casa	8
			Jubilado (a)	9
			Desempleado (que puede trabajar)	10
			Desempleado (que NO puede trabajar)	11
			Otro (Especificar) _____	12
			No aplica (no tenía padre ni tutor)	97
			No sabe/No recuerda	98
			Rehúsa responder	99

31	¿Cuál fue el nivel de educación más alto alcanzado por su PADRE o TUTOR? (LEER LAS OPCIONES)	No fue a la escuela / Analfabeto	1
		Lee, pero no fue a escuela	2
		Primaria incompleta	3
		Primaria completa	4
		Secundaria incompleta	5
		Secundaria completa	6
		Estudios técnicos o superiores incompletos	7
		Estudios técnicos o superiores completos	8
		No aplica (no tenía padre ni tutor)	97
		No sabe/No recuerda	98
		Rehúsa responder	99

MODULO: SOCIOECONÓMICO Código del entrevistado: P [ ][ ] -- [ ][ ][ ]

32	¿Cuál de las siguientes opciones describe mejor la actividad laboral principal de su MADRE o TUTORA cuando tenía 10 a 12 años? (LEER LAS OPCIONES)	Empleada(o) del gobierno		1
		Empleada(o) del sector privado		2
		Independiente	Comerciante	3
			Manual (obrero)	4
			Campesino	5
		No remunerado		6
		Estudiante		7
		Ama(o) de casa		8
		Jubilado (a)		9
		Desempleado (que puede trabajar)		10
		Desempleado (que NO puede trabajar)		11
		Otro (Especificar) _____		12
		No aplica (no tenía madre ni tutora)		97
		No sabe/No recuerda		98
		Rehúsa responder		99
33	¿Cuál fue el nivel de educación más alto alcanzado por su MADRE o TUTORA? (LEER LAS OPCIONES)	No fue a la escuela / Analfabeto		1
		Lee, pero no fue a escuela		2
		Primaria incompleta		3
		Primaria completa		4
		Secundaria incompleta		5
		Secundaria completa		6
		Estudios técnicos o superiores incompletos		7
		Estudios técnicos o superiores completos		8
		No aplica (no tenía madre ni tutora)		97
		No sabe/No recuerda		98
		Rehúsa responder		99
34	¿Cuántos ambientes de su vivienda eran usados para dormir?	[ ][ ] Número de ambientes		
35	¿Cuantas personas en total dormían en estas habitaciones?	[ ][ ] Número de personas		
36	¿Cuál era la fuente principal de abastecimiento de agua que utilizaban en su hogar en ese entonces? (LEER LAS OPCIONES)	Caño dentro de la vivienda		1
		Pozo en la casa o lote		2
		Caño o pilón de uso público		3
		Pozo público		4
		Manantial		5
		Río/Acequia		6
		Camión, tanque o aguatero		7
		_____Otro		8
		No sabe/No recuerda		98
37	¿Qué tipo de servicio higiénico tenía su hogar en ese entonces? (LEER LAS OPCIONES)	Conectado a red pública dentro de la vivienda		1
		Conectado a red pública fuera de la vivienda		2
		Letrina o pozo ciego propia		3
		Letrina o pozo ciego común		4
		Río / Acequia / Canal		5
		No hay servicio / Matorral / Campo		6
		No sabe/No recuerda		98
38	¿Tenía en su hogar...: (LEER LAS OPCIONES Y MARQUE LAS OPCIONES QUE APLIQUEN)	Cocina a gas ?		1
		Radio?		2
		Televisor blanco y negro?		3
		Televisor a color?		4
		Refrigerador?		5
		Teléfono fijo?		6
		Bicicleta?		7
		Motocicleta?		8
		Carro?		9
		No sabe/No recuerda		98

MODULO: SOCIOECONÓMICO Código del entrevistado: P | | -- | | | |

39	Mayormente, ¿qué tipo de combustible se utilizaba para cocinar? (LEER LAS OPCIONES)	Electricidad	1
		Gas	2
		Kerosene	3
		Carbón	4
		Leña	5
		Bosta/Carca/Tusa (Coronta)	6
		Rehúsa responder	99
40	¿Cuál era el material predominante de los pisos de su vivienda en ese entonces? (LEER LAS OPCIONES)	Piso natural: Tierra / Arena	1
		Piso rústico: Entablado	2
		Piso cemento no acabado	3
		Piso acabado: Parquet / Vinílicos / Losetas / Cemento acabado	4
		No sabe/No recuerda	98
41	¿Cuál era el material predominante de las paredes exteriores de su vivienda en ese entonces? (LEER LAS OPCIONES)	Ladrillo o bloque de cemento	1
		Adobe	2
		Madera	3
		Triplay	4
		Estera	5
		Otro	6
		No sabe/No recuerda	98
42	¿Cuál era el material predominante en los techos de su vivienda en ese entonces? (LEER LAS OPCIONES)	Concreto	1
		Madera	2
		Tejas	3
		Calamina	4
		Eternit	5
		Caña con torta de barro	6
		Ichu	7
		Otro	8
		No sabe/No recuerda	98

MODULO: FACTORES DE RIESGO Código del entrevistado: P | | -- | | | |

Apellidos y nombres del entrevistado:_____ **Fecha:**_____/_____/_____

Consumo de tabaco		Respuesta	
1	¿Alguna vez ha probado cigarrillos, aunque sea una o dos pitadas?	Si	1
		Si, una sola vez para probar (Pase a la pregunta 7)	2
		No (Pase a la pregunta 7)	3
2	¿Ha fumado por lo menos 100 cigarrillos en toda su vida?	Si	1
		No	2
		No recuerdo/No estoy seguro	98
3	¿Actualmente fuma Ud. cigarrillos?	Sí (uno o más cigarrillos diariamente)	1
		Ocasionalmente (menos de un cigarrillo por día)	2
		No (he dejado de fumar)	3
4	¿Qué edad tenía cuando comenzó a fumar cigarrillos por primera vez en su vida?		Años
		No recuerda/No está seguro	98
5	¿Cuándo fue la última vez que fumó un cigarrillo?	Menos de 1 mes	1
		Entre 1 y 6 meses (Pase a la preg. 7)	2
		Entre 6 y 12 meses (Pase a la preg. 7)	3
		Un año y más (Pase a la preg. 7)	4
		No recuerda/No está seguro	98
6	¿Cuántos cigarrillos fumó en total en los últimos treinta días?		Cigarrillos último mes
		No recuerda/No está seguro	98

Consumo de alcohol		Respuesta	
7	¿Con qué frecuencia consumió alguna bebida alcohólica en el último año?	Nunca (Pasar a la pregunta 11)	1
		Una o menos veces al mes	2
		De 2 a 4 veces al mes	3
		De 2 a 3 veces a la semana	4
		4 o más veces a la semana	5
		Rehúsa responder	99
8	En el último año, ¿con qué frecuencia ha tomado 6 o más tragos que contengan alcohol en una misma ocasión?	Nunca	1
		Menos de una vez al mes	2
		Mensualmente	3
		Semanalmente	4
		A diario o casi a diario	5
		Rehúsa responder	99
9	¿Con qué frecuencia tiene Ud. resaca?	Nunca o casi nunca	1
		Menos de una vez al mes	2
		Una vez al mes	3
		Varias veces al mes	4
		Una vez a la semana	5
		Varias veces a la semana	6
		Todos los días	7
		Rehúsa responder	99
10	Si tuviera que calificar su consumo de alcohol, Ud. diría que mayormente es: (MARCAR SOLO UNA)	Acompañando las comidas	1
		Mayoría de fines de semana o vacaciones	2
		Momentos o motivos ocasionales	3
		Rehúsa responder	99

MODULO: FACTORES DE RIESGO Código del entrevistado: P | | -- | | | |

Antecedentes de enfermedad y tratamiento		Respuesta			
11	¿Ha sufrido Ud. alguna vez de alguna de éstas enfermedades? (DIAGNOSTICADA POR ALGÚN PROFESIONAL DE LA SALUD) (LEER LAS OPCIONES Y MARCAR TODAS LAS QUE APLICAn) (HASTA ANTES DE LA TOMA DE LA MUESTRA DE SANGRE PARA ESTE ESTUDIO)	DIAGNÓSTICO		Profesional: 1 (Médico) 2 (Enfermero) 3 (Farmacéutico) 4 (Otro…)	Lugar del diagnóstico: 1 (Hospital) 2 (C. o P. de salud) 3 (Clínica) 4 (Otro ...)
		Presión arterial elevada	1 [Sí] 2 [No]		
		Enfermedad del corazón	1 [Sí] 2 [No]		
		Diabetes	1 [Sí] 2 [No]		
		Derrame cerebral (DCV)	1 [Sí] 2 [No]		
		Tuberculosis	1 [Sí] 2 [No]		
		Asma	1 [Sí] 2 [No]		
12	Toma alguna medicación o tratamiento especifico para controlar ... (HASTA ANTES DE LA TOMA DE LA MUESTRA DE SANGRE PARA ESTE ESTUDIO)	TOMA PARA...	1 (Si) 2 (No) 3 (No aplica)	MEDICAMENTO Indicar el medicamento que toma	
		Presión arterial elevada			
		Enfermedad del corazón			
		Diabetes			
		Derrame cerebral (DCV)			
		Tuberculosis			
		Asma			

Percepciones sobre obesidad y estado de salud		Respuesta	
13	¿Considera Ud. que para su edad su peso es …?	1	Bajo de peso
		2	Normal
		3	Sobrepeso
		4	Obeso
14	Comparado con otras personas de su edad ¿Su estado de salud es…?	1	Muy bueno
		2	Bueno
		3	Regular
		4	Malo
		5	Muy Malo

MODULO: MIGRACIÓN Código del entrevistado: P | | -- | | | |

11	¿Quién tomó la decisión de irse? (REFERIDA A LA PRIMERA SALIDA)	Yo mismo	1
		Mis padres / mis hermanos	2
		Mis familiares de mi pueblo	3
		Mis familiares en otra ciudad	4
		Mi pareja	5
		Otras personas	6
		Especificar quién:_____	
12	¿Qué edad tenía cuando dejó su lugar de origen, por primera vez, por un periodo de más de 6 meses? (COLOCAR 00 PARA LOS MENORES DE 1 AÑO)		(Años)
		Esta información es:	
		Exacta	1
		Aproximada	2
13	¿Ud. inició sus estudios en su lugar de origen?	Sí	1
		No	2
		Rehúsa responder	99
		No aplica/No procede	98
14	¿Qué nivel de educación tenia cuando dejó su lugar de origen, por primera vez, por un periodo de más de 6 meses? (REFERIDA A LA PRIMERA SALIDA)	No fue a la escuela / Analfabeto	1
		Lee, pero no fue a escuela	2
		Primaria incompleta	3
		Primaria completa	4
		Secundaria incompleta	5
		Secundaria completa	6
		Estudios técnicos o superiores incompletos	7
		Estudios técnicos o superiores completos	8
		Rehúsa responder	99
		No aplica/No procede	98
15	Esta salida de su lugar de origen, ¿tuvo algún impacto en su educación?	Sí, pudo continuar estudios	1
		Sí, porque tuvo que parar estudios	2
		No cambió en nada	3
		Rehúsa responder	99
		No aplica/No procede	98
16	¿Qué edad tenía cuando regresó a su lugar de origen para quedarse hasta la actualidad? (SOLO PARA EL GRUPO AYACUCHANOS RETORNANTES A SU LUGAR DE ORIGEN)		Años
		Esta información es:	
		Exacta	1
		Aproximada	2
17	¿En cuántos lugares en total ha vivido por más de 6 meses seguidos desde que dejó su lugar de origen hasta antes de establecerse en este lugar? (NO CONTAR LA RESIDENCIA ACTUAL)		Lugares
		Esta información es:	
		Exacta	1
		Aproximada	2
		No recuerda	98
		Rehúsa responder	99
18	De todas sus mudanzas fuera del distrito donde se encontraba, ¿alguna ha sido por motivo de terrorismo o violencia política?	Sí	1
		No	2
		Rehúsa responder	99
19	En el último año, ¿ha vivido Ud. temporalmente fuera de su casa por un periodo de 2 o más meses?	Sí	1
		No	2 (Pasa a pregunta 21)
		Rehúsa responder	99 (Pasa a pregunta 21)

MODULO: MIGRACIÓN Código del entrevistado: P | | -- | | | |

20	¿Cuántos meses en total?		Meses
		Esta información es:	
		Exacta	1
		Aproximada	2
21	En promedio, ¿cuántos años de su vida ha vivido en una zona rural?		Años
		Esta información es:	
		Exacta	1
		Aproximada	2
22	En promedio, ¿cuántos años de su vida ha vivido en una zona urbana?		Años
		Esta información es:	
		Exacta	1
		Aproximada	2
23	¿Qué edad tenía cuando llegó a Lima? (SOLO PARA LOS MIGRANTES EN LIMA)		Años
		Esta información es:	
		Exacta	1
		Aproximada	2
		No aplica	97
24	¿Cuánto tiempo en total lleva viviendo en Lima? (SOLO PARA LOS MIGRANTES EN LIMA)		Años
		Esta información es:	
		Exacta	1
		Aproximada	2
		No aplica	97
25	¿Cuánto tiempo lleva viviendo en esta propiedad?		Años
		Esta información es:	
		Exacta	1
		Aproximada	2

Preferencias de lugar para vivir		Respuesta	
26	Si Ud. tuviera la oportunidad, ¿dónde preferiría vivir? (LEER OPCIONES)	Comunidad	1
		Pueblo	2
		Cuidad pequeña	3
		Ciudad grande	4
27	¿Cuál sería la razón principal por tal preferencia? (MARCAR SOLO UNA)	Económica / Trabajo	1
		Servicios (educación, diversión, salud, accesibilidad)	2
		Razones familiares	3
		Modo o estilo de vida	4
		Otra	5
		Especificar: _____	

Appendix D.
Survey questionnaires: Participant's rejection short-form

CUESTIONARIO PARA SUJETOS QUE RECHAZAN EL ESTUDIO

Sujeto ID

☐ Centro
☐ Comunidad
☐ Casa

1. Razón de no-entrevista

☐ No disponible ☐ Rehusar ☐ Otro

2. ¿La persona está dispuesta a completar el cuestionario? ☐ No ☐ Si

3. ¿Cuántas personas viven en su casa? ☐ En total

4. ¿Cuál es su edad? ☐ Años

5. Grado de instrucción del paciente

☐ 1. No fue a la escuela ☐ 2. Primaria incompleta ☐ 3. Primaria completa

☐ 4.Secundaria incompleta ☐ 5.Secundaria completa ☐ 6.Estudios técnicos o superiores incompletos ☐ 7.Estudios técnicos o superiores completos ☐ 8. Rehusa

6. ¿Ud. fuma actualmente? ☐ 1. Si ☐ 2.No

7. ¿Ud. consume alcohol usualmente durante los fines de semana? ☐ 1. Si ☐ 2. No

8. Historia médica del paciente

Nota: Si no está seguro de uno o más de los siguientes ítems por favor marcar la casilla de desconocido.

	Desconocido	No	Si
Diabetes	☐	☐	☐
Cualquier enfermedad cardiovascular	☐	☐	☐
Cáncer	☐	☐	☐

9. ¿Le ha dicho algún médico que tiene usted la presión arterial alta o hipertensión?

☐ **1. Si** ☐ **2. No**

Si su respuesta es NO pasar a la pregunta 11

10. Siendo diagnosticada de Hipertensión alta, Ud. recibe alguna medicación

☐ **1. Si** ☐ **2. No**

11. ¿Cuál es el material predominante de los pisos de su vivienda?

☐ 1. Piso natural: Tierra / Arena ☐ 2. Piso rústico: Entablado ☐ 3. Piso cemento no acabado

4. Piso acabado: Parquet /Vinílicos /Losetas / Cemento acabado

12. ¿Cuál es el material predominante de las paredes exteriores de su vivienda?

☐ 1. Ladrillo ☐ 2.Adobe ☐ 3. Madera ☐ 4.Triplay ☐ 5. Estera ☐ 6. Otro

13. ¿Cuál es el material predominante del techo de su vivienda?

☐ 1.Concreto ☐ 2.Madera ☐ 3.Tejas ☐ 4. Calamina ☐ 5. Caña con torta de barro

☐ 6.Otro

14. **¿Vive Ud. en el lugar en que nació?**

☐ 1. Si ☐ 2.No

Si la respuesta es Si, terminar la entrevista.

15. ¿**Qué edad tenía cuando dejó el lugar?** ☐ Años

16. ¿Cuál fue el motivo principal de su salida de su lugar de origen?

☐ 1.Estudios ☐ 2.Trabajo ☐ 3.Amenaza de terrorismo ☐ 4.Temor al terrorismo

☐ 5.Maltrato, violencia familiar y/o terrorismo ☐ 6.Enfermedad ☐ **7.**Matrimonio

☐ 8.Alquiler de vivienda en otro lugar ☐ 9. Alojamiento con familiares

☐ 10.Comprar terreno/mudarse a casa propia ☐ 11.Otro.....

17. Nombre del entrevistador:

Nombres y Apellidos

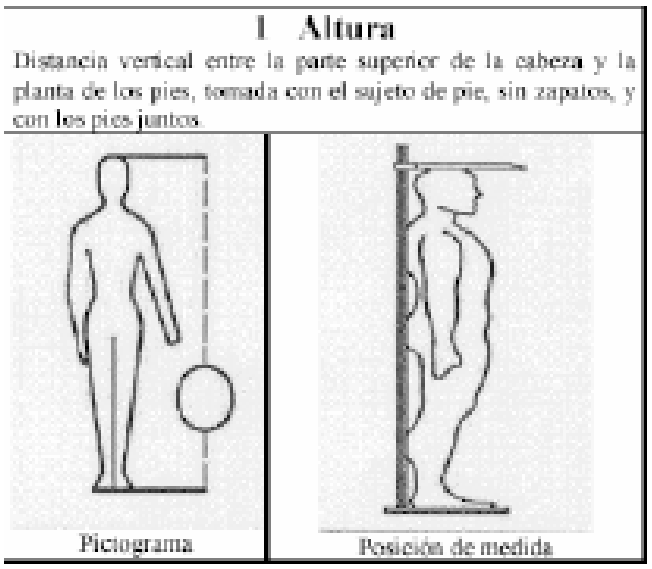
Appendix E.
Standard operations procedures for measurements

MANUAL DE PROCEDIMIENTOS PARA LA TOMA DE MEDICIONES

1. **PESO (balanza digital):**
   - Colocar la balanza en un lugar fijo y plano.
   - Si no se cuenta con una superficie lisa, colocar la balanza en la base del tallímetro.
   - Explicar a la persona que debe estar en posición de pie, con ambos pies ligeramente separados dentro de la balanza, y mirando de frente.
   - Solicitar a la persona que se saque los zapatos.
   - Ayudar a la persona a subirse a la balanza con la menor cantidad de ropa posible.
   - Leer el peso en kilos y gramos, realizar el descuento respectivo de la ropa.
   - Informar resultados a la persona y registrar resultado en la ficha.

2. TALLA DE PIE:

- Explicar el procedimiento a la persona que se está atendiendo y buscar un lugar plano en el cual se pueda apoyar la parte vertical del tallímetro hasta que se encuentre fijo en la superficie. Usar el nivelador adjunto para corroborar que la superficie donde se apoya el tallímetro sea completamente horizontal.



- Pedir a la persona que se coloque sin zapatos con los pies ligeramente separados, con la mirada al frente[1], los brazos pegados al cuerpo y dando la espalda al tallímetro. Los talones, pantorrillas, glúteos, espalda y cabeza deben estar pegados al tallímetro
- Apoyar la superficie horizontal móvil del tallímetro en la parte más alta de la cabeza, sin ejercer presión y pedir a la persona que tome aire.
- Leer la talla en centímetros.
- Informar a la persona y registrar resultados en la ficha.

***Para el registro de los datos: todas las medidas de longitud** –talla, talla sentado, largo de pierna, circunferencia de cintura y cadera- **se redondearán al último milímetro completado (1mm).** En el caso de pliegues cutáneos, se redondearán al último 0.2mm completado.*

[1] Para mantener la mirada al frente constatar que la parte superior del conducto auditivo externo este en el mismo plano horizontal con el suelo de la orbita y esto sea perpendicular al plano vertical del tallímetro (Plano Horizontal de Frankfurt)

3. **TALLA SENTADO:**

- Es la distancia entre el punto más alto de la cabeza y el piso medida en cm.
- La medición se realiza con la persona sentada en una superficie perfectamente plana y horizontal, y de preferencia, que permita el apoyo de los pies, manteniendo los músculos de las piernas y glúteos relajados. Las rodillas deben estar flexionadas en una posición cómoda. La cabeza es sostenida en el plano de Frankfurt. Deslizar la superficie horizontal del tallímetro hasta que toque la cabeza.
- Descontar la altura de la silla donde está sentado el individuo.

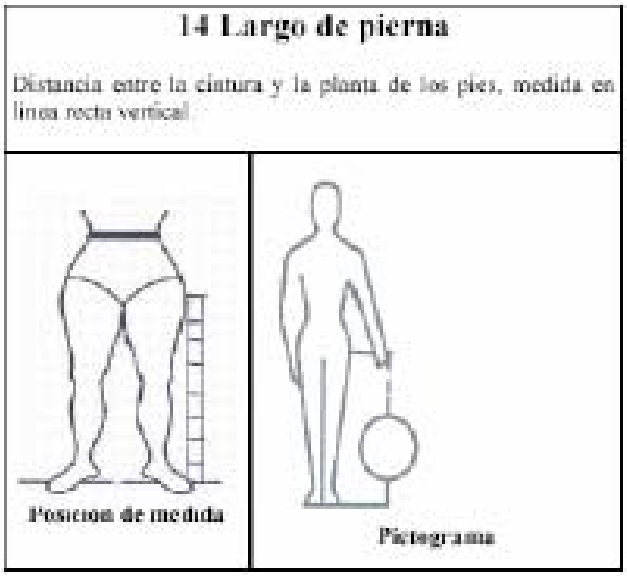
*Para el registro de los datos: todas las medidas de longitud –talla, talla sentado, largo de pierna, circunferencia de cintura y cadera- **se redondearán al último milímetro completado (1mm).** En el caso de pliegues cutáneos, se redondearán al último 0.2mm completado.*

En general, se recomienda utilizar una silla estándar única con una medida de altura de la pata ya establecida y medir a todas las personas por igual con la misma silla.

*Si se usan diferentes sillas, registrar la altura de cada una de ellas. Por precaución, **siempre debe registrarse la talla sentada y la altura de la silla utilizada.***

4. **LARGO DE PIERNA:**

- La longitud de pierna es la distancia que existe entre la cintura hasta la planta de los pies (Ver figura).
- El sujeto debe colocarse sin zapatos en posición vertical apoyándose en ambos pies.
- La cinta debe pasar sobre el maléolo externo (tobillo), permaneciendo paralela a la pierna hasta llegar al borde inferior del pie.

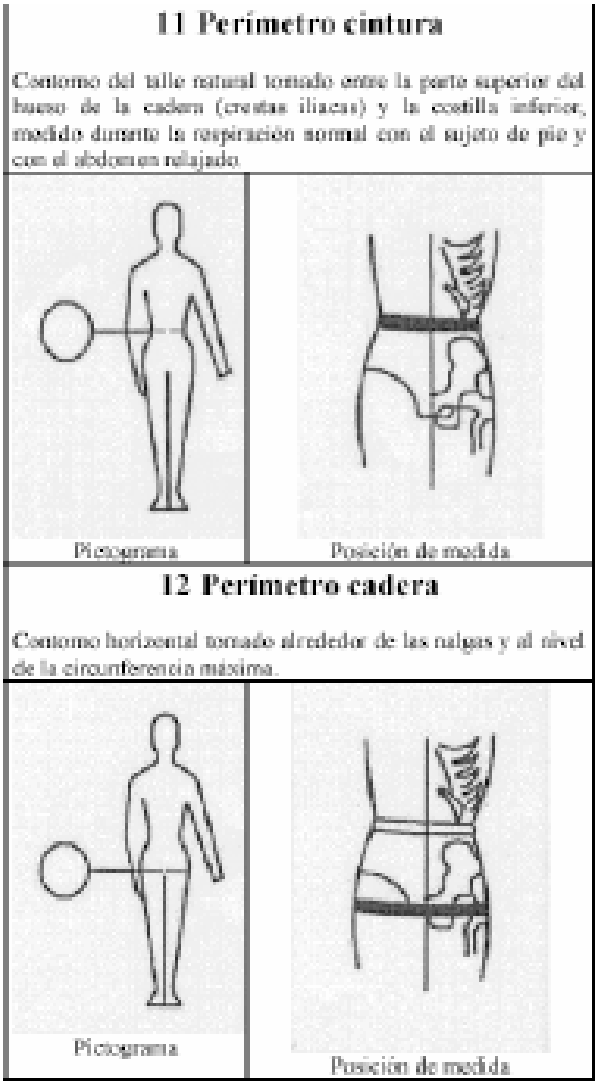


- Finalmente, se realiza la lectura en centímetros registrándose la misma en centímetros.

Para el registro de los datos: todas las medidas de longitud –talla, talla sentado, largo de pierna, circunferencia de cintura y cadera- se redondearán al último milímetro completado (1mm). *En el caso de pliegues cutáneos, se redondearán al último 0.2mm completado.*

5. **INDICE CINTURA CADERA:**

- La medición de las circunferencias se realiza con una cinta métrica, en posición de pie y con la persona en ropa interior.
- La circunferencia abdominal se mide en el punto medio entre la ultima costilla y la cresta ilíaca.
- Para medir la circunferencia glútea pasar la cinta métrica por la parte mas prominente de los glúteos (nalgas).
- El índice cintura cadera se calcula de acuerdo a la siguiente fórmula: ICC = circunferencia abdominal / circunferencia glútea.



***Para el registro de los datos: todas las medidas de longitud** –talla, talla sentado, largo de pierna, circunferencia de cintura y cadera- **se redondearán al último milímetro completado (1mm).** En el caso de pliegues cutáneos, se redondearán al último 0.2mm completado.*

PULSO Y PRESION ARTERIAL:

- Antes de iniciar todo se deben reducir todas las distracciones y ruidos cercanos (perros, alto volumen, etc.). En el caso de los niños, tomar la presión arterial cuando no esté llorando.
- Previo a la toma de la presión arterial tomar del pulso del paciente. Para esto se colocan los dos primeros dedos sobre la arteria radial (en la muñeca) por espacio de un minuto.
- La presión arterial conviene medirla en el brazo derecho, estando el paciente sentado, cómodo y relajado con los pies en el suelo y apoyado en el respaldar de la silla. Debe haber descansado unos 5 minutos y no haber consumido café, alcohol o haber fumado en los 30 minutos anteriores.
- El mango se aplica en la mitad del brazo derecho (el borde inferior queda unos 2 a 3 cm. sobre el pliegue del codo). Debe quedar bien aplicado y no suelto (ya que esto último favorecería lecturas falsamente elevadas). El brazo debe estar desnudo, sin ropa que comprima o dificulte su colocación. Conviene que el brazo esté apoyado sobre una mesa o que cuelgue relajado al lado del cuerpo. El mango del tensiómetro (bolsa de goma) debe quedar ubicada de tal forma que justo la mitad de ella quede sobre la arteria braquial (borde interno del brazo). Además, el brazo con el manguito una vez colocado debe estar a la altura del corazón.
- La medición debe realizarse teniendo la campana del estetoscopio apoyada en el pliegue cubital, sobre la arteria braquial.
- La aparición de los primeros ruidos correspondientes a latidos del pulso determina la presión sistólica auscultatoria. Después de identificar la presión sistólica auscultatoria, se sigue desinflando el manguito hasta que desaparecen los ruidos. Este momento corresponde a la presión diastólica. En ocasiones, primero los ruidos se atenúan y luego desaparecen. En general se considera como la presión diastólica el momento en que los ruidos desaparecen.

- Si se realiza más de una toma de Presión Arterial hacerlo con un intervalo mínimo de 5 minutos.
- Registrar en la ficha los valores tanto de presión sistólica como diastólica obtenidos en cada una de las tomas
- Informar a la persona el valor de Presión Arterial obtenido y en caso de ser varias tomas informar el promedio.

Es importante recalcar a la persona ***que los valores obtenidos en la medición de la presión arterial de ninguna manera deben ser interpretados como un diagnóstico.*** *Para tener un diagnostico definitivo debe recomendarse la visitar a su médico.*

- Si no tenemos el mango adecuado para una persona muy obesa, no medir la presión pues nos dará un valor errado y registrar que el mango no funciono de manera adecuada.
- En pediatría, se debe utilizar un manguito de manómetro de tamaño adecuado que cubra 80% a 100% la circunferencia del brazo y, por lo menos el 40% del largo del brazo.

6. **PLIEGUES CUTANEOS:**

- Para realizar la medición se utiliza el lipocalibrador o caliper, y su resultado se obtiene en milímetros.
- El pliegue cutáneo (usualmente conocido como "rollos" o "gordos") debe ser sujetado fuertemente entre el pulgar y el índice izquierdo del observador. Ambos dedos deben estar separados lo suficiente entre si como para permitir la inclusión de todo el tejido adiposo subyacente en el pliegue.
- Se toma el instrumento con la mano derecha, aplicándolo sobre el pliegue a 1cm de los dedos de la mano izquierda, de tal manera que solamente las caras del caliper -y no la de los dedos del observador- ejerzan presión sobre el pliegue.
- Cuando el caliper está en posición correcta, el observador relaja los dedos de su mano derecha para que el instrumento pueda ejercer su máxima presión. La toma del pliegue con los dedos de la mano izquierda debe ser mantenida.
- La lectura se efectúa sobre el dial hasta el último quinto de mm (0,2 mm) completo. **Una vez aplicado el caliper sobre los pliegues, mantenerlo por 2 segundos antes de efectuar la lectura**.
- Los puntos más utilizados comúnmente son:
  a) **TRICIPITAL:** El brazo debe estar relajado y ligeramente flexionado, con la palma hacia adelante. La medición se toma a nivel del punto medio entre el acromion (hueso prominente del hombro) y el olécranon (hueso del codo), en la superficie posterior del músculo triceps, sobre una línea paralela al brazo, que pasa por el olécranon. El pliegue debe tomarse alrededor de 1 cm por encima del nivel al cual se efectuará la medición
  b) **BICIPITAL:** En la cara anterior del brazo a la misma altura que la medición del pliegue tricipital.
  c) **SUBESCAPULAR:** La persona debe pararse con la espalda desnuda y los brazos relajados a los costados. El antropometrista pasa su índice izquierdo a lo largo del borde medial de la escápula hacia el ángulo

inferior; el pulgar, entonces, toma el pliegue. Éste es ligeramente oblicuo, a un centímetro por debajo del ángulo inferior de la escápula derecha.

d) **SUPRAILIACO**: Dos centímetros por encima de la cresta iliaca izquierda (punto más alto del hueso de la cadera), en la línea media axilar.

- La toma de los pliegues tricipital y bicipital se realizará en el **brazo no dominante** (brazo que se use menos) y se debe registrar en la ficha en que brazo se hizo la medición.
- **Se deben medir todos los pliegues necesarios hasta completar un circuito.** Repetir el circuito de mediciones 3 veces. No se puede medir el mismo pliegue dos veces seguidas.
- Registrar las 3 medidas en las fichas.

Para el registro de los datos: *todas las medidas de longitud –talla, talla sentado, largo de pierna, circunferencia de cintura y cadera- se redondearán al último milímetro completado (1mm).* ***En el caso de pliegues cutáneos, se redondearán al último 0.2mm completado.***

7. Estandarización de las mediciones

*La **estandarización de las personas que realizan las mediciones debe realizarse periódicamente**. En nuestro caso este procedimiento se realizará mensualmente. Para este proceso de estandarización no se usará un Gold Standard único sino el promedio de todas las observaciones realizadas por el mismo observador. Cada observador deberá medir a un mínimo de 10 sujetos, dos veces por sujeto.*

Appendix F.
Laboratory performance overview

In this section, and as general examples, independent quality control evaluations for the laboratory used in the study are presented. These evaluations correspond to blind assessment of samples for the glucose, HDL, total cholesterol and insulin tests.

BIO-RAD

Laboratory Performance Overview
Assayed Chemistry

Lab 196801 | **Lot 14110**

MEDLAB CANTELLA COLICHON SAC
William Cherre CONTROL DE CALIDAD
AV. Santa Cruz 367
Miraflores
L 17 LIMA 1

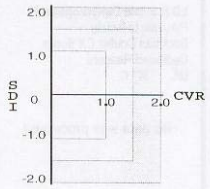
Data For: 11–2005
Lot Exp: 01–31–2006
Printed: 01–11–2006
Page 5

UNITY

Level 1

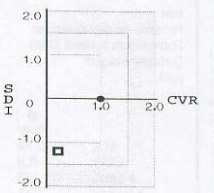
Digoxin
EIA
Abbott AxSYM
Dedicated Reagent
ng/mL

No data was processed.



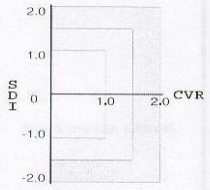
GGT (Gamma Glutamyltransferase)
G–glutamyl–carboxy–nitroanilide–IFCC
Roche Hitachi Systems (USA)
IFCC/RPPHS (CRM 470) Standardization
U/L 37° C

	● Peer	□ Method
SDI:	0.00	–1.21
CVR:	1.0	0.2
# Labs:	1	39
# Points:	5	908

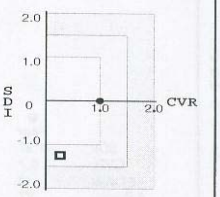


Level 2

No data was processed.



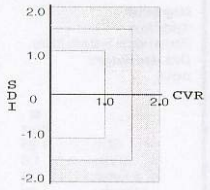
	● Peer	□ Method
SDI:	0.00	–1.25
CVR:	1.0	0.3
# Labs:	1	40
# Points:	5	864



Level 1

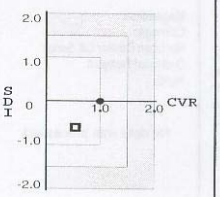
GGT (Gamma Glutamyltransferase)
G–glutamyl–p–nitroanilide
Beckman Coulter CX Series
Dedicated Reagent
U/L 37° C

No data was processed.



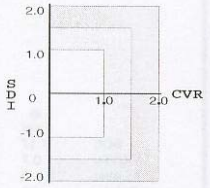
Glucose
Glucose oxidase, hydrogen peroxide (Trinder)
Roche Hitachi Systems (USA)
IFCC/RPPHS (CRM 470) Standardization
mg/dL

	● Peer	□ Method
SDI:	0.00	–0.61
CVR:	1.0	0.5
# Labs:	1	70
# Points:	5	1171

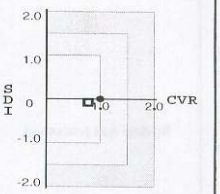


Level 2

No data was processed.



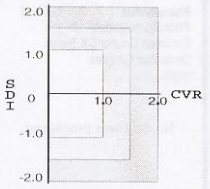
	● Peer	□ Method
SDI:	0.00	–0.09
CVR:	1.0	0.8
# Labs:	1	68
# Points:	5	1127



Level 1

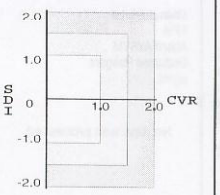
Glucose
Glucose oxidase, oxygen consumption
Beckman Coulter CX Series
Dedicated Reagent
mg/dL

No data was processed.



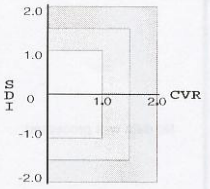
Iron
Ferrozine–no deproteinization
Beckman Coulter CX Series
Dedicated Reagent
µg/dL

No data was processed.

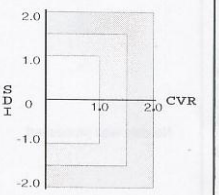


Level 2

No data was processed.



No data was processed.



COPYRIGHT 1991 BIO–RAD LABORATORIES, ALL RIGHTS RESERVED.

Laboratory Performance Overview
Assayed Chemistry

BIO-RAD

Lab 196801

MEDLAB CANTELLA COLICHON SAC
William Cherre CONTROL DE CALIDAD
AV. Santa Cruz 367
Miraflores
L 17 LIMA 1

Lot 14110

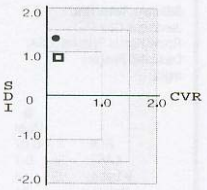
Data For: 11-2005
Lot Exp: 01-31-2006
Printed: 01-11-2006
Page 4

UNITY

Level 1

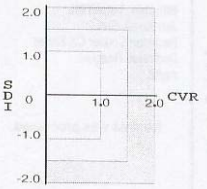
Cholesterol, HDL
Direct measure-PEG
Roche Hitachi Systems (USA)
Dedicated Reagent
mg/dL

	Peer	Method
SDI:	1.31	0.86
CVR:	0.2	0.2
# Labs:	5	38
# Points:	95	617



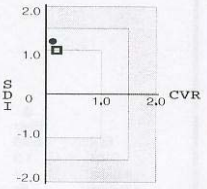
Cholesterol, HDL
Direct measure, polymer-polyanion
Beckman Coulter CX Series
Dedicated Reagent
mg/dL

No data was processed.

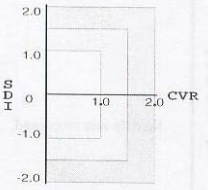


Level 2

	Peer	Method
SDI:	1.20	1.00
CVR:	0.1	0.2
# Labs:	5	36
# Points:	93	600



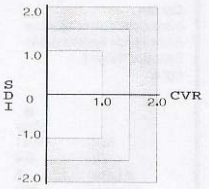
No data was processed.



Level 1

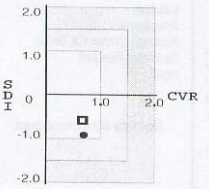
Cholesterol, Total
Cholesterol oxidase, esterase, peroxidase
Beckman Coulter CX Series
Dedicated Reagent
mg/dL

No data was processed.



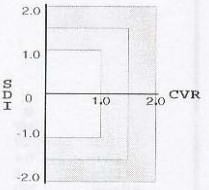
Cholesterol, Total
Cholesterol oxidase, esterase, peroxidase
Roche Hitachi Systems (USA)
Dedicated Reagent
mg/dL

	Peer	Method
SDI:	−0.92	−0.59
CVR:	0.7	0.7
# Labs:	11	247
# Points:	218	4936

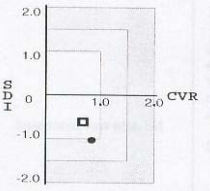


Level 2

No data was processed.



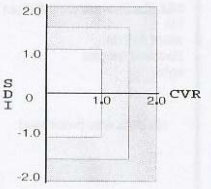
	Peer	Method
SDI:	−1.03	−0.62
CVR:	0.8	0.7
# Labs:	11	239
# Points:	219	4736



Level 1

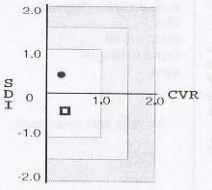
Creatinine
Alkaline picrate-kinetic
Beckman Coulter CX Series
Dedicated Reagent
mg/dL

No data was processed.



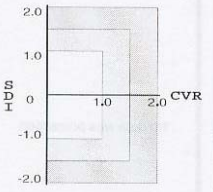
Creatinine
Alkaline picrate-kinetic
Roche Hitachi Systems (USA)
Dedicated Reagent
mg/dL

	Peer	Method
SDI:	0.42	−0.41
CVR:	0.3	0.3
# Labs:	5	221
# Points:	84	5515

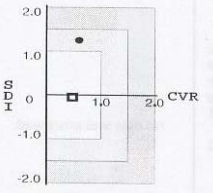


Level 2

No data was processed.



	Peer	Method
SDI:	1.24	−0.06
CVR:	0.6	0.5
# Labs:	4	216
# Points:	71	5360



COPYRIGHT 1991 BIO-RAD LABORATORIES, ALL RIGHTS RESERVED.

BIO-RAD

Posicionamiento del Laboratorio
Immunoassay Plus

Lab. 196801

MEDLAB CANTELLA COLICHON SAC
William Cherre CONTROL DE CALIDAD
AV. Santa Cruz 367
Miraflores
L 17 LIMA 1

Lote 40160

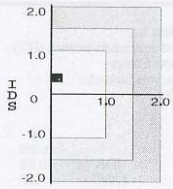
Datos de: 01-2006
Caducidad: 06-30-2007
Impresos: 03-28-2006

UNITY

Nivel 1

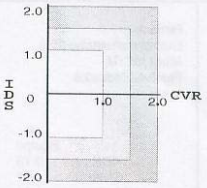
Insulina
Electroquimioluminiscencia (ECL)
Roche Elecsys
Reactivos Dedicados
µUI/mL

	● Equipo	□ Método
IDS:	0.40	0.37
CVR:	0.1	0.1
N° Lab.:	10	13
N° Valores:	52	85



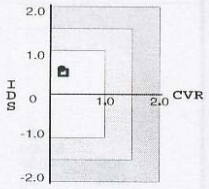
LH
Electroquimioluminiscencia (ECL)
Roche Elecsys
Reactivos Dedicados
mUI/mL

No han sido procesados valores.

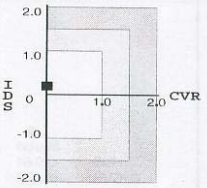


Nivel 2

	● Equipo	□ Método
IDS:	0.58	0.49
CVR:	0.2	0.2
N° Lab.:	10	12
N° Valores:	46	64

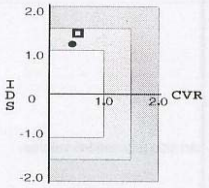


	● Equipo	□ Método
IDS:	0.21	0.19
CVR:	0.0	0.0
N° Lab.:	25	40
N° Valores:	315	707

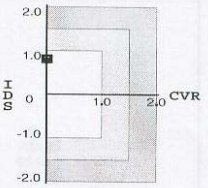


Nivel 3

	● Equipo	□ Método
IDS:	1.14	1.38
CVR:	0.4	0.5
N° Lab.:	8	11
N° Valores:	35	67



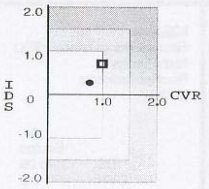
	● Equipo	□ Método
IDS:	0.77	0.81
CVR:	0.0	0.0
N° Lab.:	23	38
N° Valores:	281	689



Nivel 1

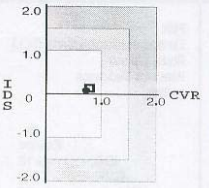
Prolactina
Electroquimioluminiscencia (ECL)
Roche Elecsys
Reactivos Dedicados
ng/mL

	● Equipo	□ Método
IDS:	0.27	0.70
CVR:	0.8	1.0
N° Lab.:	21	37
N° Valores:	217	624



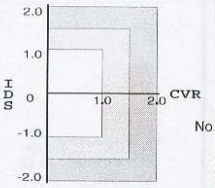
PSA, Libre
Electroquimioluminiscencia (ECL)
Roche Elecsys
Reactivos Dedicados
ng/mL

	● Equipo	□ Método
IDS:	0.08	0.13
CVR:	0.7	0.8
N° Lab.:	13	22
N° Valores:	148	284

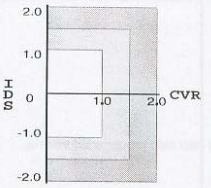


Nivel 2

No han sido procesados valores.

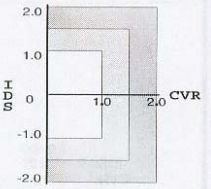


No han sido procesados valores.

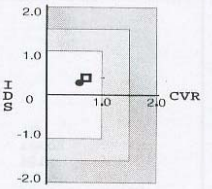


Nivel 3

No han sido procesados valores.

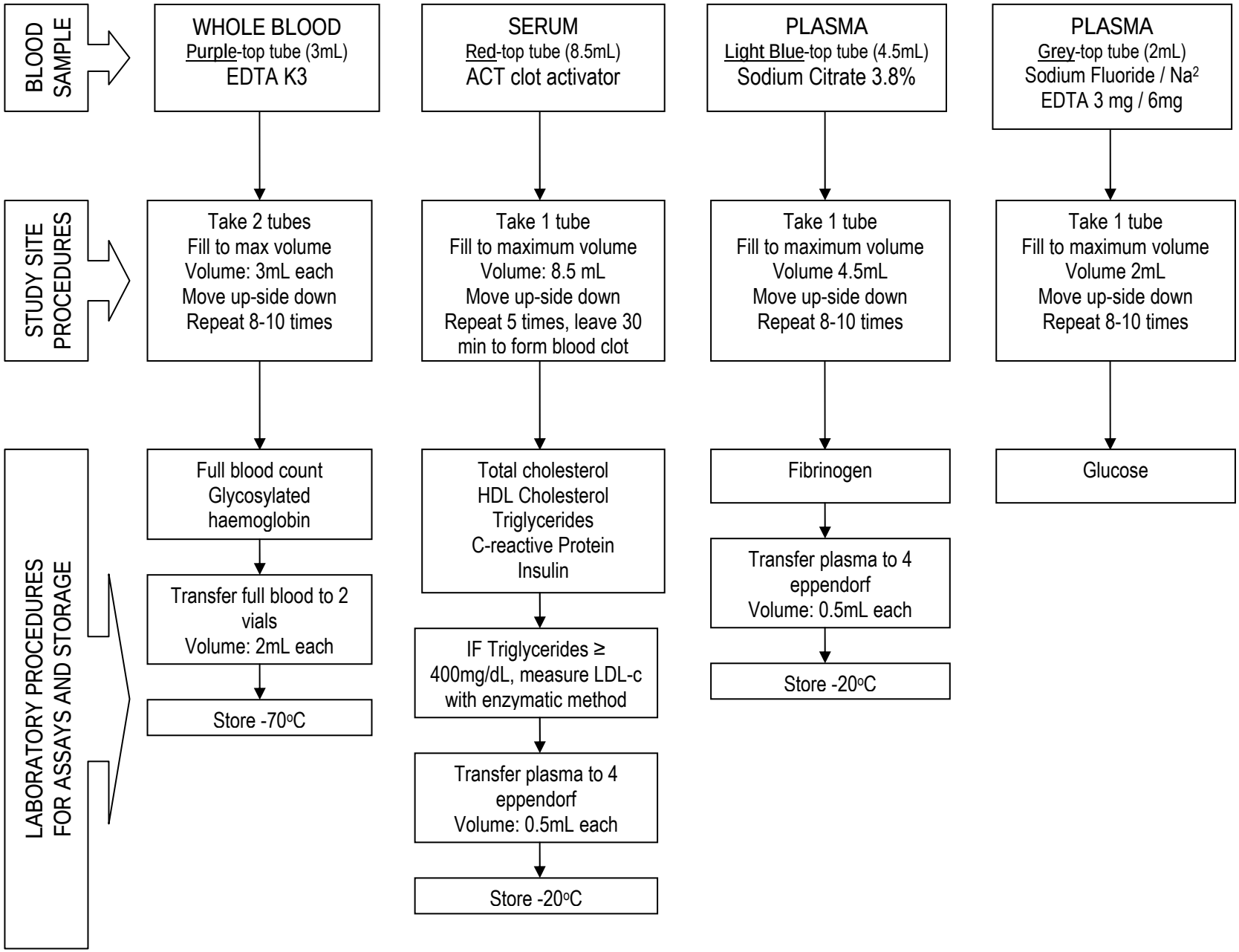


	● Equipo	□ Método
IDS:	0.28	0.38
CVR:	0.6	0.7
N° Lab.:	13	20
N° Valores:	145	264



Derechos de copia 1991 Bio-Rad Laboratories, Todos los derechos reservados.

Appendix G.
Field and laboratory processes for blood testing and storage of samples


BLOOD SAMPLE
STUDY SITE PROCEDURES
LABORATORY PROCEDURES FOR ASSAYS AND STORAGE
WHOLE BLOOD
Purple-top tube (3mL)
EDTA K3
Take 2 tubes
Fill to max volume
Volume: 3mL each
Move up-side down
Repeat 8-10 times
Full blood count
Glycosylated haemoglobin
Transfer full blood to 2 vials
Volume: 2mL each
Store -70°C
SERUM
Red-top tube (8.5mL)
ACT clot activator
Take 1 tube
Fill to maximum volume
Volume: 8.5 mL
Move up-side down
Repeat 5 times, leave 30 min to form blood clot
Total cholesterol
HDL Cholesterol
Triglycerides
C-reactive Protein
Insulin
IF Triglycerides ≥ 400mg/dL, measure LDL-c with enzymatic method
Transfer plasma to 4 eppendorf
Volume: 0.5mL each
Store -20°C
PLASMA
Light Blue-top tube (4.5mL)
Sodium Citrate 3.8%
Take 1 tube
Fill to maximum volume
Volume 4.5mL
Move up-side down
Repeat 8-10 times
Fibrinogen
Transfer plasma to 4 eppendorf
Volume: 0.5mL each
Store -20°C
PLASMA
Grey-top tube (2mL)
Sodium Fluoride / Na2
EDTA 3 mg / 6mg
Take 1 tube
Fill to maximum volume
Volume 2mL
Move up-side down
Repeat 8-10 times
Glucose