Title and Author List:

Cohort profile: Dehghan prospective cohort study (DehPCS) on non-communicable diseases in a Kurdish community in the west of Iran

Runin Title: DehPCS on non-communicable diseases

Farhad Moradpour¹, Ebrahim Ghaderi¹, Ghobad Moradi¹, Mojdeh Zarei², Amjad Mohamadi Bolbanabad¹, Bakhtiar Piroozi¹, Azad Shokri¹

1. Social Determinants of Health Research Center, Research Institute for Health Development, Kurdistan University of Medical Sciences, Sanandaj, Iran.
2. Deputy of research and technology, Kurdistan University of Medical Sciences, Sanandaj, Iran.

Corresponding author

Azad Shokri, PhD: Social Determinants of Health Research Center, Research Institute for Health Development, Kurdistan University of Medical Sciences, Sanandaj, Iran.

Emaol: Azad_shokri@yahoo.com

Tel: +989187762166

Epub ahead of print

Title and Author List:

Cohort profile: Dehgolan prospective cohort study (DehPCS) on non-communicable diseases in a Kurdish community in the west of Iran.

Abstract:

The Dehgolan Prospective Cohort Study (DehPCS) was established to investigate and identify risk factors related to the most prevalent Non-communicable disease. In addition, to examine health status, socio-economic, behavioral, nutritional, and environmental exposures, DehPCS also attempt to collecting, analyzing and store blood, urine, nails and, hair samples to facilitate genetic studies and discovery of biomarkers and another biological determinant of NCD. In general, 3996 adult participants aged 35-70 were enrolled from the general population, between February 2018 and March 2019. 43.7% of them were female. The first follow-up wave was conducted on 3995 participants. Wide ranges of information include socioeconomic, lifestyle, nutrition, habit, physical exam, medication use and, medical history was collected. Proxy variables such as BMI index, MET score, Assets index and, macro and micronutrients were calculated. The most common self-reported diseases were kidney stones, hypertension, and fatty liver respectively. The prevalence of diabetes and hypertension among participants was 9.3% and 33.4 respectively. All data, samples and, measurements are collected again at the five-year intervals. So, we will be able to explore time-dependent changes in the NCD risk factors. DehPCS can be used to study the relationship between genetic, lifestyle, socioeconomic status, and environmental risk factors, and the most important non-communicable diseases by case-cohort and nested case-control designs that will be implemented on the cohort infrastructures. Researchers can also submit their pre-proposals via the following web address: <http://c.ddrc.ac.ir/persianaccess/Account/Login>

Keywords: Cohort study; non communicable disease; PERSIAN cohort; risk factor

Key Messages

- DehPCS is designed to examine NCDs in the Kurdish population who are located in a wide area between the northwest of the Zagros Mountains and the eastern Taurus Mountains, including Turkey, Iraq, Iran and Syria.
- 3996 adult participants aged 35-70 were enrolled from general population.
- All data, samples and measurements of the registration phase will collected again in the fifth, tenth and fifteenth years.
- After labeling, the aliquots were placed in a freezer at minus 70 °C to continue to study among this population for years.

Introduction

Epidemiological transition and demographic changes around the world have shifted attention to NCDs as a priority. Chronic diseases such as cardiovascular diseases (CVD), diabetes, cancers, and chronic lung diseases are generally responsible for higher rates of morbidity and mortality than all other causes. The incidence and death resulted from NCDs in low and middle-income countries have been increasing in recent years (1).

NCDs in Iran, as a middle-income country, are responsible for 83.48% (84.29% -82.64%) of all deaths and 78.09% (76.05% - 80.05%) of the disease burden. It has been estimated that in 2019, 326507 deaths occurred due to NCDs in Iran and during the last 20 years, deaths because of NCDs have increased by 18.40% in Iran. Also, the probability of the premature death (between 15 to 49 years) due to NCDs in Iranians has been 65.28 per 100,000 people (52.92%) (2).

Ischemic heart diseases (IHD) in Iran with a ratio of 26.28% are the leading cause of death. They are followed by cancers, stroke, and diabetes and kidney diseases with a ratio of 17.08%, 10.46% and 7.03%, respectively (3). Studies show that more than 5 million Iranians suffer from diabetes, which indicates an increase of about 113% compared to 2005. High blood pressure, high BMI, high fasting blood sugar (FBS), smoking, and high levels of LDL cholesterol as important risk factors for NCDs respectively occupy the first to fifth ranks of the disease burden with 42.12% (2).

The high number of cases and deaths due to NCDs in Iran has attracted the attention of politicians and been given priority by the government. Despite effective health programs, mobilization of facilities, inter-, and intra sectorial cooperation, a national action plan for the prevention and control of NCDs has also been developed in Iran. However, an increase in the urban population, lifestyle changes and an increase in the life expectancy in recent years as well as the prediction of rapid population aging in the coming years have more complicated the situation (4) (4, 5).

Population-based prospective cohorts are ideal studies to examine the effects of the genetic susceptibility, occupation, environment, SES, nutrition, and other aspects of lifestyles on the occurrence of NCDs. For this purpose, to investigate the environmental and ethnic effects and to identify major modifiable risk factors in different parts of Iran, the comprehensive and uniform Prospective Epidemiological Research in Iran (PERSIAN) with the support of the Ministry of Health and Medical Education (MOHME) and, the Digestive Diseases Research Institute (DDRI) launched in 2014 (6). Currently, 18 cohorts, including DehPCS, operate in different parts of the country according to the same protocol. DehPCS and Ravansar non-communicable study (RaNCD) (7) are centers that both study the Kurdish ethnicity. But the physical environment, and the environmental exposures of the inhabitants of these two regions are completely different.

The goals of DehPCS include the following: 1) to determine the incidence of major NCDs (e.g. cardiovascular, cerebrovascular, cancers, diabetes, chronic lung disease, chronic kidney disease, and liver disease) in the Kurdish population of 35 to 70 years old in Dehghan, Kurdistan province, in the west of Iran, 2) to determine the hospital admission trend and, all cause-mortality/ cause specific mortality due to NCDs, 3) to evaluate time trends of lifestyle, SES, environmental factors, nutrition, and the outcomes of interest in the Kurdish population, 4) to comparison and clarification of the relationship between risk factors (e.g. pesticide, opioid use, nutrition, dietary pattern, economic inequality) and NCDs, 5) to establishment a biobank for basic and genetic studies and discovery of biomarkers and other biological determinants of NCDs and, 6) to establishment of a population-based structure and cooperation with other PERSIAN Cohort Centers to increase the strength of the findings of the cohort data set.

DehPCS is designed to examine NCDs in the Kurdish population. Geographically, the Kurdish people are located in a wide area between the northwest of the Zagros Mountains and the eastern Taurus Mountains, including the east and southeast of Turkey, the north and northeast of Iraq, the west and northwest of Iran and the north of Syria with an approximate area of 392,000 square meters (Fig 1) (6). The Kurdish population is estimated at 41 million people, including 10 to 12 million people who inhabit in Iran (8). The cohort study of DehPCS is located in Dehghan city in the south of Kurdistan province. Kurdistan province with an area of 29,137 square kilometers has a population of about 1,600,000 people. The province capital is Sanandaj which with a population of 580,000 people is located at 45 km of the west of Dehghan. Dehghan County, the center of which is Dehghan city, has a population of about 68,000 people. The city of Dehghan has a population of about 26,000 people, almost all of whom are of Kurdish descent. There are one Hospital, 2 Comprehensive Urban Health Service Centers, 4 Comprehensive Rural Service Centers and 41 health houses in Dehghan County (9).

The city of Dehghan was considered for this study for the following reasons:

- Dehghan is located at 45 km from Kurdistan University of Medical Sciences and provides easy access for researchers on the Sanandaj-Tehran communication route.
- The majority of the Dehghan population is Kurdish and they are similar to the Kurds of other regions in terms of the culture, food and lifestyle.
- Due to low migration in Dehghan, study sensors in the follow-up phase are minimized.
- Easy access to the population and familiarity of the study staff with the population reduce costs and increase the response rate.

Figure 1: location of the Kurdistan province and Dehghan district in Iran.

Study participants

The enrolment phase started in February 2019. The pilot phase of the study lasted until May 2019, with enroll of 984 people. Participants in this study include only people who live in the urban area of Dehghan. Inclusion criteria for participants include: a) the age of 35 to 70 years, b) having Iranian citizenship (based on the national card and birth certificate), c) having a residence history of at least one year with a duration of at least 9 months per year, d) being eager to participate in the study and completing the informed consent form and having the necessary ability to communicate with the team of interviewers. Blind, deaf and dumb people as well as individuals with mental disorders (such as untreated psychosis) that unable to go through the questioning process, were excluded from the study. The age range of 35-70 years old (adults) was picked for three reasons. 1- people in this age group are more likely well-established lifestyle and behaviors, 2- in this age group, people reach the study outcome of interest within a reasonable time, and 3- they are active and energetic enough to participate in the study.

To select the sample, the study team conducted a dedicated census. Systematic random one-stage cluster sampling was used to select participants. We decided to include 50% of all 721 block to the study. So, the sampling ratio was set to 2. The first block was selected by random selection of a value equal to or less than the sampling ratio. The next blocks were selected by adding the row number of the previous block to the sampling ratio. All eligible individual was attended to participate in the study. Sampling started from the first building on the right side of the closest street to the cohort study center. A local, trained and Kurdish language speaking interviewer by going door to door of all the urban area households, conducted a census, registered the addresses of households and assigned a special code to each household. Information of household members including name, age, gender, position in the family and their contact numbers were recorded. The goals of the study and its implementation process were discussed with family members and their questions were answered. When one or more eligible family members agreed to enter the study, a pamphlet containing information about the purpose of the study, the method of its implementation, the applied tests, and the conditions for the laboratory tests (such as fasting, tests performed on blood and urine samples) was given to them. Also, by inviting participants, a certain time was set for their presence in the cohort study center. Individuals, who did not attended the study on time, were not immediately excluded from the study. They only were excluded if being absent after three calls and setting a new appointment.

A reminder phone call was made the day before their arrival. One week before the attending, they were advised not to cut their nails, not to dye their hair, and being fast for at least 8 hours.

About 9,000 people aged 35-70 live in the Dehghan area. Based on the available resources and in coordination with the central team, the members of the research team decided to include 4,000 people in the study. All eligible people were invited to enter the study. Reasons for refusing to participate for those who did not wish to enter the study, or who did not attend the Cohort Study Center after the appointment were recorded. Of all those invited, 404 (9.18%) refused to participate in the study for the following reasons: a) lack of sufficient time and busywork (41%), b) lack of confidence in the study (17%), c) rumors of AIDS transmission and the increased risk of other diseases such as Influenza and Corona (14%), d) caregiving of the child or old at home (9%), e) fear of sampling and testing (6%), f) other cases such as travel, family dissatisfaction, forgetting the attendance date (12%).

When attending the cohort center, eligibility criteria of the participants were approved by the trained interviewer. The informed consent forms of the study were completed, and then the identity details of participants and their age were registered online based on their national and identity cards. Each participant was given a unique 11-digit identification code. After registration, participants were referred to the laboratory for blood, urine, hair, and nail samples. (Figure 2 A)

Ethical Consideration

The design and implementation of the study with ethic cod of IR.MUK.REC.1396/93 were approved by the ethics committee of Kurdistan University of Medical Sciences.

Measurements

Follow up:

First follow-up wave started one year after enrolment in February 2020. Participants follow-up are done annually, both actively (by phone) and passively (self-reports, reports of cancer and disease registries, private physicians, forensic physicians, insurance institutes and laboratories). In most cases, follow-ups are done by phone and by the interviewer study physician, at one-year intervals or when the desired outcome occurs. If the person or his/her relatives do not answer the phone, after six attempts in two weeks (three different days each week), the executive team of follow up will first refer to the person's postal address and if the person is not present at the place, they refer to his/her relatives' postal addresses to implement face to face the follow-up. Also, if necessary and according to the field work schedule, in order to complete the follow-up by re-inviting the participants, a face-to-face interview will be conducted in the field (Figure 2 B).

In cases of the favorable outcome, outcome review forms (for chronic NCDs) and verbal autopsy (in case of death) are completed by the physician. Living participants are invited to visit the cohort site for further investigation. To reduce recall bias, participants are asked to provide hospitalization information and documentation related to diagnostic and treatment procedures and medications. Also, clinical, laboratory, and disease profile information for each case (the occurrence of outcomes) in all hospitals and laboratories are followed up by the inquiring physician, and a copy of it is prepared. The information recorded on these forms, along with all previous follow-up questionnaires, laboratory test results, existing biological samples, death certificate images, and other necessary medical records is reviewed for the final diagnosis by reviewer physicians in the Outcome Review Committee which consists of two physicians (internal medicine specialists) who complete the final diagnosis form based on the evidence, depending on the outcome type or the death cause. In cases where the final diagnosis of the two physicians does not match, the third reviewer (an internal medicine or other specialist) will be consulted.

To investigate changes in risk factors or protectors, all data, samples and measurements of the registration phase will be collected again in the fifth, tenth and fifteenth years. Blood samples are also collected as soon as possible for those with the desired outcome.

Participation in the study ends due to unwillingness to continue cooperation, emigration or lack of access to the participant, defects in recalling information or unrealistic reports, and death. In case of unwillingness to continue the participation, the executive team of the follow-up tries to re-invite the participant to the study by explaining the benefits of participating, including improvement of the health of the individual and community, access to mostly free medical services in the study centers for early diagnosis (and treatment) of cancers and other diseases, the benefit from recommendations for the prevention of chronic diseases, especially heart attacks and strokes. In the case of the emigrates to the outside of the study area or lack of access to the participants at the time of follow-up, the executive team of the follow-up tries to re-enter the participants to the study according to the mentioned identity information and access ways.

At present, we have completed the first wave of the follow-up phase. The response rate in the first wave of follow-up was 3995 (99.97%) (One person was unwilling to continue the study). Second follow-up wave has been started but is not completed yet.

Figure 2: A, Algorithm of the baseline measurements of DehPCS, 2018-2019. B, Algorithm for assessing the individual's vital condition and medical events in the annual follow-up of DehPCS

Instruments and measurements:

Upon arrival, all participants were enrolled in online software. Unique 11-digit codes under the name of PERSIAN Code Identification (PCID) were assigned to them by the software. Because of fasting they were immediately refer to biological sampling.

Laboratory evaluation: 25 ml of the blood sample was taken by the standard method (sitting and using a tourniquet) with vacuum tubes. 15 ml of urine was collected in sterile containers. About 300 strand of size 3 cm of hair from the back of the head and at least 10 nail samples were taken from the participants. Blood samples were divided into different aliquots using a 3,000 rpm centrifuge for 10 to 15 minutes (Table 1). After labeling (an 11-digit PCID code), the aliquots were placed in a freezer at minus 70 °C. In addition, small amounts of blood samples were used for the following tests: CBC, blood lipid profile, blood glucose profile, alkaline phosphate, creatine, BUN, gamma glutamine transpeptidase, alanine transaminase and aspartate transaminase. Urine samples were analyzed for urine pH, blood, glucose, bilirubin, nitrate, ketones, ascorbic acid, leukocytes, microalbumin and the specific gravity. The results of all analyzes were stored in the database and provided to the participant as an incentive. Storage equipment was supported 24 hours a day by automatic monitoring sensors, and by an uninterruptible power supply.

Table 1:

Anthropometry: to standardize anthropometry and reduce the systematic error, anthropometric measurements were performed after sample collection and when the participant was still fasting. Participants were asked to remove shoes, heavy clothes, and accessories. The Seka standing hand scale and the Seka inelastic tape meter with an accuracy of 0.1 cm were used for the measurements. The waist circumference, hip circumference, wrist circumference, height (in cm) and weight (kg) were measured according to the National Health Protocol (10).

Questioning: a wide range of questions were collected in the study of DehPCS. A total of 536 questions in the three areas of general, medicine, and nutrition were asked face-to-face by trained interviewers. Whenever possible, Standard validated questionnaires for Iranian population were used to collecting data (11-13).

General questions included the general and demographic characteristics (42 questions), SES (27 questions), work history (7 questions), lifestyles (36 questions), sleep and physical activity (48 questions), and personal habits (14 questions). Also, the level of participants' trust in the health care system was measured by 2 questions (Table 2). The data of this section were collected after measuring the anthropometric characteristics of the participants.

The medical status included the fertility history (33 questions), medical history (81 questions), medications (3 questions), family history of diseases (27 questions), dental and oral health (9 questions), physical examinations and blood pressure (25 questions) and general health (29 questions) (Table 2).

Participants' nutritional status was assessed using three questionnaires of food frequency (112 questions), dietary supplements (10 questions), and eating habits (31 questions). Food frequency questionnaire is a tool to check the consumption of food and beverages in the past year. The questionnaire included items suitable for the Iranian population such as bread, cereals, meats, legumes, dairy products, fruits and vegetables, sugars and sweets, and a variety of oils. Local foods and breads were also added to the questionnaire. In this section, questions such as cooking methods and food storage were also asked (Table 2).

In the annual follow-up of each participant, first the ways of accessing the person such as the address and telephone number of the person and his/her relatives are updated. All medical records over the past year including vital signs, hospital admissions, events, diagnostic and therapeutic measures, medication regimens, and the final result are registered.

Table 2:

Key findings (and publications)

3985 (99.7) out of 3996 participants in the DehPCS were Kurdish ethnicity. 1748 (43.7%) of whom were men, 1786 (44.7%) were middle-aged adults and only 486 (12.2%) had more than 60 years old (elderly). 3671 (91.9%) were married and 44(1.1%) were single (never married). 1247(31.2%) were illiterate and only 514(12.9%) had a university degree. Internet penetration rate was 1113(28.1). SES based on wealth index shows 811 (20.4%) participants were classified in poorest quintile and, 663 (16.7) of them fall in the richest level (Table 3).

Table 3:

Overall, 2991(75.1%) of all participants especially woman, had an overweight or obese BMI score, and Only 944 (23.7%) of them were normal weight. WHO classification of the waist-to-hip ratio (WHR) showed 575 (14.4%) of participants especially woman had a WHR more than 1, so at higher risk for heart disease. In total, only 547(13.8%) people had a physical activity equal to or higher than metabolic equivalent task (MET) minute per 24 hour of value 2700. Much of them were men. 1020 (25.6) of the participants consumed equal to or more than the upper tolerable level of 5 grams salt per day (Table 4).

Table 4:

Most prevalent chronic diseases were kidney stone, hypertension, fatty liver, psychiatric disorders, chronic headaches, thyroid diseases, and diabetes respectively. According to the taking medication by participants, as well as choking fasting plasma glucose and blood pressure measurements, hypertension prevalence was 1334 (33.4%; male: 32.6%, female: 34.0%) and diabetes was 373 (9.3%; male: 7.9%, female: 10.5%). Apart from kidney stones, which are a different disease, other NCDs were more common among women than men (Table 5).

Table 5:

Strengths and weaknesses

The study of DehPCS, as a part of the large Persian Cohort Study, made it possible to investigate the relationship between geographical environment and ethnicity on the occurrence of NCDs. Also, this study, along with the study of Ravansar, was the first prospective cohort study to examine the major risk factors for NCDs in the Kurdish population as an independent ethnic group. It should be noted that the rate of participation in this study was at a high level.

In this study, some baseline variables were collected on the basis of participants' reports. So, it might have limitations due to information bias, like other similar studies. Also, people who are eager to participate in long-term studies, such as cohort ones, usually differ from other people of the community in some respects. These people are usually concerned about their health, so limits the study in terms of selection bias.

Data Accessibility

Although the data of this study are not freely accessible, in order to increase the scientific output, foreign researchers can access this data for secondary analysis by submitting a request under the terms and conditions of the Persian Cohort Data Publication. All necessary information about how to access the data such as authorship rules, data transfer agreements and samples and the data dictionary is available on our website at <https://persiancohort.com>. Researchers can also submit their pre-proposals via the following web address: <http://c.ddrc.ac.ir/persianaccess/Account/Login>

Competing interest: The authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization: PM, SC, KK, VS, PJ, AA. Data curation: FM, MZ, AMB, BP, AS. Formal analysis: FM,EG, GM, AS. Writing – original draft: FM, AS, GM. Writing – review & editing: EG, BP,AMB,MZ.

Acknowledgments: Special tanks to central committee who consultant us without any expectation in return. We wish to acknowledge the help provided by the technical and support staff in the DehPCS.

Funding: The Iranian Ministry of Health and Medical Education has contributed to the funding used in the PERSIAN Cohort through Grant no 700/534.

1. World Health Organization. Global status report on noncommunicable diseases 2014. World Health Organization; 2014. Report No.: 9241564857 Contract No.: WHO/NMH/NVI/15.1.
 2. Global burden of Diseases. Washington, Institute of Health Metric and Evaluation; 2019 [2021 May 28]. Available from: <http://vizhub.healthdata.org/gbd-compare/>.
 3. Global burden of Diseases. Washington, Institute of Health Metric and Evaluation; 2019 [Available from: <http://ghdx.healthdata.org/gbd-results-tool/>].
 4. Ahmadi B, Alimohammadian M, Yaseri M, Majidi A, Boreiri M, Islami F, et al. Multimorbidity: epidemiology and risk factors in the Golestan cohort study, Iran: a cross-sectional analysis. *Medicine*. 2016;95(7).
 5. Peykari N, Hashemi H, Dinarvand R, Haji-Aghajani M, Malekzadeh R, Sadrolsadat A, et al. National action plan for non-communicable diseases prevention and control in Iran; a response to emerging epidemic. *Journal of Diabetes & Metabolic Disorders*. 2017;16(1):1-7.
 6. Poustchi H, Egtesad S, Kamangar F, Etemadi A, Keshtkar A-A, Hekmatdoost A, et al. Prospective epidemiological research studies in Iran (the PERSIAN Cohort Study): rationale, objectives, and design. *American journal of epidemiology*. 2018;187(4):647-55.
 7. Pashdar Y, Najafi F, Moradinazar M, Shakiba E, Karim H, Hamzeh B, et al. Cohort profile: Ravansar Non-Communicable Disease cohort study: the first cohort study in a Kurdish population. *International journal of epidemiology*. 2019;48(3):682-3f.
 8. Bois T, Minorsky, V. and MacKenzie, D.N., Kurds, Kurdistan", in: *Encyclopaedia of Islam*, Second Edition, Edited by: P. Bearman, Th. Bianquis, C.E. Bosworth, E. van Donzel, W.P. Heinrichs. First published online: 2012. Available from: http://dx.doi.org/10.1163/1573-3912_islam_COM_0544.
 9. The Kurdish Population by the Kurdish Institute of Paris, 2017 estimate [Available from: <https://www.institutkurde.org/en/info/the-kurdish-population-1232551004>].
 10. Center for Diseases Control. National Health and Nutrition Examination Surveys (NHANES) Anthropometry Procedure Manual 2007. Available from: https://www.cdc.gov/nchs/data/nhanes/nhanes_07_08/manual_an.pdf.
 11. Vasheghani-Farahani A, Tahmasbi M, Asheri H, Ashraf H, Nedjat S, Kordi R. The Persian, last 7-day, long form of the International Physical Activity Questionnaire: translation and validation study. *Asian journal of sports medicine*. 2011;2(2):106.
 12. Asghari G, Rezazadeh A, Hosseini-Esfahani F, Mehrabi Y, Mirmiran P, Azizi F. Reliability, comparative validity and stability of dietary patterns derived from an FFQ in the Tehran Lipid and Glucose Study. *British journal of nutrition*. 2012;108(6):1109-17.
 13. Nourbala AA, Bagheri YS, Mohammad K. The validation of general health questionnaire-28 as a psychiatric screening tool. *Hakim research journal*. 2009;11(4):47-53.
-

Table 1: Biological samples characteristics stored in DehPCS biobank

Sample	Aliquots type	Aliquots count (Sample amount in 1 ml)	Store condition (°C)
Blood	Whole blood	4	-70
	Plasma	6	-70
	Buffy coat	2	-70
	Serum	3	-70
Urine	-	1	-20
Hair	-	About 300 strand in size of 1-3 cm	In foil and zip up bag with silica gel in cool and dry storage
Nail	-	10 finger and toenails	In foil and zip up bag with silica gel in cool and dry storage

Table 2: measures and instruments used in baseline data collection and 5-yearly assessment of DehPCS

Categories	groups	Measures & Instruments
General	General & demographic characteristics	Birth date, birth place, gender, age, ethnicity, contact address, marriage status, type and number of marriage (first degree, second degree, ...), participant's status in the household, family size, number of bedrooms,
	Socio-economic characteristics	Education, employment status, assets, job history, Status of the house currently lives in, The participant's home area, Wealth score index, source of income, dwelling status, annual reading status,
	environmental exposure	drinking water status, Use of a kitchen hood or ventilator, contact with animals, fuel use for cooking and heating, access to landline, cellophane, internet, trip status, hygiene status and its facilities, housing type/structure, exposure to agriculture/household pesticides,
	Life style	Circadian rhythm of food and sleep, physical activity, sleep quality, shift working, daily activity
Medical	Personal Habit	History of smoking and passive smoking, alcohol consumption, drug use
	Reproductive history	Menstrual cycle, history of live birth, abortion, stillbirth, history of breast feeding, history of infertility, oophorectomy, tubectomy, hysterectomy, history of contraceptive usage, history of hormone replacement therapy, history of breast exam, mammography and papsmear
	General health & medical history	General health status including Physical symptoms, anxiety, social function and depression, self rate health status, family history, diabetes and renal disease, CVD, liver disease, chronic lung disease, thyroid disease, kidney disease, cancers, psychological disorders, neurological disease, other sign and symptoms, history of transfusion, hospitalization history, surgery, disability and amputation
	Medication use	type, time frame and duration of medication use
	Dental health	History of brushing and flossing, number of teeth, decayed, missing, and filling teeth (DMF), oral lesion
	Blood pressure & physical exam anthropometry	Right and left blood pressure, heart rate, Alopecia, Hirsutism, iris Color, physical or sensational disabilities, spinal disorders Height, weight, hip circumference, waist circumference, wrist circumference, BMI
Nutrition	Blood & urine analysis	Complete blood count, liver function tests, renal function tests, lipid profile, FBS,
	Food frequency	125 item Food frequency questionnaire including in the past years; bread & cereals, beans, meat and its products, dairy products, fruits & vegetables, oils & oilseeds, sugars, spices
	Supplement use	Type and amount of supplement use in the past year
	Dietary habits	Methods of cooking food, food preparation and preserving, dietary habit in the past year, any food allergy

Table 3: Demographic characteristics and social-economic status of participants in DehPCS

Variables	Total n (%)	Male n (%)	Female n (%)	P-value (Chi ²)
Age (Years)				
35-45	1786 (44.7)	733 (41.9)	1053 (46.8)	0.005
46-60	1724 (43.0)	782 (44.8)	942 (41.9)	
61-70	486 (12.7)	233 (13.3)	253 (11.3)	
Marital status				
Single	44 (1.1)	4 (0.2)	40 (1.8)	< 0.001
Married	3671 (91.9)	1729 (86.4)	1942 (98.9)	
Widowed	252 (6.3)	243 (10.8)	9 (0.5)	
Divorced	28 (0.7)	5 (0.3)	23 (1)	
Education years				
Illiterate	1247 (31.2)	185 (10.9)	1062 (47.2)	< 0.001
1-5	1112 (27.8)	458 (26.2)	654 (29.1)	
6-12	1123 (28.1)	691 (39.5)	432 (19.2)	
university	514 (12.9)	414 (23.7)	100 (4.5)	
Consanguineous marriage				
Yes	809 (20.5)	363 (20.8)	446 (20.2)	0.63
No	3143 (79.5)	1381 (79.2)	1762 (79.8)	
Use Mobile				
Yes	3632 (91.7)	1712 (99.1)	1920 (85.9)	< 0.001
No	330 (8.3)	15 (0.9)	315 (14.1)	
Use internet				
Yes	1113 (28.1)	671 (38.9)	442 (19.7)	< 0.001
No	2849 (71.9)	1056 (61.1)	1793 (80.3)	
Assets index				
Q1(Poorest)	811 (20.4)	-	-	-
Q2	824 (20.8)	-	-	
Q3	933 (25.0)	-	-	
Q4	679 (17.1)	-	-	
Q5(Richest)	663 (16.7)	-	-	

Table 4: Anthropometric indices, nutrition and substance use in DehPCS

Variables	Total n (%)	Male n (%)	Female n (%)	P-value (Chi ²)
BMI ^a				
Under-weight (<18.5)	45 (1.1)	30 (1.7)	15 (0.7)	< 0.001
Normal or health weight (18.5-24.9)	944 (23.7)	608 (34.9)	336 (15)	
Over-weight (25-29.9)	1705 (42.8)	774 (44.5)	931 (41.6)	
Obese (>30)	1286 (32.3)	329 (18.9)	957 (42.7)	
WHR (WHO) ^b				
Healthy	1229 (30.9)	574 (33)	655 (29.3)	< 0.001
Over-weight	2176 (54.7)	998 (57.3)	1178 (52.6)	
At risk for heart disease	575 (14.4)	169 (9.7)	406 (18.1)	
Smoking				
Yes	927 (23.5)	787 (45.7)	140 (6.3)	< 0.001
No	3024 (76.5)	636 (54.3)	2088 (93.7)	
Use alcohol				
Yes	481 (12.2)	437 (25.4)	44 (2)	< 0.001
No	3470 (87.8)	1286 (74.6)	2184 (98)	
Use drug				
Yes	448 (11.3)	424 (24.6)	24 (1.1)	< 0.001
No	3502 (88.7)	1298 (75.4)	2204 (98.9)	
MET-minute (daily) ^c				
1440-2100	1100 (27.8)	610 (35.4)	490 (21.9)	< 0.001
2101-2699	2312 (58.4)	636 (36.8)	1676 (75.1)	
≥ 2700	547 (13.8)	480 (28.7)	67 (3.0)	

24 hour sleep duration					
	< 7	1584 (40.0)	743 (43.0)	841 (37.7)	0.001
	≥ 7	2376 (60.0)	985 (57.0)	1391 (62.3)	
Fried food use (meals per month)					
	< 4	961(24.2)	381 (21.9)	580 (25.9)	0.009
	4-12	2834 (71.2)	1267 (72.9)	1567 (69.9)	
	>12	185 (4.6)	90 (5.2)	95 (4.2)	
Dyslipidaemia					
	Yes	3228 (80.8)	1301 (74.4)	1927 (85.7)	< 0.001
	No	768 (19.2)	447 (25.6)	321 (14.3)	
Salt use (gram/day)					
	< 5	2958 (74.4)	1253 (72.1)	1705 (76.1)	0.005
	≥ 5	1020 (25.6)	484 (27.9)	536 (23.9)	

^a Body Mass Index ^b Waist to Hip Ratio, healthy: Less than 0.80 and 0.90 in women and men respectively; Over-weight: between 0.8-1 for women and 0.9-1 for men; at risk for heart disease: Upper than 1 in women and men ^c Metabolic Equivalent Task

Table 5: Prevalence of chronic and underlying diseases diagnosed by physician, base on self-report in DehPCS

Variables		Total n (%)	Male n (%)	Female n (%)	P-value (Chi²)
Diabetes	Yes	522 (13.2)	151(9)	371 (16.6)	< 0.001
	No	3440 (86.8)	1576 (91)	1864 (83.4)	
Hypertension	Yes	774 (19.5)	226 (13.1)	548 (24.5)	< 0.001
	No	3188 (80.5)	1501 (86.9)	1687 (75.5)	
Cardiac ischemic	Yes	300 (7.6)	118 (6.8)	182 (8.1)	0.068
	No	3662 (92.4)	1609 (93.2)	2053 (91.9)	
Stroke	Yes	62 (1.56)	27(1.56)	35(1.56)	0.99
	No	3914 (98.44)	1708 (98.44)	2206 (98.44)	
Myocardial Infarction	Yes	73 (1.8)	48 (2.7)	25 (1.1)	< 0.001
	No	3889 (98.2)	1679 (97.3)	2210 (98.9)	
Chronic Lung Diseases ^a	Yes	163 (4.1)	46 (2.7)	117 (5.2)	< 0.001
	No	3799 (95.9)	1681 (97.3)	2118 (94.8)	
Thyroid Diseases ^b	Yes	535 (13.5)	64 (3.7)	471 (21.1)	< 0.001
	No	3427 (86.5)	1663 (96.3)	1764 (78.9)	
Fatty Liver	Yes	762 (19.2)	269 (15.5)	493 (22)	< 0.001
	No	3214 (80.8)	1466 (84.5)	1748 (78)	
kidney stone	Yes	813(20.5)	440 (25.4)	373 (16.6)	< 0.001
	No	3163 (79.5)	1,295 (74.6)	1,868(83.4)	
Chronic headaches	Yes	608 (15.3)	99 (5.7)	509 (22.7)	< 0.001
	No	3368 (84.7)	1636 (94.3)	1732 (77.3)	
Psychiatric disorders ^c	Yes	672(16.9)	257(14.8)	415 (18.5)	0.002
	No	3304(83.1)	1478 (85.2)	1826 (81.5)	

^a Including tuberculosis, asthma, etc ^b Including hypothyroidism, hyperthyroidism, goiter and thyroid nodules ^c Diagnosed with any kind of psychiatric disorder, other than depression

: // %