Analysis of dietary patterns and cross-sectional and longitudinal associations with hypertension, high body mass index and type 2 diabetes in Peru

Carmelia Alae-Carew[1*], Pauline Scheelbeek[1], Rodrigo M. Carrillo-Larco[2,3], Antonio Bernabé-Ortiz[4], William Checkley[5,6], J. Jaime Miranda[2]

[1]Department of Population Health, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK.
[2]CRONICAS Centre of Excellence in Chronic Diseases, Universidad Peruana Cayetano Heredia, Lima, Peru; School of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru.
[3]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.
[4]CRONICAS Centre of Excellence in Chronic Diseases, Universidad Peruana Cayetano Heredia, Lima, Peru; School of Public Health and Administration, Universidad Peruana Cayetano Heredia, Lima, Peru.
[5]Division of Pulmonary and Critical Care, School of Medicine, Johns Hopkins University, Baltimore, USA
[6]Center for Global Non-Communicable Disease Research and Training, Johns Hopkins University, Baltimore, USA

*Corresponding author: Email carmelia.alae-carew@lshtm.ac.uk

Short Title: Dietary Patterns and Health Outcomes in Peru

Acknowledgements

The authors would like to thank Dr Rosemary Green of LSHTM for her advice and guidance on the use of Mplus for latent class analysis. Thanks is also owed to the CRONICAS staff members and field teams for their collection and provision of the data.

Financial Support

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflict of Interest

None

Authorship

The research question was formulated by C.A.-C. in consultation with P.S. and J.J.M. The data analysis was carried out primarily by C.A.-C. with guidance on the statistical analysis strategy from P.S. and advice on the dataset and aspects of the CRONICAS cohort study from R.M.C.L. The article was written by C.A.-C. incorporating contributions from P.S., J.J.M and R.M.C.L. Contributions to writing were also provided by A.B.-O. and W.C. in addition to guidance on statistical methods and analysis.

Ethical Standards Disclosure

This study was conducted according to the guidelines laid down in the Declaration of Helsinki and all procedures involving research study participants were approved by the Ethics Committees at Universidad Peruana Cayetano Heredia and A.B. PRISMA in Lima, Peru, and Institutional Review Board at the Johns Hopkins Bloomberg School of Public Health, in Baltimore, USA. Verbal informed consent was obtained from all subjects.

Abstract

Objective

To determine if specific dietary patterns are associated with risk of hypertension, type 2 diabetes mellitus (T2DM), and high BMI in four sites in Peru.

Design

We analysed dietary patterns from a cohort of Peruvian adults in four geographical settings using latent class analysis. Associations with prevalence and incidence of hypertension, T2DM and high BMI were assessed using Poisson regression and generalised linear models, adjusted for potential confounders.

Setting

Four sites in Peru varying in degree of urbanisation.

Subjects

Adults aged ≥35 (n=3,280).

Results

We identified four distinct dietary patterns corresponding to different stages of the Peruvian nutrition transition, reflected by the foods frequently consumed in each pattern. Participants consuming the "stage 3" diet, characterised by high proportional consumption of processed foods, animal products and low consumption of vegetables, mostly consumed in the semi-urban setting, showed the highest prevalence of all health outcomes (hypertension 32.1%; T2DM 10.7%; and high BMI 75.1%). Those with a more traditional "stage 1" diet characterised by potato and vegetables, mostly consumed in the rural setting, had lower prevalence of hypertension (PR 0.57, 95% CI 0.43-0.75), T2DM (PR 0.36, 95% CI 0.16-0.86) and high BMI (PR 0.55, 95% CI 0.48-0.63) compared to the "stage 3" diet. Incidence of hypertension was highest among individuals consuming the stage 3 diet (63.75 per 1000 person-years, 95% CI 52.40-77.55).

Conclusions

This study found more traditional diets were associated with a lower prevalence of three common chronic diseases, whilst prevalence of these diseases was higher with a diet high in processed foods and low in vegetables.

Keywords

Dietary pattern analysis; latent class analysis; Peru; non-communicable diseases; cardiometabolic risk factors

Introduction

Unhealthy diet is one of the main modifiable risk factors for the predominant non-communicable diseases (NCDs), as it can contribute to the development of conditions such as hypertension and type 2 diabetes mellitus (T2DM), as well as overweight and obesity which are precursors to many NCDs[1]. The rise of unhealthy diets in low- and middle-income countries (LMICs) has been in part attributed to a rapid ‘nutrition transition’, in which traditional diets are being replaced by consumption of more energy dense Westernised foods, including animal products high in saturated fat, and processed foods high in salt, oils and refined sugar[2, 3].

The nutrition transition in Peru, an upper-middle income country, has been taking place at different rates throughout the country due to its diverse geography, nutrition profile and levels of urbanisation[4]. The diet in rural areas of Peru remains in line with a more traditional dietary pattern, which is mostly comprised of potatoes and other tubers[5], and starchy seeds such as quinoa[6], although there are substantial regional differences in diet. However, as 77.3% of its inhabitants now live in urban areas[7], this is no longer the case for much of the country; urbanisation being one the drivers underpinning the change in consumption that characterises the nutrition transition towards a Westernised dietary pattern[3]. Examination of food intake trends inferred from food balance sheets over the 1990’s have shown increasing consumption of animal products, saturated fat and sugar[2], while energy supply from cereals, roots and tubers has declined as Gross Domestic Product (GDP) per capita has improved[8]. Mapping of the “stages” of transition in 2012 based on prevalence of stunting and obesity suggested that some coastal regions of the country had moved away from traditional diets and showed low levels of stunting and a high burden of obesity and other nutrition related chronic disease. Diets in most other regions of Peru were much more in transition, showing a typical “double-burden” profile with a substantial burden of both stunting and obesity among people consuming a diet similar to the traditional Peruvian diet and people consuming more Westernised diets respectively [4].

In 2014, NCDs accounted for 66% of deaths in Peru[7]. Hypertension and T2DM are two of the leading risk factors for many of the major NCDs, and further characterisation of the estimated local burden of these risk factors is required to develop and support local and regional strategies for prevention and control of NCDs[9]. Previous studies in Peru have shown that both overweight and obesity increased the risk of hypertension and T2DM, with obesity being the leading risk factor for both[10, 11]. The contribution of obesity to disease risk was found to vary in different parts of the country, and similarly variations in incidence and prevalence of hypertension[10, 11] and T2DM[11, 12] were found amongst different parts of the country that vary in degree of urbanisation. In 2012, the prevalence of overweight and obesity in Peru was >30% in all but one of its 25 regions[4]. However, there has been very little examination of the role of diet in development of obesity and other NCD determinants in a country with extensive environmental and cultural diversity.

The link between nutrition and disease is increasingly being investigated using dietary pattern analysis which takes into account many of the complexities associated with examining the diet; it is inclusive of eating behaviours, food synergy, and nutrient interactions[13]. Whilst it does not take the place of study of single food components, dietary pattern analysis can provide valuable information on the overall effects of diet in order to predict disease risk or aid in a comprehensive approach to prevention strategies[13, 14]. No previous dietary pattern analyses have been performed on Peruvian dietary data, a country where diets are heterogeneous and can be difficult to classify. Therefore this study was undertaken to achieve the following objectives: to characterise dietary patterns in three different locations in Peru; to examine the cross-sectional relationship of these dietary patterns with hypertension, T2DM and body mass index (BMI); and to investigate changes in disease risk over time by assessing the longitudinal association between baseline dietary patterns and the three outcomes. By doing this we aim to contribute to the evidence on the link between dietary patterns and NCDs in resource-constrained settings to further inform targeted intervention strategies.

Methods

Study Design and Setting

The CRONICAS Cohort Study is a longitudinal ongoing cohort study of non-communicable disease progression in distinct geographic areas of Peru, which has been described elsewhere[15]. In short, four study sites were included in the study: two sea-level sites (a highly urbanised area of Lima, the capital, and a semi-urban setting in Tumbes, in the north of the country) and two high-altitude sites (an urban and a rural site in Puno in the south of the country). Beginning in September 2010, the study sought to characterise the baseline prevalence and rate of progression of cardiopulmonary diseases and their risk factors among these different populations. To date, the study has had two rounds of follow-up over a 30-month period. The first follow up comprised repeated clinical measurements only, whilst the second follow up consisted of repeated clinical measurements and blood sampling. For this analysis, we used data from baseline and second follow-up, hence covering a 30-month period.

Participants

Using the most recent census available in all study sites, participants were randomly selected using a sex and age-stratified sampling strategy. Only one subject per household was selected. Those eligible were 35 years or older and had to be a permanent resident in the selected area. Those who were pregnant, bedridden, unable to provide consent, had active tuberculosis, or had a physical disability that would prevent measurement of clinical outcomes were excluded. The CRONICAS Cohort Study protocol and informed consent forms were approved by the Ethics Committees at Universidad Peruana Cayetano Heredia and A.B. PRISMA in Lima, Peru, and Institutional Review Board at the Johns Hopkins Bloomberg School of Public Health, in Baltimore, USA.

Data Collection

At baseline, participants were visited at home by fieldworkers to verify eligibility criteria and obtain informed consent. Socioeconomic, dietary and lifestyle information was gathered by fieldworker-administered paper-based questionnaire, which was adapted from the WHO STEPwise approach for non-communicable disease risk factor surveillance[16]. Within this was contained a short-version of a food frequency questionnaire (FFQ), based on a similar questionnaire designed by an earlier study[17], to obtain information on consumption frequency of certain foods and beverages, selected either because they are commonly eaten foods in Peru, or known to be linked to chronic disease[18]. In completing the FFQ participants were asked to report how many times a month or week they consumed foods and beverages within 23 categories (Supplement A). Subjects were subsequently seen for clinical assessment including blood sampling by a trained technician according to standardised techniques and protocols. Blood pressure was measured using a previously validated automatic monitor (model: OMRON HEM-780)[19]. Three readings were taken and an average of the last 2 measurements was used in the analysis. Blood samples were analysed for fasting plasma glucose level using an enzymatic colorimetric method (GOD-PAP; Modular P-E/Roche-Cobas, Grenzach-Whylen, Germany). Weight was measured using a body composition analyser (model: TBF-300A; TANITA Corporation, Tokyo, Japan). Second follow up comprised repeated completion of subsections of the baseline questionnaire (not including FFQ) and repeated clinical assessment including blood sampling following the same procedures conducted at baseline.

Variables

The outcomes of interest at baseline and follow up were hypertension, T2DM and high BMI. Hypertension was defined as systolic blood pressure (SBP) ≥140 mmHg or diastolic blood pressure (DBP) ≥90 mmHg in participants younger than 60 years old, and SBP ≥150mmHg or DBP ≥90 mmHg in those aged 60 years or older, as per the eighth Joint National Committee (JNC8) target blood pressure recommendations[20]. Subjects self-reporting a diagnosis of hypertension by a physician or currently prescribed anti-hypertensive medications were also included in the definition of hypertension. T2DM was defined as fasting plasma glucose level ≥126 mg/dL[21], self-reported diagnosis by a physician, or currently prescribed medications for T2DM. High BMI was considered to be BMI ≥25 kg/m^2 to incorporate both overweight (≥25.0-29.9 kg/m^2) and obese (≥30.0 kg/m^2) subjects as per WHO international classification of BMI[22]. All outcome variables were binary.

Dietary patterns were derived from responses to the short version FFQ. Data were aggregated into 14 food groups (Supplement A) from which daily frequency of consumption was calculated. Foods were grouped based on similarities in the way the foods are consumed, and similar health impacts. Unrealistic extreme intake values of over 10 times per day were excluded. Each food group variable was split into four categories (zero consumption and tertiles of frequency) to reflect proportional consumption frequency of each food group prior to dietary pattern analysis[23, 24].

Analyses were adjusted for age (35-44, 45-54, 55-64, 65-74, >75 years old), sex, ethnicity (native Quechua/Aymara, half Spanish/Native, other), education level (none, primary, secondary and further education), currently employed, socioeconomic status (based on wealth index indicator derived from assets and household facilities; categorised into tertiles), smoking (never, former, current), heavy alcohol consumption (≥2 nights in the past month of heavy drinking, defined as ≥6 alcoholic drinks), physical activity (low, moderate, high; based on metabolic equivalent score), and television (TV) watching (<2 or ≥ 2 hours per day). Baseline measurements were assumed to stay constant throughout the 30-month follow-up period except age, which was updated to age at last follow up for follow-up analyses.

Statistical Methods

Latent class analysis (LCA)[23, 25-27] was used to identify dietary patterns from a set of observed categorical variables. Meaningful latent classes or subgroups of individuals were created based on shared patterns of consumption. Individuals were assigned to a certain dietary pattern group based on similar dietary characteristics. LCA was performed using Mplus version 7.4 (Muthen & Muthen, Los Angeles, CA, USA) using the 14 categorical food group variables. A series of models were generated with increasing number of classes from one to seven. Model selection was based on Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) to compare goodness of fit, entropy to assess the certainty of classification, and pattern interpretability of each model[23, 28]. From the selected model, each individual's most likely class membership, as determined by posterior probabilities, was exported into the dataset as a new variable. Dietary patterns for each of the classes were described based on the conditional probabilities of reported food groups within each class[26, 29].

Tabulation and univariate analyses were used to explore the prevalence of the outcomes and check for collinearity between the exposure variables. Generalised linear models assuming logistic-normal distribution were used to explore the relationship between dietary patterns and prevalence of the three disease outcomes, adjusting for potential confounders using forward selection stepwise regression; likelihood ratio tests (LRTs) were used to assess goodness of fit. The effect of site was assessed using variables for urbanisation (urban, semi-urban, rural) and site (sea-level, altitude) within the models; goodness of fit was determined by AIC and BIC. Stratified analysis by site was also carried out to further explore within site associations, however there was not enough data to perform individual analysis in each site. For all models, prevalence ratio (PR) and 95% confidence intervals (CI) were obtained for each dietary pattern category, with the dietary pattern showing the highest outcome prevalence chosen as the reference group. Ethnicity was not included in the final model due to collinearity with socioeconomic status and education level.

Generalised linear models assuming Poisson distribution using random censoring were used to determine the association of overall and site-specific dietary patterns with incidence of the three outcomes, generating crude and adjusted incidence risk ratio (IRR) and 95% CIs. Subjects who already had the condition of interest at baseline were excluded. Analyses of the relationship between dietary patterns and hypertension, T2DM and obesity were performed using STATA 15.0 (StataCorp LLC, College Station, Texas, USA). Analysis code can be made available on request (https://datacompass.lshtm.ac.uk/).

Results

Population's characteristics

Of the 3,280 subjects, 1,064 (32.4%) were residing in Lima, 599 (18.3 %) urban Puno, 586 (17.9%) rural Puno and 1,031 (31.4%) Tumbes. Baseline data on hypertension was complete for 3,266 (99.6%) subjects, diabetes 3,134 (95.6%) subjects, and BMI 3,112 (94.9%) subjects (Figure 1). Detailed characteristics of the study population at baseline are shown in Table 1.

Table 1: Baseline participant characteristics of Peruvian adults in the CRONICAS cohort study 2010-2013 (*n*=3,280) and their distribution amongst the study sites

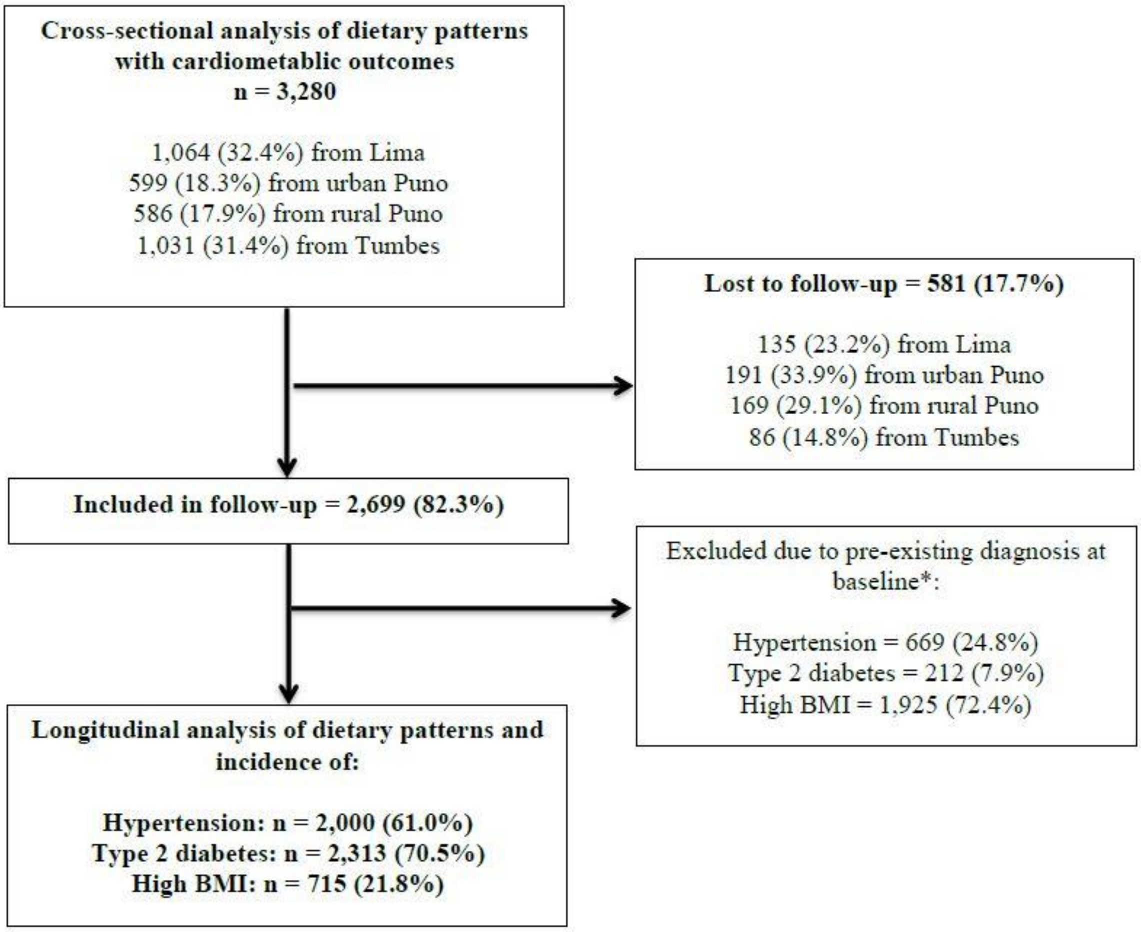
Sociodemographic characteristic		Overall No. (%)	Lima No. (%)	Urban Puno No. (%)	Rural Puno No. (%)	Tumbes No. (%)
Sex	Female	1,694 (51.7)	550 (51.7)	312 (52.1)	313 (53.4)	519 (50.3)
	Male	1,583 (48.3)	514 (48.3)	286 (47.8)	271 (46.3)	512 (49.7)
Age Group (years)	35-44	785 (23.9)	250 (23.5)	143 (23.9)	131 (22.4)	261 (25.3)
	45-54	835 (25.5)	288 (27.1)	146 (24.4)	150 (25.6)	251 (24.4)
	55-64	832 (25.4)	269 (25.3)	154 (25.7)	148 (25.3)	261 (25.3)
	65-74	550 (16.8)	196 (18.4)	112 (18.7)	94 (16.0)	148 (14.4)
	75-92	278 (8.5)	61 (5.7)	44 (7.4)	63 (10.8)	110 (10.7)
Education Level	None	242 (7.4)	86 (8.1)	27 (4.5)	68 (11.6)	61 (5.9)
	Primary	1,258 (38.4)	369 (34.7)	69 (11.5)	308 (52.6)	512 (49.7)
	Secondary	1,071 (32.7)	424 (39.9)	165 (27.6)	175 (29.9)	307 (29.8)
	Further	707 (21.6)	191 (17.3)	338 (56.4)	35 (6.0)	150 (14.6)
Currently Employed	Yes	2,114 (64.5)	693 (65.1)	373 (62.3)	449 (76.6)	599 (58.1)
	No	1,164 (35.5)	370 (34.8)	226 (37.7)	137 (23.4)	431 (41.8)
Socioeconomic Status	Lowest	1,052 (32.1)	130 (12.2)	144 (24.0)	422 (72.0)	356 (34.5)
	Middle	1,108 (33.8)	390 (36.7)	167 (27.9)	149 (25.4)	402 (39.0)
	Highest	1,120 (34.2)	544 (51.1)	288 (48.1)	15 (2.6)	273 (26.5)
Smoking	Never	1,857 (56.6)	414 (38.9)	382 (63.8)	444 (75.8)	617 (56.6)
	Former	1,052 (32.1)	498 (46.8)	157 (26.2)	104 (17.8)	293 (28.4)
	Current	370 (11.3)	152 (14.3)	60 (10.0)	37 (6.3)	121 (11.7)
Heavy Alcohol Consumption	No	3,106 (94.7)	1,006 (94.6)	560 (93.5)	569 (97.1)	971 (94.2)
	Yes	174 (5.3)	58 (5.5)	39 (6.5)	17 (2.9)	60 (5.8)
Physical Activity	Low	1,036 (31.6)	201 (18.9)	126 (21.0)	150 (25.6)	559 (54.2)
	Moderate	1,812 (55.2)	645 (60.6)	390 (65.1)	362 (61.8)	415 (40.3)
	High	428 (13.1)	217 (20.4)	81 (13.5)	73 (12.5)	57 (5.5)
TV Watching	<2 hours per day	1,875 (57.2)	543 (51.0)	325 (54.3)	501 (85.5)	506 (49.1)
	≥2 hours per day	1,403 (42.8)	520 (48.9)	274 (45.7)	84 (14.3)	525 (50.9)
Hypertension	No	2,433 (74.2)	772 (72.6)	443 (74.0)	502 (85.7)	716 (69.5)
	Yes	833 (25.4)	287 (27.0)	148 (24.7)	83 (14.2)	315 (30.6)
T2DM*	No	2,862 (87.3)	951 (89.4)	475 (79.3)	520 (88.7)	916 (88.9)
	Yes	272 (8.3)	86 (8.1)	50 (8.4)	21 (3.6)	115 (11.2)
High BMI*†	No	904 (27.6)	225 (21.2)	132 (22.0)	304 (51.9)	243 (23.6)
	Yes	2,208 (67.3)	783 (73.6)	411 (68.6)	253 (43.2)	761 (73.8)

T2DM, type 2 diabetes mellitus; BMI, body mass index.

*Values do not add up due to missing data.

†High BMI defined as ≥25kg/m^2 to incorporate both overweight and obesity.

Fig 1: Inclusion of participants at baseline and follow up of the CRONICAS cohort study 2010-2013



*Incidence calculations performed separately for each outcome; therefore numbers represent those excluded from calculations for the specified outcome only.

Dietary Patterns

The four-class LCA model was the preferred model to take forward in the analysis (Supplement B) based on optimal BIC, AIC and entropy values, and pattern interpretability. Dietary patterns were determined from the percentage distributions of the frequency of intake categories (none, low, moderate, high) of each food group within each class (Supplement C). Table 2 shows the summarised dietary intake patterns which can be labelled as follows based on the stage of the nutrition transition reflected by the diet: stage 1, traditional diet consisting of high starch and low fat foods with low diversity of food groups consumed; stage 2, elements of the traditional diet remain with increasing range of high fibre foods as well as high fat foods consumed; stage 3, higher in processed foods and animal products which contain high fat and sugar, with less of the traditional high fibre foods; and stage 4, with high diversity of food group consumption including high fibre, high fat and high sugar foods, indicative of a fully transitioned diet.

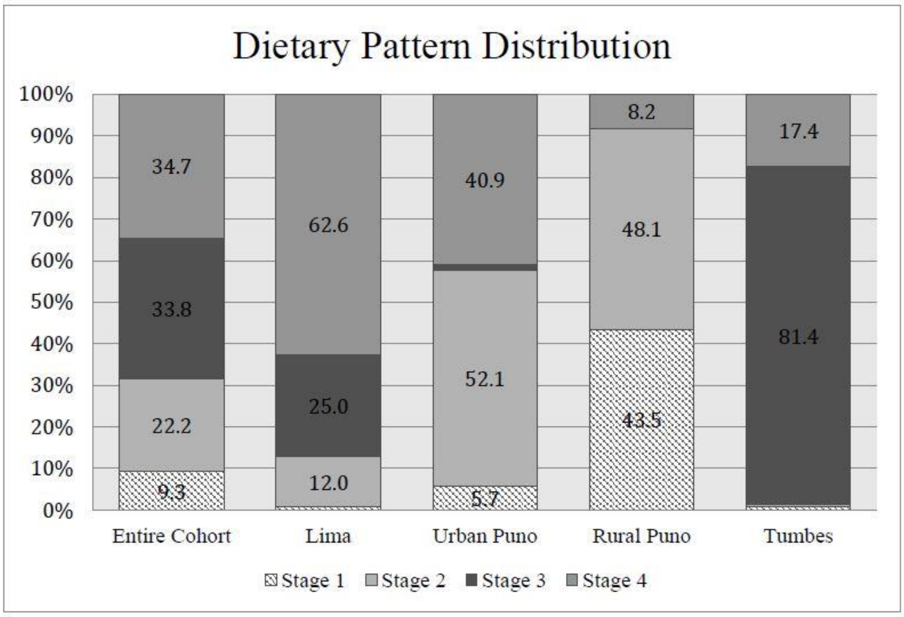
Table 2: Summarised food group dietary patterns, obtained using latent class analysis, among Peruvian adults in the CRONICAS cohort study 2010-2013. Percentages are the conditional probability of a class member falling into the stated category of intake frequency of the stated food group

Food Group	Class One Stage 1 diet	%	Class Two Stage 2 diet	%	Class Three Stage 4 diet	%	Class Four Stage 3 diet	%
Whole grains			Low whole grains	66.5				
Refined grains	Low refined grains	84.5	Low refined grains	79.4	Moderate refined grains	47.5	Moderate refined grains	66.2
Dairy	Low dairy	46.9	Low dairy	60.5	High dairy	41.8		
Red meat			Low red meat	42.1	Moderate red meat High red meat	36.2 35.4	Moderate red meat	52.2
Poultry	Low poultry No poultry	48.8 41.3	Low poultry	72.8	High poultry	44.7	Moderate poultry	53.0
Vegetables (veg)	No green veg No raw veg High cooked veg	61.8 68.1 46.9	Low green veg Low raw veg Moderate cooked veg	53.0 48.3 55.6	High green veg Moderate raw veg Moderate cooked veg	50.3 40.5 50.6	Low cooked veg	81.0
Fruit	Low fruit	55.2	Low fruit	43.4	High fruit	60.2	Moderate fruit	40.4
Seafood							High seafood	44.8
Potatoes	High potato	73.8	High potato	65.8			Moderate potato	49.1
Legumes	No legume	46.9	Low legume	61.7	Moderate legume	47.9	Moderate legume	55.3
Egg	Low egg	45.6	Low egg	63.6			Low egg	60.8
UPF	Low UPF	81.7	Low UPF	58.7	High UPF	46.7	Moderate UPF High UPF	45.4 44.4

UPF, ultra-processed foods.

In Figure 2, we show the distribution of the dietary patterns amongst the study areas. In Lima 62.6% of the study population were likely to fall within the stage 4 dietary pattern. In urban Puno, the stage 2 pattern had the highest prevalence at 52.1%. In rural Puno the majority of the study population were either the stage 2 (48.1%) or stage 1 (43.5%) pattern. In Tumbes, it was the stage 3 pattern that the majority of the population (81.4%) were likely to fall into.

Fig 2: Overall dietary pattern prevalence amongst Peruvian adults in the CRONICAS cohort study 2010-2013 and distribution according to site



Site-specific dietary pattern analysis was also performed. However, characteristics of the site-specific dietary patterns did appear to reflect the patterns obtained by dietary pattern analysis of the cohort in its entirety (Supplement D). Therefore, the overall dietary patterns were used in further analyses rather than the site-specific patterns.

Association of Dietary Patterns with Cardiometabolic Outcomes

A total of 833 (25.5%) study participants had hypertension at baseline, 272 (8.7%) had T2DM and 2,208 (71.0%) had a BMI ≥25kg/m^2. Prevalence of each outcome varied by dietary pattern (Table 3): prevalence of hypertension and T2DM was highest among participants with the stage 3 dietary pattern. High BMI was more prevalent in those of the stage 4 (76.6%), stage 3 (75.1%) and stage 2 (67.7%) dietary patterns than those with the stage 1 pattern (41.3%). Participants with the stage 1 dietary pattern had the lowest prevalence of all three outcomes. Though numbers were small, some site-specific patterns were observed within Lima, where prevalence of hypertension was highest among those with the stage 4 dietary pattern (33.6%) (Supplement E).

Table 3: Baseline prevalence of cardiometabolic outcomes by dietary pattern among Peruvian adults in the CRONICAS cohort study 2010-2013

	Hypertension *n*=3,266		T2DM *n*=3,134		High BMI* *n*=3,112	
	%	p value†	%	p value†	%	p value†
Stage 1	18.2%		3.9%		41.3%	
Stage 2	20.1%		6.9%		67.7%	
Stage 3	32.1%	<0.001	10.7%	=0.001	75.1%	<0.001
Stage 4	24.5%		8.9%		76.6%	

T2DM, type 2 diabetes mellitus; BMI, body mass index.

*High BMI defined as ≥25kg/m^2 to incorporate both overweight and obesity.

†Significance test derived from Pearson's Chi^2 test

Prevalence Rates

Crude and adjusted models of the association of dietary patterns with prevalence of cardiometabolic outcomes are shown in Table 4. The co-variates "site" and "level of urbanisation" were introduced separately into the models, however made little difference to model fit; hence the simpler model without these co-variates was selected to report results. There was no evidence of effect modification between either of these two co-variates and other potential confounders.

Overall, the stage 3 diet was associated with the highest burden of disease in the cardiometabolic outcomes and was therefore used as a reference diet to estimate prevalence ratios.

Hypertension

After adjusting for potential confounders, those with the stage 1 diet were shown to be 39% less likely to have hypertension compared to those with the stage 4 diet (PR 0.61, 95% CI 0.46-0.81); those with a stage 2 diet were shown to be 43% less likely to be hypertensive as compared to the stage 4 group (PR 0.57, 95% CI 0.48-0.68). Within Lima, a reduction in prevalence of approximately 30% was reported for those with the stage 4 diet (PR 0.70, 95% CI 0.58-0.86) and the stage 2 diet (PR 0.69, 95% CI 0.51-0.93) as compared to the stage 3 reference diet. In rural Puno the stage 1, 2 and 4 dietary patterns were associated with a much lower prevalence of hypertension (stage 1: PR 0.08, 95% CI 0.02-0.27; stage 2, PR 0.05, 95% CI 0.01-0.19; stage 4: PR

0.16, 95% CI 0.05-0.52) compared with the stage 3 dietary pattern (Supplement F). However, study power was lower when running the models by site, creating more uncertainty in the estimates.

T2DM

The prevalence of T2DM was almost 3 times higher in those with the stage 3 diet compared to those with the stage 1 diet (PR 0.29, 95% CI 0.09-0.99). There was no statistically significant difference in prevalence of T2DM between the two other diets (stage 2 and 4), however a (non-significant) trend could be observed in both overall and site specific datasets in which the lowest prevalence of T2DM was found among those with a stage 1 diet followed by stage 2, stage 4 and stage 3.

High BMI

Those with the stage 1 diet had the lowest prevalence of high BMI, which was 37% lower (PR 0.63, 95% CI 0.55-0.72) in comparison to those with the stage 4 reference diet.

Incidence Rates

A total of 2,669 (81.4%) participants completed follow up of blood pressure; 2,536 (76.9%) for blood sugar; and 2,619 (79.8%) completed follow up of BMI. Of those who returned for follow up, 669 were excluded from hypertension incidence calculations, 212 from T2DM incidence calculations, and 1,925 from incidence of high BMI calculations due to pre-existing diagnosis at baseline. Overall 203 new cases of hypertension were identified, 111 new cases of T2DM, and 139 new cases of overweight and obesity.

Incidence of hypertension was highest among individuals consuming the stage 3 dietary pattern (63.75 per 1000 person-years, 95% CI 52.40-77.55). In comparison to this diet, a reduction in hypertension incidence of 66%, 56% and 31% was found for those with the stage 1, 2 and 4 diets respectively, after adjustment for confounders (stage 1: IRR 0.34, 95% CI 0.18-0.64; stage 2: IRR 0.44, 95% CI 0.29-0.66; stage 4: IRR 0.69, 95% CI 0.51-0.94). Incidence of T2DM was highest among those consuming the stage 3 dietary pattern (24.92 per 1000 person-years, 95% CI 18.89-32.89), and incidence of high BMI was highest amongst those with the stage 2 pattern (133.77 per 1000 person-years, 95% CI 96.49-185.44). There was no evidence for a difference in the risk of T2DM or high BMI amongst the dietary patterns.

Table 4: Association between dietary pattern and hypertension, T2DM and high BMI* at baseline and follow up among Peruvian adults in the CRONICAS cohort study 2010-2013. Average follow up period 30 months.

		Incidence per 1000 person-years (95% CI)	Crude		Fully Adjusted†	
			PR (95% CI)	IRR (95% CI)	PR (95% CI)	IRR (95% CI)
Hypertension		46.07 (40.15-52.86)				
	Stage 3	63.75 (52.40-77.55)	1 (Reference)	1 (Reference)	1 (Reference)	1 (Reference)
	Stage 1	28.88 (15.03-55.51)	**0.57 (0.43-0.75)**	**0.36 (0.19-0.71)**	**0.61 (0.46-0.81)**	**0.34 (0.18-0.64)**
	Stage 2	31.77 (21.63-46.65)	**0.63 (0.53-0.75)**	**0.44 (0.29-0.67)**	**0.57 (0.48-0.68)**	**0.44 (0.29-0.66)**
	Stage 4	39.39 (31.39-50.50)	**0.77 (0.67-0.87)**	**0.63 (0.47-0.84)**	**0.76 (0.67-0.86)**	**0.69 (0.51-0.94)**
T2DM		21.34 (17.72-25.70)				
	Stage 3	18.91 (13.76-25.99)	1 (Reference)	1 (Reference)	1 (Reference)	1 (Reference)
	Stage 1	13.36 (5.02-35.61)	**0.36 (0.16-0.86)**	0.55 (0.20-1.51)	**0.29 (0.09-0.99)**	0.64 (0.22-1.87)
	Stage 2	21.42 (13.67-33.59)	**0.64 (0.45-0.91)**	1.00 (0.59-1.72)	0.73 (0.53-1.01)	1.00 (0.55-1.80)
	Stage 4	24.92 (18.89-32.89)	0.84 (0.66-1.07)	1.31 (0.87-1.98)	0.86 (0.67-1.09)	1.32 (0.84-2.05)
High BMI		95.21 (80.63-112.43)				
	Stage 3	89.57 (67235-119.21)	1 (Reference)	1 (Reference)	1 (Reference)‡	1 (Reference) ‡
	Stage 1	80.12 (49.81-128.89)	**0.55 (0.48-0.63)**	0.71 (0.43-1.19)	**0.63 (0.55-0.72)**	0.78 (0.45-1.37)
	Stage 2	133.77 (96.49-185.44)	**0.90 (0.85-0.96)**	1.25 (0.85-1.83)	**0.93 (0.88-0.99)**	1.32 (0.87-1.99)
	Stage 4	85.93 (62.78-117.61)	1.02 (0.97-1.07)	0.92 (0.63-1.35)	0.97 (0.92-1.02)	0.95 (0.64-1.42)

T2DM, type 2 diabetes mellitus; BMI, body mass index; PR, prevalence ratio; IRR, incidence risk ratio. Results with *p* value <0.05 shown in bold.

*High BMI defined as ≥25kg/m^2 to incorporate both overweight and obesity.

† Adjusted for age, sex, education level, currently working, socioeconomic status, smoking, heavy drinking, physical activity, TV watching and high BMI.

‡ Adjusted for age, sex, education level, currently working, socioeconomic status, smoking, heavy drinking, physical activity, and TV watching.

Discussion

This present study demonstrates an association between food group consumption patterns and hypertension, T2DM and high BMI in a cohort of Peruvian adults, and that the distribution of these patterns differ between the diverse geographic areas of Peru. An interesting trend in prevalence of cardiometabolic outcomes amongst the different diets was seen, with prevalence lowest in the more traditional stage 1 and 2 diets, and higher in the stage 3 and 4 dietary patterns, which contain more processed foods and animal products. Thus, our results support the use of dietary pattern analysis in furthering understanding of the impact of dietary risk factors in NCD development.

Although no dietary diversity scoring was undertaken within this study, and therefore no formal conclusions on the impact of dietary diversity can be drawn, the dietary patterns most prevalent in highly urbanised Lima and semi-urban Tumbes incorporated higher consumption of a greater number of food groups in comparison to highland Puno, inclusive of ultra-processed foods (UPFs,) animal products, refined grains, seafood and fruit. This is in keeping with the benefits of greater availability and affordability of such products that result from urbanisation. Dietary diversity is encouraged in public health messages to ensure requirements of essential nutrients are met(30), however health outcomes may differ depending on the nature of the foods comprising the different interpretations of diversity(31). For example, patterns of consumption typical of a Western diet high in a range of animal products, refined carbohydrates and processed foods are linked to higher rates of NCDs and NCD-related mortality, whereas diets high in diverse fruits, vegetables and whole grains and low in processed foods are associated with better outcomes(32-35). Consistent with this, results presented here show that the diets with comparatively higher intake frequency of UPFs, refined grains and animal products were associated with greater prevalence of hypertension, T2DM and high BMI, and incidence of hypertension.

Rates of sale of UPF products in Peru were amongst the fastest growing in Latin America from 2000-2013, and fast-food purchases (typically high in refined carbohydrates and processed meat) doubled during this time making Peru one of the biggest consumers of fast-foods in the region(36). Consumption of UFPs is associated with an increase in BMI(37), higher prevalence of obesity(38), and incidence of hypertension(39). Sugar-sweetened beverages, a form of UPF, are the preferred choice of drink by an alarming amount throughout the country(40), and their consumption is also associated with obesity and T2DM(41, 42). In Peru, sugar-sweetened beverage consumption is much higher in urban and coastal areas than in rural or highland parts of the country(43), giving further support to the finding of this present study that the dietary patterns more prevalent in the urban and semi-urban coastal settings had higher risk of cardiometabolic outcomes.

Interestingly, of the dietary patterns more in line with the typical Western diet, the stage 3 pattern most consumed in Tumbes was associated with a higher prevalence of the cardiometabolic outcomes than the stage 4 pattern more prevalent in Lima. This is in keeping with previous studies demonstrating that the semi-urban coastal area of Tumbes had a worse cardiometabolic risk profile than the urban capital Lima(10, 11); but is somewhat surprising given the extensive research linking degree of urbanisation to obesity and NCD risk, especially in LMIC settings(44, 45). As the predominant dietary pattern of Lima had a higher comparative frequency of fruit and vegetable intake, it may be that a more diverse diet inclusive of these foods confers some protection. This can be supported by many previous studies using dietary pattern analysis to examine the relationship between diet and NCD risk(24, 32, 35, 46).

The predominant dietary pattern of Tumbes was the only pattern with high frequency of fish and seafood consumption, consistent with Tumbes being a coastal area known for its seafood dishes. Fish, in particular oily fish, are an important source of omega 3 fatty acids which are known to have cardiovascular and inti-inflammatory benefits, and accordingly fish consumption has been associated with lower cardiovascular risk profiles when included as part of the "healthier" diet in dietary pattern studies(32, 35). However this effect may depend on the type of fish and how it's cooked, as well as other components of the diet, such as fruit, vegetables and wholegrains. In Tumbes the population may eat healthy fish however the diet does not appear to be complimented by a high vegetable and wholegrain intake, and the potential benefits may be limited by consumption of UPFs, as increased consumption of UPFs has previously been associated with a greater risk of hypertension(39). As the dietary patterns in the current study are model driven based on consumption frequency of multiple food groups, it is difficult to distinguish the negative effect of the high salt and saturated fat content of the UPFs from the potentially positive effect of fish, or absent effect of other components.

The dietary pattern with the most favourable cardiometabolic risk profile, the stage 1 diet, was most widely consumed in rural Puno. Incidence of hypertension and T2DM are lowest at this highland site in comparison to Lima and Tumbes(10, 12), which would be consistent with a diet lower in UPFs and higher in vegetables. Low calorie intake suggested by the low dietary diversity may also help to explain the lower incidence of T2DM and hypertension, for which overweight and obesity are risk factors(47). A previous study in Puno reported a mean calorie intake of 1300 kcal/day(18), in keeping with the lower prevalence of overweight and obesity associated with the dietary patterns commonly consumed in this high altitude area. Whilst a low cardiometabolic risk profile is encouraging, the low mean calorie intake is likely to indicate a high level of undernutrition and micronutrient deficiency in the area, especially if there is a low diversity of food groups consumed as our findings suggest. In the 2012 Peruvian nutrition transition mapping, Puno was one of the many areas undergoing rapid nutrition transition and experiencing the double-burden of nutrition related disease[4]. Programs and policies in Peru therefore face the challenge of addressing micronutrient deficiency by increasing availability and affordability of a diverse range of foods associated with reduced health risk, such as fruits, vegetables and whole grains[32, 35, 46], whilst being careful not to contribute to the increasing burden of obesity and NCDs by promotion of high sugar and high fat foods.

The exact length of time for dietary changes to be reflected in clinical outcomes is unknown, therefore longitudinal analysis was performed in order to explore the role of diet as a risk factor for cardiometabolic outcomes over time. However results at follow up were less robust than for cross-sectional analysis due to a relatively short follow-up period and smaller sample size. Nonetheless the stage 3 diet mostly consumed in semi-urban Tumbes was associated with the highest incidence of hypertension, consistent with Tumbes having the highest incidence of hypertension amongst the study sites in a previous study[10]. This suggests an interplay between diet, location and chronic disease, further exploration of which was limited in this current study by paucity of data when individually analysing the association of dietary patterns with the outcomes in each site, but could be the focus of future studies.

This is the first study to use latent class analysis for the investigation of dietary patterns in Peru, and substantially adds to previous studies by examining the prospective relationship of diet with disease burden in regions of the country. In support of studies done elsewhere, this study has found an association between dietary pattern and prevalence of hypertension, T2DM[48] and obesity[24, 46, 49, 50]. Use of this method in the investigation of diet as a risk factor for disease has allowed for examination of the effect of diet as a whole, rather than focussing on specific nutrients and food groups taken out of the context of how food is consumed.

However, there are a number of limitations to consider when interpreting the findings of this study. First, the results may have been subject to selection bias due to loss to follow up and missing data. Although sample size was large, final numbers in the analysis of the follow up data were small due to loss to follow up, missing data, and exclusions of those who already had the outcomes of interest (see Figure 1). Second, assessment of diet relied on self-reporting in response to a short FFQ, which can lead to non-differential measurement bias that may weaken the strength of the associations, therefore the estimates presented here may be conservative. Third, energy intake was not adjusted for, nor was portion size taken into account in this study, which can make the results of dietary pattern analysis easier to interpret[27]. Latent class analysis was therefore based on frequency of daily consumption rather than quantity in this case. While not able to give an accurate quantitative description of the diet, it did give a qualitative indication of the composition of the diet. Fourth, ethnicity was not included in the final adjusted model due to collinearity with other risk factors. This may have resulted in a portion of the genetic component of disease risk being unaccounted for, leading to an over-estimation of the association between diet and the cardiometabolic outcomes. However the role of genetics in this particular study may be relatively small, as participants were mainly of native or mixed native background and therefore likely to have shared ancestry[(51)]. Lastly, this cohort had a short follow up period, though because the exact length of time for development of chronic diseases to occur in response to diet is unknown, our results still provide interesting and relevant findings.

In conclusion, this study revealed clear dietary patterns indicative of consumption habits and diet composition in four different settings in Peru. The distribution of dietary patterns was found to closely reflect the distribution of disease risk profile amongst the settings, demonstrating that diet may explain some of the variation in disease prevalence amongst the different areas in addition to urbanisation and socioeconomic status, acknowledging that these are all closely linked. Characterising local diets can contribute toward development of locally-relevant guidelines for health promotion and disease prevention, for example by discouraging the consumption of commonly eaten foods associated with adverse health outcomes (for instance through tax strategies), and informing policies to improve access to a diverse range of food groups associated with lower disease risk.

References

1. World Health Organization. A prioritized research agenda for prevention and control of noncommunicable diseases. Geneva, Switzerland: WHO, 2011.
2. Bermudez OI, Tucker KL. Trends in dietary patterns of Latin American populations. Cadernos de Saúde Pública. 2003;19:S87-99.
3. Popkin BM. The nutrition transition and its health implications in lower income countries. Public Health Nutrition. 1998;1(1):5-21.
4. Chaparro MP, Estrada L. Mapping the nutrition transition in Peru: evidence for decentralized nutrition policies. Revista Panamericana de Salud Pública. 2012;32:241-4.
5. Berti PR, Fallu C, Cruz Agudo Y. A systematic review of the nutritional adequacy of the diet in the Central Andes. Revista Panamericana de Salud Pública. 2014;36:314-23.
6. Oyarzun PJ, Borja RM, Sherwood S, et al. Making sense of agrobiodiversity, diet, and intensification of smallholder family farming in the highland Andes of Ecuador. Ecology of food and nutrition. 2013;52(6):515-41.
7. World Health Organization. Noncommunicable diseases (NCD) country profiles , 2014 [20th July 2017]. Available from: http://www.fao.org/faostat/en/#country/170.
8. Food and Agriculture Organization. FAOSTAT: Peru: Country Indicators 2017 [4th September 2017]. Available from: http://www.fao.org/faostat/en/#country/170.
9. World Health Organization. 2008-2013 Action Plan for the Gobal Strategy for the Prevention and Control of Noncommunicable Diseases. Geneva, Switzerland: WHO, 2008.
10. Bernabé-Ortiz A, Carrillo-Larco RM, Gilman RH, et al. Impact of urbanisation and altitude on the incidence of, and risk factors for, hypertension. Heart. 2017;103(11):827-33.
11. Bernabe-Ortiz A, Carrillo-Larco RM, Gilman RH, et al. Contribution of modifiable risk factors for hypertension and type-2 diabetes in Peruvian resource-limited settings. Journal of Epidemiology and Community Health. 2016;70(1):49-55.
12. Bernabe-Ortiz A, Carrillo-Larco RM, Gilman RH, et al. Geographical variation in the progression of type 2 diabetes in Peru: The CRONICAS Cohort Study. Diabetes Research and Clinical Practice. 2016;121:135-45.
13. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Current Opinion in Lipidology. 2002;13(1):3-9.
14. Fung TT, Willett WC, Stampfer MJ, et al. Dietary patterns and the risk of coronary heart disease in women. Archives of Internal Medicine. 2001;161(15):1857-62.
15. Miranda JJ, Bernabe-Ortiz A, Smeeth L, et al. Addressing geographical variation in the progression of non-communicable diseases in Peru: the CRONICAS cohort study protocol. BMJ Open. 2012;2(1):e000610.
16. World Health Organization. WHO STEPwise approach to surveillance (STEPS) manual Geneva, Switzerland [24th August 2017]. Available from: http://www.who.int/chp/steps/manual/en/.
17. O'Donnell M, Xavier D, Diener C, et al. Rationale and design of INTERSTROKE: a global case-control study of risk factors for stroke. Neuroepidemiology. 2010;35(1):36-44.
18. McCloskey ML, Tarazona-Meza CE, Jones-Smith JC, et al. Disparities in dietary intake and physical activity patterns across the urbanization divide in the Peruvian Andes. The International Journal of Behavioral Nutrition and Physical Activity. 2017;14:90.
19. Coleman A, Steel S, Freeman P, et al. Validation of the Omron M7 (HEM-780-E) oscillometric blood pressure monitoring device according to the British Hypertension Society protocol. Blood Pressure Monitoring. 2008;13(1):49-54.
20. James PA, Oparil S, Carter BL, et al. 2014 evidence-based guideline for the management of high blood pressure in adults: Report from the panel members appointed to the eighth joint national committee (jnc 8). JAMA. 2014;311(5):507-20.
21. World Health Organization. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia: report of a WHO/IDF consultation. Geneva, Switzerland: WHO, 2006.
22. World Health Organization. Obesity and overweight factsheet. Geneva, Switzerland: WHO, 2016.
23. Leech RM, Worsley A, Timperio A, et al. Temporal eating patterns: a latent class analysis approach. The International Journal of Behavioral Nutrition and Physical Activity. 2017;14:3.

24. Joy EJ, Green R, Agrawal S, et al. Dietary patterns and non-communicable disease risk in Indian adults: secondary analysis of Indian Migration Study data. Public Health Nutrition. 2017;20(11):1963-72.
25. Noor SWB, Ross MW, Lai D, et al. Use of latent class analysis approach to describe drug and sexual HIV risk patterns among injection drug users in Houston, Texas. AIDS and Behavior. 2014;18(3):276-83.
26. Padmadas SS, Dias JG, Willekens FJ. Disentangling women's responses on complex dietary intake patterns from an Indian cross-sectional survey: a latent class analysis. Public Health Nutrition. 2007;9(2):204-11.
27. Fahey MT, Thane CW, Bramwell GD, et al. Conditional Gaussian mixture modelling for dietary pattern analysis. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2007;170(1):149-66.
28. Tein J, Coxe S, Cham H. Statistical power to detect the correct number of classes in latent profile analysis. Structural Equation Modeling : a multidisciplinary journal. 2013;20(4):640-57.
29. Berlin KS, Williams NA, Parra GR. An introduction to latent variable mixture modeling (part 1): overview and cross-sectional latent class and latent profile analyses. Journal of Pediatric Psychology. 2014;39(2):174-87.
30. Steyn N, Nel J, Nantel G, et al. Food variety and dietary diversity scores in children: are they good indicators of dietary adequacy? Public Health Nutrition. 2006;9(5):644-50.
31. de Oliveira Otto MC, Padhye NS, Bertoni AG, et al. Everything in moderation-dietary diversity and quality, central obesity and risk of diabetes. PLoS One. 2015;10(10):e0141341.
32. Heidemann C, Schulze MB, Franco OH, et al. Dietary patterns and risk of mortality from cardiovascular disease, cancer, and all causes in a prospective cohort of women. Circulation. 2008;118(3):230-7.
33. Mente A, de Koning L, Shannon HS, et al. A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. Archives of Internal Medicine. 2009;169(7):659-69.
34. Mozaffarian D, Hao T, Rimm EB, et al. Changes in diet and lifestyle and long-term weight gain in women and men. The New England Journal of Medicine. 2011;364(25):2392-404.
35. van Dam RM, Rimm EB, Willett WC, et al. Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. Annals of Internal Medicine. 2002;136(3):201-9.
36. Pan-American Health Organization. Ultra-processed food and drink products in Latin America: Trends, impact on obesity, policy implications. Washington D.C.: PAHO, 2015.
37. Asfaw A. Does consumption of processed foods explain disparities in the body weight of individuals? The case of Guatemala. Health Economics. 2011;20(2):184-95.
38. Canella DS, Levy RB, Martins APB, et al. Ultra-processed food products and obesity in Brazilian households (2008–2009). PLoS One. 2014;9(3):e92752.
39. Mendonca RD, Lopes AC, Pimenta AM, et al. Ultra-processed food consumption and the incidence of hypertension in a Mediterranean cohort: the Seguimiento Universidad de Navarra project. American Journal of Hypertension. 2017;30(4):358-66.
40. Instituto Nacional de Estadística e Informática. Consumo de Alimentos y Bebidas 2008-2009. Lima, Peru: INEI, 2012.
41. Hu FB, Malik VS. Sugar-sweetened beverages and risk of obesity and type 2 diabetes: Epidemiologic evidence. Physiology & Behavior. 2010;100(1):47-54.
42. Schulze MB, Manson JE, Ludwig DS, et al. Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. JAMA. 2004;292(8):927-34.
43. Lazaro M. Guias alimentarias para la poblacion Peruana, 2013 [4th September 2017]. Available from: http://www.ins.gob.pe/repositorioaps/0/0/not/temdif32599/PPT%20Gu%C3%ADas%20alimentarias.pdf.
44. Popkin BM, Adair LS, Ng SW. Now and then: The global nutrition transition: the pandemic of obesity in developing countries. Nutrition Reviews. 2012;70(1):3-21.
45. Albala C, Vio F, Kain J, et al. The nutrition transition in Latin America: the case of Chile. Nutrition Reviews. 2001;59(6):170-6.
46. Ganguli D, Das N, Saha I, et al. Major dietary patterns and their associations with cardiovascular risk factors among women in West Bengal, India. British Journal of Nutrition. 2011;105(10):1520-9.

47. Guh DP, Zhang W, Bansback N, et al. The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis. BMC Public Health. 2009;9(1):88.
48. Daniel CR, Prabhakaran D, Kapur K, et al. A cross-sectional investigation of regional patterns of diet and cardio-metabolic risk in India. Nutrition Journal. 2011;10:12.
49. Satija A, Hu FB, Bowen L, et al. Dietary patterns in India and their association with obesity and central obesity. Public Health Nutrition. 2015;18(16):3031-41.
50. Pou SA, del Pilar Díaz M, De La Quintana AG, et al. Identification of dietary patterns in urban population of Argentina: study on diet-obesity relation in population-based prevalence study. Nutrition Research and Practice. 2016;10(6):616-22.
51. Mao X, Bigham AW, Mei R, et al. A genomewide admixture mapping panel for Hispanic/Latino populations. The American Journal of Human Genetics.80(6):1171-8.

Supplement A: Food groups contained within the short version food frequency questionnaire and their corresponding dietary pattern analysis food category used to determine diets among Peruvian adults in the CRONICAS cohort study 2010-2013

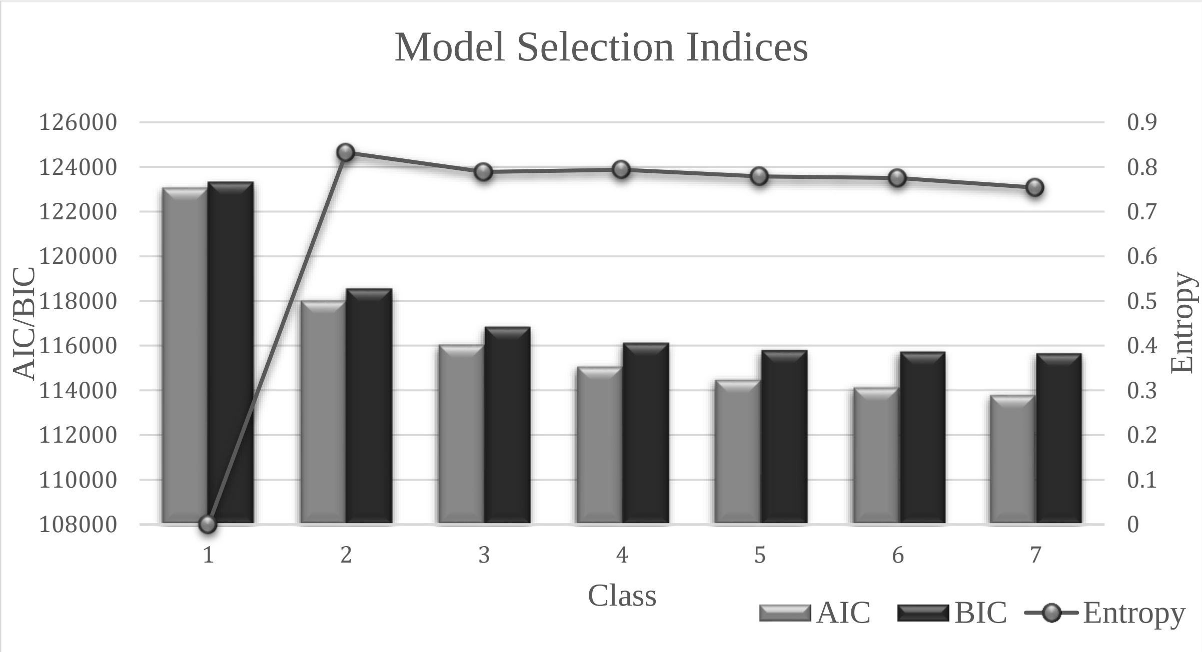
Food groups within the questionnaire	**Food groups used for dietary pattern analysis**
Whole grains	Whole grains
Refined/milled grains	Refined grains
Dairy products	Dairy
Meat	Red meat
Organ meats	Red meat
Poultry	Poultry
Fish and Seafood	Seafood
Eggs	Eggs
Pizza	UPFs
Leafy green vegetables	Leafy green vegetables
Other raw vegetables	Raw vegetables
Other cooked vegetables	Cooked vegetables
Legumes, nuts and seeds	Legumes
Pickled food	Not used*
Deep fried foods	UPFs
Potatoes	Potatoes
Salty snacks	UPFs
Fruits	Fruit
Ice cream and pudding	UPFs
Desserts/sweet snacks	UPFs
Confectionary, sugars and syrups	UPFs
Fruit drinks	UPFs
Carbonated beverages†	UPFs

UPF, ultra-processed foods.

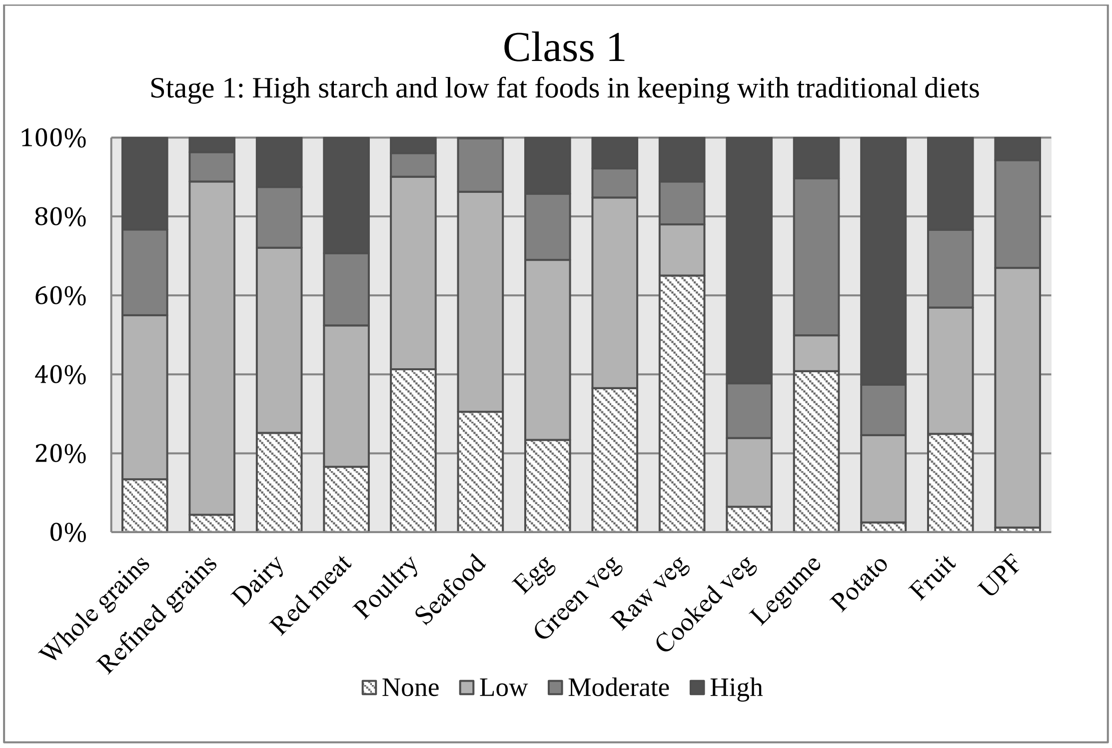
*Pickled foods not retained in analysis as not eaten by 89.5% of study population.

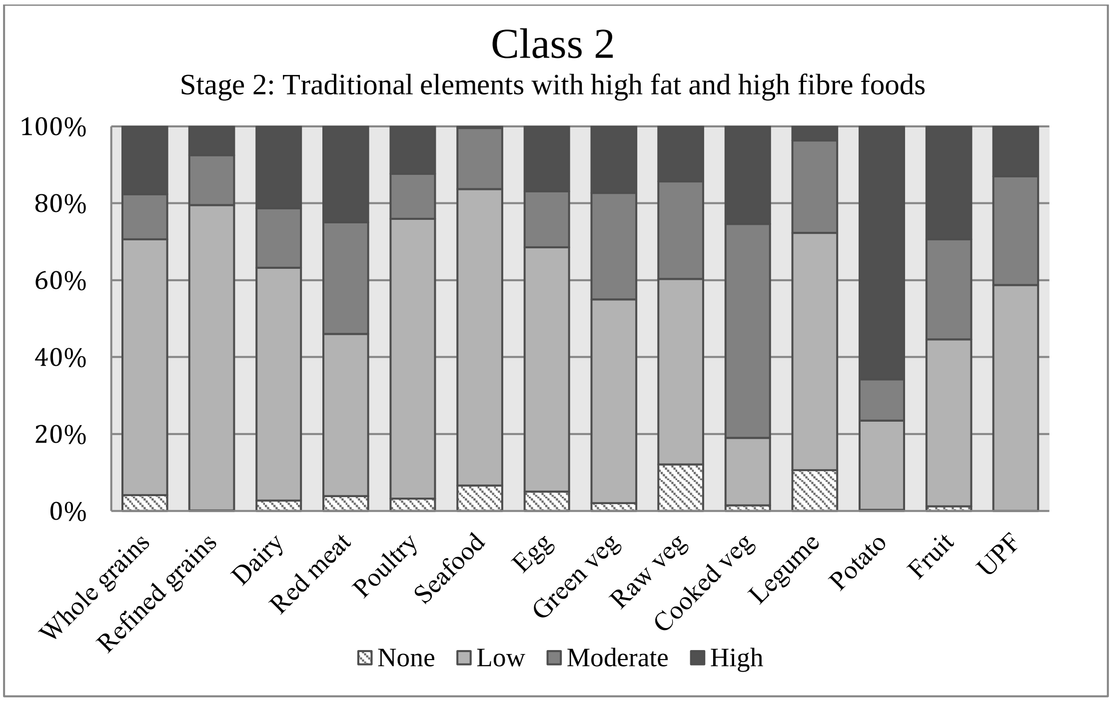
†Includes sugar-sweetened beverages.

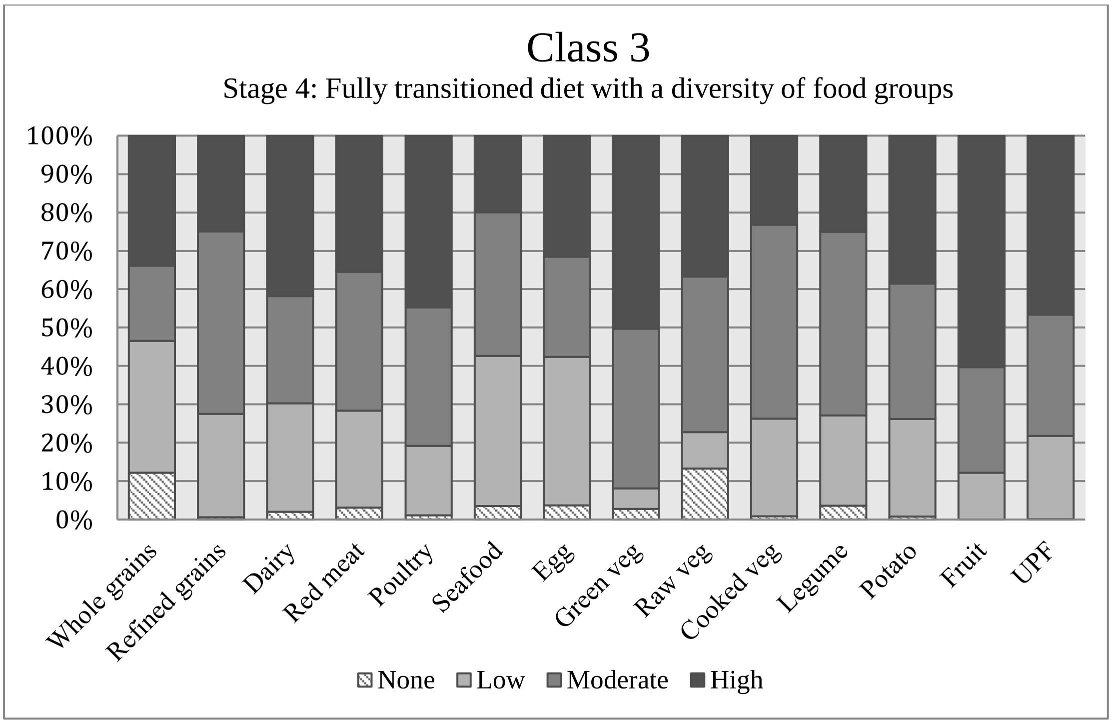
Supplement B: Goodness of fit criteria for selection of the optimal latent class analysis model in the identification of dietary patterns among Peruvian adults in the CRONICAS cohort study 2010-2013

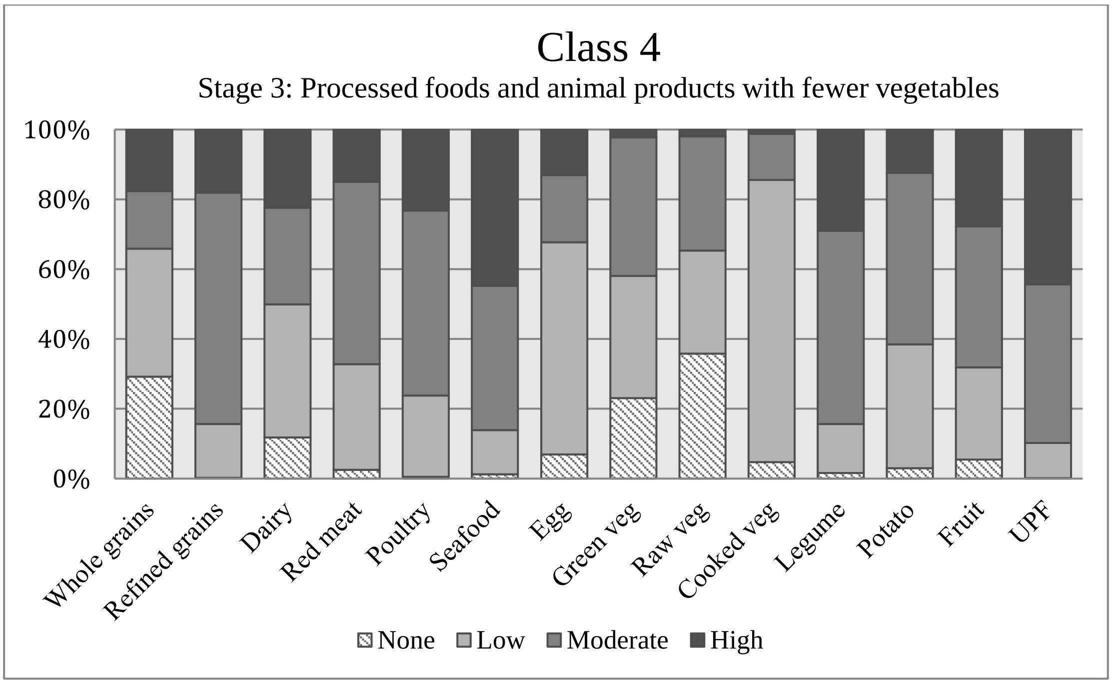


Supplement C: Percentage distribution of frequency of intake categories of food groups within each class. Used to determine dietary patterns amongst Peruvian adults in the CRONICAS cohort study 2010-2013






Class 3
Stage 4: Fully transitioned diet with a diversity of food groups
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Whole grains
Refined grains
Dairy
Red meat
Poultry
Seafood
Egg
Green veg
Raw veg
Cooked veg
Legume
Potato
Fruit
UPF
None
Low
Moderate
High


Class 4
Stage 3: Processed foods and animal products with fewer vegetables
100%
80%
60%
40%
20%
0%
Whole grains
Refined grains
Dairy
Red meat
Poultry
Seafood
Egg
Green veg
Raw veg
Cooked veg
Legume
Potato
Fruit
UPF
None
Low
Moderate
High

Supplement D: Distribution of site-specific dietary patterns from four study sites amongst overall dietary patterns of entire cohort of Peruvian adults in the CRONICAS cohort study 2010-2013. Site specific dietary patterns labelled according to most frequently eaten food groups within each pattern.

Dietary Patterns of the Cohort	**Dietary Patterns of Lima (*n=1,064*)**				
	Processed foods & animal products No. (%)	*Fruit & poultry No. (%)*	*High diversity No. (%)*	*Fruit, vegetables & animal products No. (%)*	*Total No. (%)*
Stage 1	6 (2.0)	1 (0.6)	0	1 (0.4)	9 (0.8)
Stage 2	34 (11.2)	88 (50.6)	0	6 (2.5)	128 (12.0)
Stage 3	121 (39.8)	66 (37.9)	74 (21.45)	1 (0.4)	262 (24.6)
Stage 4	143 (47.0)	19 (10.9)	271 (78.6)	233 (96.7)	666 (62.6)

Dietary Patterns of the Cohort	**Dietary Patterns of Urban Puno (*n=599*)**				
	Potato & low diversity No. (%)	*Potato, fruit, diary & low diversity No. (%)*	*Potato, fruit, vegetables & low diversity No. (%)*	*Dairy, processed foods, fruit & high diversity No. (%)*	*Total No. (%)*
Stage 1	24 (48.0)	5 (3.3)	0	5 (1.9)	34 (5.7)
Stage 2	21 (42.0)	138 (92.0)	119 (93.0)	34 (12.6)	312 (52.1)
Stage 3	0	1 (0.7)	0	7 (2.6)	8 (1.3)
Stage 4	5 (10.0)	6 (4.0)	9 (7.0)	225 (83.0)	245 (40.9)

Dietary Patterns of the Cohort	**Dietary Patterns of Rural Puno (*n=586*)**				
	Potato, vegetables & animal products No. (%)	*Wholegrains, potato, fruit, vegetables & red meat No. (%)*	*Potato & low diversity No. (%)*	*Potato & cooked vegetables No. (%)*	*Total No. (%)*
Stage 1	21 (44.7)	38 (25.3)	10 (5.2)	186 (94.9)	255 (43.5)
Stage 2	7 (14.9)	82 (54.7)	183 (94.8)	10 (5.1)	282 (48.1)
Stage 3	1 (20.3)	0	0	0	1 (0.2)
Stage 4	18 (38.3)	30 (20.0)	0	0	48 (8.2)

Dietary Patterns of the Cohort	**Dietary Patterns of Tumbes (*n=1,031*)**				
	Seafood, processed foods & high diversity No. (%)	*Seafood, processed foods & legumes No. (%)*	*Seafood, potatoes, fruit & animal products No. (%)*	*Seafood & processed foods No. (%)*	*Total No. (%)*
Stage 1	0	0	0	7 (1.54)	7 (0.7)
Stage 2	3 (1.1)	0	1 (0.8)	2 (0.4)	6 (0.6)
Stage 3	164 (58.6)	176 (100.0)	59 (49.2)	440 (96.7)	839 (81.4)
Stage 4	113 (40.4)	0	60 (50.0)	6 (1.32)	179 (17.4)

Supplement E: Baseline prevalence of cardiometabolic outcomes by dietary pattern among Peruvian adults in the CRONICAS cohort study 2010-2013 separated by study site

		Hypertension		*T2DM*		*High BMI**	
		%	p value†	%	p value†	%	p value†
LIMA		*n*=1,059		*n*=1,037		*n*=1,008	
	Stage 1	*No observations*		*No observations*		87.5%	
	Stage 2	27.3%		6.4%		81.1%	
	Stage 3	33.6%	0.017	8.2%	0.678	68.2%	0.356
	Stage 4	24.8%		8.7%		78.2%	
URBAN PUNO		*n*=591		*n*=525		*n*=543	
	Stage 1	24.2%		3.7%		66.7%	
	Stage 2	25.6%	0.830	9.0%	0.316	75.1%	0.643
	Stage 3	37.5%		25.0%		75.0%	
	Stage 4	23.9%		10.3%		77.5%	
RURAL PUNO		*n*=585		*n*=541		*n*=557	
	Stage 1	16.9%		3.3%		36.8%	
	Stage 2	10.7%	0.011	4.7%	0.805	53.7%	0.001
	Stage 3	100% (*n*=1)		*No observations*		100% (*n*=1)	
	Stage 4	18.8%		2.3%		40.0%	
TUMBES		*n*=1,031		*n*=1,031		*n*=1,004	
	Stage 1	57.1%		28.6%		42.9%	
	Stage 2	16.7%	0.158	16.7%	0.416	83.3%	0.145
	Stage 3	31.5%		11.3%		75.3%	
	Stage 4	25.7%		9.5%		79.0%	

T2DM, type 2 diabetes mellitus; BMI, body mass index.

*High BMI defined as ≥25kg/m^2 to incorporate both overweight and obesity.

†Significance test derived from Pearson's Chi^2 test.

Supplement F: Association between dietary pattern and hypertension, T2DM and high BMI* at baseline and follow up among Peruvian adults in the CRONICAS cohort study 2010-2013, separated by site

		PR (95% CI)	
		Crude	Fully Adjusted†
LIMA	*Hypertension* (*n*=1,059)		
	Stage 3	1 (Reference)	1 (Reference)
	Stage 1	*No observations*	*No observations*
	Stage 2	0.81 (0.59-1.13)	**0.69 (0.51-0.93)**
	Stage 4	**0.74 (0.60-0.91)**	**0.70 (0.58-0.86)**
	T2DM (*n*=1,037)		
	Stage 3	1 (Reference)	1 (Reference)
	Stage 1	*No observations*	*No observations*
	Stage 2	0.78 (0.33-1.84)	0.88 (0.14-5.61)
	Stage 4	1.07 (0.66-1.72)	4.01 (0.63-25.4)
	High BMI (*n*=1,008)		
	Stage 3	1 (Reference)	1 (Reference)‡
	Stage 1	1.18 (0.84-1.66)	1.21 (0.87-1.69)
	Stage 2	1.10 (0.98-1.23)	1.09 (0.97-1.22)
	Stage 4	1.06 (0.97-1.15)	1.02 (0.94-1.11)
URBAN PUNO	*Hypertension* (*n*=591)		
	Stage 3	1 (Reference)	1 (Reference)
	Stage 1	0.65 (0.24-1.77)	0.30 (0.10-0.89)
	Stage 2	0.68 (0.30-1.56)	0.41 (0.19-0.89)
	Stage 4	0.64 (0.28-1.47)	0.41 (0.18-0.91)
	T2DM (*n*=525)		
	Stage 3	1 (Reference)	*No observations*
	Stage 1	0.15 (0.01-3.29)	*Too few observations*

		Stage 2	0.36 (0.15-0.89)	*Too few observations*
		Stage 4	0.41 (0.17-1.01)	*Too few observations*
	High BMI (*n*=543)			
		Stage 3	1 (Reference)	1 (Reference) ‡
		Stage 1	0.89 (0.56-1.42)	0.93 (0.60-1.45)
		Stage 2	1.0 0.67-1.50)	1.00 (0.69-1.47)
		Stage 4	1.03 (0.69-1.55)	1.00 (0.68-1.47)
RURAL PUNO	*Hypertension* (*n*=585)			
		Stage 3	1 (Reference)	1 (Reference)
		Stage 1	**0.17 (0.08-0.35)**	**0.08 (0.02-0.27)**
		Stage 2	**0.11 (0.05-0.23)**	**0.05 (0.01-0.19)**
		Stage 4	**0.19 (0.08-0.44)**	**0.16 (0.05-0.52)**
	T2DM (*n*=541)			
		Stage 3	*No observations*	*No observations*
		Stage 1	*No observations in reference group*	*No observations in reference group*
		Stage 2	*No observations in reference group*	*No observations in reference group*
		Stage 4	*No observations in reference group*	*No observations in reference group*
	High BMI (*n*=557)			
		Stage 3	1 (Reference)	1 (Reference)‡
		Stage 1	**0.37 (0.14-0.98)**	0.40 (0.15-1.05)
		Stage 2	0.54 (0.20-1.42)	0.51 (0.19-1.34)
		Stage 4	0.40 (0.14-1.12)	0.37 (0.13-1.03)
TUMBES	*Hypertension* (*n*=1,031)			
		Stage 3	1 (Reference)	1 (Reference)
		Stage 1	1.82 (0.99-3.33)	1.93 (1.15-3.22)

	Stage 2	0.53 (0.06-4.84)	1.00 (0.35-2.85)
	Stage 4	0.82 (0.62-1.08)	0.80 (0.63-1.02)
T2DM (*n*=1,031)			
	Stage 3	1 (Reference)	1 (Reference)
	Stage 1	2.52 (1.09-5.8)	4.65 (2.40-8.98)
	Stage 2	1.47 (0.32-6.76)	0.96 (0.10-9.33)
	Stage 4	0.84 (0.50-1.41)	0.67 (0.41-1.09)
High BMI (*n*=1,004)			
	Stage 3	1 (Reference)	1 (Reference)‡
	Stage 1	0.57 (0.27-1.19)	0.68 (0.35-1.32)
	Stage 2	1.11 (0.73-1.67)	1.08 (0.72-1.63)
	Stage 4	1.05 (0.96-1.15)	1.03 (0.94-1.13)

T2DM, type 2 diabetes mellitus; BMI, body mass index; PR, prevalence ratio; Results with *p* value <0.05 shown in bold.

*High BMI defined as ≥25kg/m^2 to incorporate both overweight and obesity.

† Adjusted for age, sex, education level, currently working, socioeconomic status, smoking, heavy drinking, physical activity, TV watching and high BMI.

‡ Adjusted for age, sex, education level, currently working, socioeconomic status, smoking, heavy drinking, physical activity, and TV watching