International Journal of Epidemiology, 2018, 1–9
doi: 10.1093/ije/dyy215
Cohort Profile

Cohort Profile

Cohort profile: The AZAR cohort, a health-oriented research model in areas of major environmental change in Central Asia

Sara Farhang,[1,2,3] Elnaz Faramarzi,[1] Nayyere Amini Sani,[4] Hossein Poustchi,[5] Alireza Ostadrahimi,[6] Behrooz Z Alizadeh[1,7] and Mohammad H Somi[1]*

[1]Liver and Gastrointestinal Diseases Research Center, Tabriz University of Medical Sciences, Tabriz, Iran, [2]University Medical Center Groningen, University Center of Psychiatry, Rob Giel Research Center, Groningen, The Netherlands, [3]Research Center of Psychiatry and Behavioral Sciences, [4]Department of Statistics and Epidemiology, Tabriz University of Medical Sciences, Iran, [5]Liver and Gastrointestinal Diseases Research Center, Tehran University of Medical Sciences, Tehran, Iran, [6]Nutrition Research Center, Tabriz University of Medical Sciences, Tabriz, Iran and [7]Department of Epidemiology, University Medical Center Groningen, Groningen, The Netherlands

*Corresponding author. Liver and Gastrointestinal Diseases Research Center, Tabriz University of Medical Sciences, Tabriz, Iran. E-mail: mhosseinsina@yahoo.com

Editorial decision 5 September 2018; Accepted 14 September 2018

Why was the cohort set up?

The current pattern of diseases in Central Asia and Iran has shifted from communicable and mother-child conditions, such as perinatal mortality and maternal death, towards non-communicable chronic diseases, such as metabolic, heart, respiratory and neuropsychiatric disorders. These chronic diseases now make up the greatest burden of human ill health in Iran[1,2] and generally reflect the complexity of human disease aetiology. There is evidence suggesting that a multidimensional interaction between small effects of the genome, exposome, diet, microbiome and internal microenvironment underlies the disease aetio-pathogenesis.[3]

Although dynamic interactions between host, environmental and genetic risk factors are largely population-specific, our current knowledge is limited to data that have been collected from Western countries. This could well lead to gaps in our understanding of disease mechanisms, and there may be undiscovered mechanisms in less well-investigated populations. These include a lack of knowledge about disease aetiology and pathogenesis in populations in the Middle East and Western Asia. For example, Iran is a developing country with a population of over 80 million, and it has one of the most complex cultures in the world.[4] This geographical region has played an important role in the development of modern human life,[5] but it has major differences in ecological, cultural and social aspects compared with Western countries.[6] In addition, Western Asia is currently facing rapid and sometimes disastrous changes to its ecological system: for instance the rapid shrinkage of one of the world's largest salt lakes (Lake Urmia, Figure 1), which has over 70 million people living within a radius of 500 km.[7] These natural changes are affecting the long established matrix of environmental risk factors and their combined contribution to human health in the region. A prospective cohort design is the best method to obtain valuable data on gene-environmental interactions.[8]

We hypothesized that ecological changes—in the form of natural disasters—induce rapid and sharp changes to

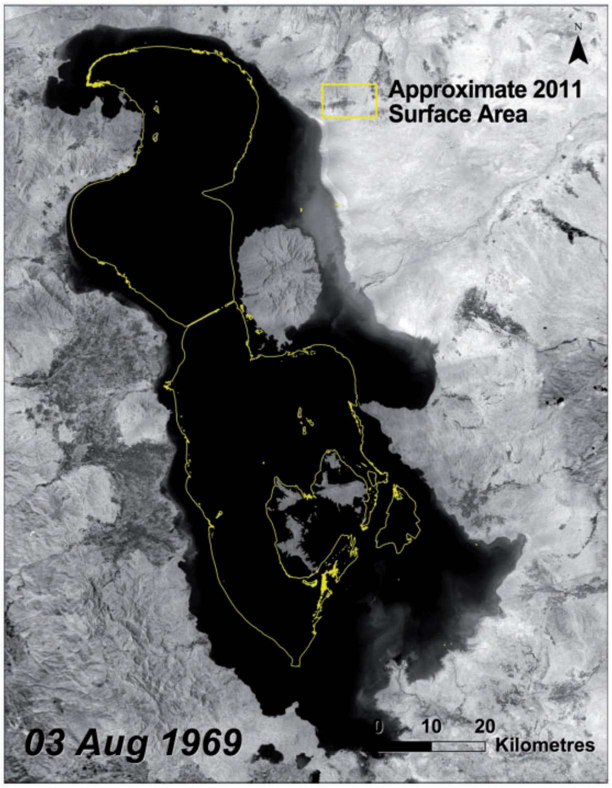

Figure 1. Change in Lake Urmia's surface area in recent years. Images for 1972–2011 from Landsat data; 2011 image was visualized by the United Nations Environmental Programme (UNEP GRID), Sioux Falls, SD.

the composition of environmental risk factors, and their interplay with host genetic susceptibility eventually affects disease frequencies. Such a cohort study could yield new understanding of human diseases, especially when a natural event occurs in a population or region that has not been the focus of scientific investigations into disease mechanisms. We would like to test this hypothesis by setting up a large population-based cohort study, the AZAR cohort, in the region of the hugely shrunk Lake Urmia. We aim to study the long-term effects of environmental changes due to natural disasters and human interventions, in interaction with genetic factors and their effects on common human diseases. The AZAR cohort focuses on the impact of the unique intervention of the Lake Urmia disaster, on the prevalence of environmental factors and their interactions and on the effects of the exposome's composition on the incidence of common diseases in a genetically stable population. The AZAR cohort aims specifically to: (i) identify environmental risk factors for common diseases by comprehensively assessing external exposures, lifestyle (including dietary patterns), social determinants and ethnic and individual factors; and (ii) provide a model for population-based studies of causal disease mechanisms by using the advantage of natural changes in regions of rapid environmental and social transition.

Who is in the cohort?

Design

The AZAR cohort is a large population-based cohort and part of the Prospective Epidemiological Research Studies in Iran (the PERSIAN Cohort Study) initiative.[9] The AZAR cohort will focus on common diseases, including cardiovascular, digestive, hepatic, renal, metabolic, psychiatric and respiratory disorders and cancers, in a region of rapid environmental change. It will provide a unique set of data on genetic predisposition, family history, phenotypes and DNA sequence analyses, together with simultaneous environmental and lifestyle factors.

Ethics, consent and permissions

The AZAR cohort has been approved by the Ethical Committee of Tabriz University of Medical Sciences (record number: tbmed.rec.1393.205). Written informed consent was obtained from all participants and they could leave the study at any time, for any reason.

Target study population

General characteristics

The AZAR cohort was established in Shabestar County in East Azerbaijan province, north-western Iran; the region lies at an altitude of 1532 m, covers an area of 2630 km^2 and its climate is influenced by its proximity to Lake Urmia. Its population was 124 499 (2011 census: 37 387 families, with a female:male ratio of 0.98:1.0), and with 48.5% of the population resident in urban areas. The majority of the population (68%) is aged between 15 and 64 years. Everyone is registered with the family physician (GP, general practitioner) programme.

The AZAR cohort has been set up in three phases: (i) pilot study; (ii) enrolment of participants; and (iii) regular follow-up of subjects for 15 years (Figure 2). The participants' age, level of education and marital status are described in Table 1. Participants are also categorized based on their socioeconomic status (SES) as determined on the basis of job title, car ownership, number of trips made (per year), type of travel, having a personal computer, home ownership and having multiple jobs. The data are given in Table 2.

Geography and ethnicity

The name AZAR comes from Azerbaijan, the name of the geographical area in north-western Iran, and also from Azeri, the ethnic name for the resident population. Iranian Azeris are the largest minority group in the country; they have a Caucasian ethnic background and comprise the world's largest Azeri population. The Azeri are a

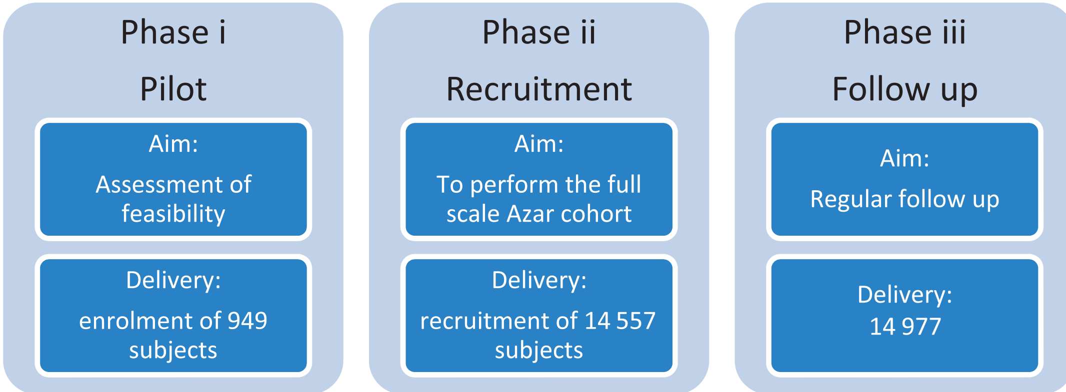

Figure 2. The three phases of the AZAR cohort.

Table 1. Baseline characteristics of study population

		Pilot			Enrolment		
		Men ($n = 433$)	Women ($n = 516$)	Total ($n = 949$)	Men ($n = 6707$)	Women ($n = 8290$)	Total ($n = 14\,997$)
Age (years)	Mean (SD)	51 ± 8.8	49.3 ± 8.5	50.0 ± 8.7	50.1 (9.3)	49.2 (9.3)	49.6 (9.3)
	35–39	52 (12)	78 (15.1)	130 (13.7)	1032 (15.4)	1474 (17.8)	2506 (16.7)
	40–44	58 (13.4)	84 (16.3)	142 (15)	1117 (16.6)	1534 (18.5)	2651 (17.7)
	45–49	77 (17.8)	116 (22.5)	193 (20.3)	1178 (17.6)	1504 (18.2)	2682 (17.9)
	50–54	93 (21.5)	105 (20.3)	198 (20.9)	1146 (17.1)	1302 (15.7)	2448 (16.3)
	55–59	75 (17.3)	63 (12.2)	138 (14.5)	979 (14.6)	1105 (13.3)	2084 (13.9)
	60–64	41 (9.5)	37 (7.2)	78 (8.2)	736 (11.0)	789 (9.5)	1525 (10.2)
	65–69	36 (8.3)	31 (6.4)	67 (7.1)	468 (7.0)	524 (6.3)	992 (6.6)
	70–85[a]	1 (0.2)	2 (0.4)	3 (0.3)	51 (0.8)	58 (0.7)	109 (0.7)
Educational level							
Illiterate		16 (3.7)	38 (7.4)	54 (5.7)	585 (8.7)	1920 (23.2)	2505 (16.7)
High school graduated		323 (74.6)	394 (76.3)	717 (75.5)	5294 (78.9)	5875 (70.8)	11 169 (74.4)
University education		94 (21.7)	84 (16.3)	178 (18.8)	828 (12.3)	495 (6)	1323 (8.8)
Marital status							
Married		425 (97.5)	450 (87.2)	875 (92.2)	6626 (98.8)	7269 (87.7)	13 895 (92.7)
Not married		8 (2.5)	66 (12.8)	74 (7.8)	81 (1.2)	1021 (12.3)	1102 (7.3)

[a]Participants over 70 were only recruited in pilot phase.

Table 2. Socioeconomic status of the total target population in AZAR cohort study

Socioeconomic status	Male ($n = 6707$)	Female ($n = 8290$)
Very high	1656 (24.8)	1196 (14.5)
High	1399 (20.9)	1477 (18)
Middle	1376 (20.6)	1543 (18.8)
Low	1223 (18.3)	1777 (21.6)
Very low	1032 (15.4)	2232 (27.1)

Turkic ethnic group who live in Iran, the independent Republic of Azerbaijan, or Turkey. European-specific haplogroups of mitochondrial DNA are reported to be present in the majority of Iranian Azeris.[10] People living in this region are known as Turkish language-speaking and a mix of eastern Transcaucasians and Medians, an ancient Persian descent of ancient historians, including Herodotus, Polybius and Strabo.[11] This ancient Turkic population lives in north-western Iran, south of Lake Urmia (Figure 1).

Although the population of Azerbaijan is culturally diverse, they have an autochthonous background with a genetic admixture derived from Central Asians of Turkmen, the Georgians and the Armenians; this is supported by genetic markers[12] and mitochondrial DNA analyses indicate that the main relationship of Turks with Iranians is through a larger West Eurasian group. Azerbaijani Turks are a mixed population of Caucasians, Iranians and Near-Easterners, Europeans, and Turkmen, in order of greater similarity.[13] A similar study indicated that Iranian Azeris are more closely related to the people of Georgia than to other Iranians, or to Armenians.[6]

Regional disease morphologies

Based on recent reports, the prevalence of cancer is in general comparable to that in other parts of Iran, and the most common sites for cancer are the gastrointestinal tract and breast.[14] Reports also indicate that cardiovascular diseases are one of the leading causes of death in this region.[15]

Major environmental change

A large part of the Azerbaijan region is mountainous, with deep valleys and fertile lowlands. This region includes an exceptional lake—Lake Urmia was the largest lake in the Middle East and the sixth largest saltwater lake on the earth, with a salt concentration that has recently risen to more than 300 g/l (eight times more salty than seawater). In the past 20 years, this lake has shrunk to 10% of its former size (Figure 1). Several reasons (damming of rivers, groundwater extraction and building of a causeway) have influenced this environmental shift, and a broad range of consequences are anticipated as a result of this rapid change.[6,16] There are ongoing efforts to restore the lake.[17]

Advisory board

The AZAR cohort's scientific board includes internists (with different subspecialties), nutritionists, pharmacologists, epidemiologists, cardiologists, geneticists, psychiatrists and environment health specialists. These principal investigators are academic researchers affiliated to Tabriz University of Medical Sciences, with well-established track records in epidemiological and cohort studies. The board is responsible for all the executive and scientific processes, including biobanking, quality assurance, the scientific direction of the cohort, financial management and fund raising.

The field complex

The cohort's head office is located in the region's main town, Shabestar. There is a reception and waiting room, and 12 rooms for taking samples, anthropometric and physical examinations, interviews and a rest room. All the rooms are arranged according to the steps in the enrolment process. There is a small laboratory for processing the samples. The biobank with seven -80°C and two -20°C freezers is located in a separate section, controlled by an uninterrupted power supply (UPS) system and with real-time temperature monitoring.

Human resources

We set up a team of trained research assistants of medi-care staff who have a medical qualification, from Tabriz University of Medical Sciences. This team includes a field supervisor, two general practitioners and one ophthalmologist, 12 interviewers in different sections, four sampling technicians and two epidemiologists. The staff perform the ongoing interviews using predefined and validated research protocols and questionnaires. All of the interviewers and laboratory staff were trained by principal investigators and/or trainers to use the standard questionnaires and screening tools, handle samples and, in the biobank organization, work with the digital data acquisition sheet and on ethical concerns.[18] Further booster training courses were also established by the same trainers to provide the best coordination.

How often have they been followed up?

The AZAR cohort is being performed in three phases: pilot, recruitment and follow-up (Figure 2).

Pilot phase

The aim of the pilot phase (2014–15) was to appraise the feasibility of the study and reveal any unmet needs for undertaking the full-scale AZAR cohort study. Specifically, the pilot phase aimed to: implement valid and reproducible methods and test the structured questionnaires in measuring environmental factors; assess the ascertainment methods; evaluate the response rate from the study population; establish follow-up routines; test the validity of measuring disease outcomes among the enrolled subjects,; implement effective procedures for biosampling; estimate the data and storage capacity needed; and set up the data acquisition, preservation and storage and quality controls. The pilot phase was established in Khameneh, a city in the Central District of Shabestar County. It was started with a meeting held in the municipality of Khameneh, where local officials also joined to show their support. The project was also announced on websites and in the local media.

Inclusion and exclusion criteria

The inclusion criteria were: (i) permanent residence in Shabestar district (minimum of 9 months); (ii) written informed consent; (iii) at least one Azeri parent; and (iv) age between 35 and 70 years at the time of enrolling in the study. Exclusion criteria were: (i) having a diagnosis of a disabling psychiatric disorder; and (ii) having a diagnosis of a disabling physical illness.

Recruitment for the pilot phase began in September 2014 (Figure 3). Invitations were sent to selected people who confirmed via telephone call to the research facility within 2 weeks. The telephone operator was trained to provide enough information. Participants were asked not to clip their nails for 1 week, to wash their hair, bring their medications and medical records and attend the intake appointment after fasting overnight. The day before the

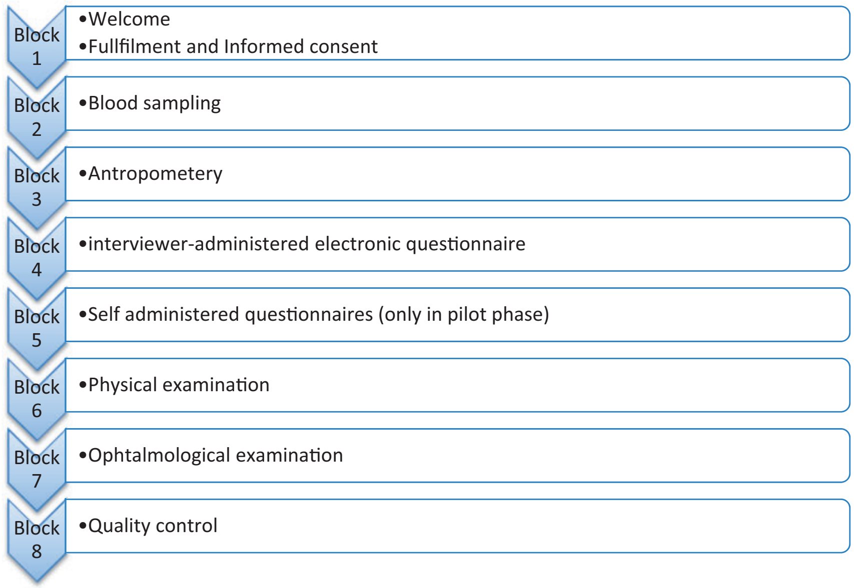

Figure 3. Flow diagram of data collection in the AZAR cohort.

appointment, a reminder call was made. We implemented a block design for data collection and biosampling, and the whole procedure took 4 h.

Enrolment phase

The AZAR cohort was launched in 2015. The enrolment phase was initiated in Shabestar city. The same procedure as when announcing the pilot phase was followed. The invitation process, questionnaires, sampling and measurements in the enrolment phase were the same as for the pilot, except for the psychiatric evaluations, which were technically not feasible.

Follow-up phase

The aim of the follow-up phase will be to record the environmental changes, incidental new diagnoses and changes in individual factors of participants over a period of 15 years. For assessing the risk and longer-term resilient factors, the whole evaluation (as described in the enrolment section) will be repeated every 5 years. The follow-up phase will be performed in three ways (Figure 4). Each participant was also instructed to contact the AZAR team in the event of hospitalization or if they received a diagnosis of a chronic disease. The databases of the field complex are also linked to local medical registries to enable searching for any diagnosis given to a participant during the follow-up period. Participants will be contacted by telephone and also through their family physician. If data are incomplete or further evaluation is needed, a face-to-face reassessment will be arranged to collect information from the family physician, municipality or disease registry systems.

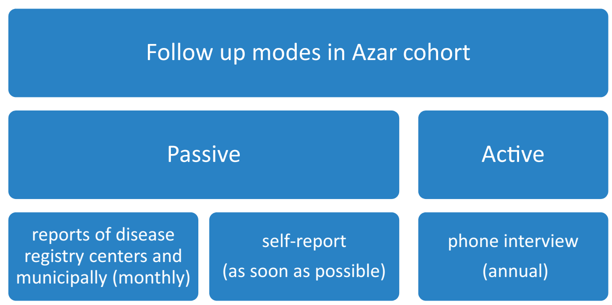

Figure 4. Modes of follow-up in the AZAR cohort.

The AZAR follow-up team completes follow-up questionnaires and records occurrence of any disease or hospital admissions that have taken place. If the participant cannot be contacted, the team visits them at their home in urban areas, or in rural areas the *Behvarz* (local health worker) contacts the participant to complete a follow-up questionnaire. In the case of a death, the follow-up team will collect the medical records about course of disease and cause(s) of death, including clinical, pathology and hospital records. The whole process has been described in the PERSIAN cohort study protocol.[9]

What has been measured?

We used a block design to perform the interviews, where one or more questionnaires were completed per block. Data collection covers several domains, including general, medical, nutrition and bio-sampling (Figure 4).

Block 1

Following a brief welcome talk, the process is explained and informed consent obtained.

Block 2

After an overnight fast of 12–14 h, 25-ml blood samples are collected from peripheral vessels, and serum, plasma and buffy coat are separated. The samples are stored in two-dimensional barcoded aliquots at –80°C for future biochemical analyses. After sampling, some biochemical parameters such as serum fasting blood glucose (FBS), complete blood count (CBC), blood urea nitrogen (BUN), creatinine, lipid profiles (total cholesterol, low- and high-density lipoprotein, triglyceride) and liver enzymes (aspartate, alanine and gamma-glutamyl aminotransferase) are measured. Urine analyses are performed. Urine, hair and nail samples are collected and stored. The procedures have been described in detail previously.[9]

Block 3

Height is recorded to the nearest 0.5 cm using a mounted tape measure, with the subject's arms hanging freely by their sides. Barefoot subjects with only light clothing have their weight recorded to the nearest 0.1 kg on a Seca scale. Body mass index is calculated and waist, hip and wrist circumferences of subjects are measured according to the USA National Institutes of Health (NIH) guidelines. Participants are then given breakfast before having a comprehensive set of interviews and examinations.

Blocks 4 and 5

An interviewer supervises an electronic questionnaire consisting of 55 questions and 482 items, covering many aspects of the participant's life that may have an impact on their health status. A few self-report questionnaires are also used to gather information in specific domains. The questions are organized into the following major categories: demographics, socioeconomic status, lifestyle factors (including physical activity, substance use), occupational history (including occupational exposures), past medical history, medicine use (past and present), family medical history, gynaecological and obstetric history (for women), oral and dental health, circadian rhythm, dietary habits (including food processing and cooking methods), psychiatric evaluations (affective symptoms, personality traits, sleep pattern) and environmental factors (including mobile use, pesticide use, housing, source of water). Dietary pattern is measured using a food frequency questionnaire (FFQ) covering 130 food items, including bread and cereals, meat and dairy products, oils, sweets, legumes, vegetables, and fruits and condiments, as well as cooking methods.

Block 6

A physical examination is performed by a trained general practitioner according to the guideline developed for the cohort.

Block 7

This block uses the information about past medical history of ophthalmological disorders collected by an optometrist. Then, together with a trained ophthalmologist, a thorough examination is made using an autorefractor keratometer,Ssnellen chart, retinoscope, tonometer and slit lamp.

Data processing and quality check

A quality check is performed for each block at four levels. A basic check is made by the software that appropriate data are recorded in each part of the questionnaire in terms of variable types, length and measurement levels. Missing data automatically alert the operator to fill in values. Furthermore, regular quality surveillance is carried out in the initial phase by an independent staff member using a validated checklist, supervised by Tabriz University. For each participant, the staff checks: three aspects of the data and specimen acquisition; general questionnaires; anthropometric measurements; medical and disease screening questionnaires; and biological samples acquisition and biobank maintenance. If data are not completed or missing, the record will be indexed and sent back to another team to fill the data gap. Only checked records can be imported into the database. The records are then evaluated by an epidemiologist and, last, data are randomly checked by the quality control administrator, who is an AZAR principal investigator with full access to the entire records database.

From each participant, 15 separate blood specimens (whole blood, serum, buffy coat, plasma) are preserved in 1 to 1.5 ml 2-D cryotube (Micronic, Lelystad, the Netherlands). The samples are kept in a specific box for each participant, labelled with their individual smart 11-digit code. All the cryotubes are stored in −80°C refrigerators. A sample of hair (approx. 200–300 strands, 1–3 cm long) and nail clippings from all fingers are collected and coded in aluminium foil packets.

What has it found?

We can report the baseline characteristics of the pilot and recruitment phases here. Table 1 shows the demographic characteristics of the participants. The pilot phase was closed after reaching the target number of 949 participants. The enrolment phase recruited a further 14 557 participants, giving 15 006 in total. After data cleaning, data from 14 997 participants was validated. The overall response rate of those invited was 93%. The prevalence of common diseases is estimated according to medical history and laboratory findings (Table 3).

What are the main strengths and weaknesses?

The main strength of the AZAR cohort study lies in the innovative concept of the study design: to see how a major

Table 3. Disease prevalence of study population at pilot ($n=949$) stage compared with all samples ($n=14\,997$)

	Pilot			Enrolment		
	Men ($n=433$)	Women ($n=516$)	Total ($n=949$)	Men ($n=6707$)	Women ($n=8290$)	Total ($n=14\,997$)
Obesity	84 (19.4)	214 (41.5)	298 (31.4)	1741 (26)	3899 (47.1)	5640 (37.6)
Type 2 diabetes	56 (12.9)	76 (14.7)	132 (13.9)	852 (12.7)	1245 (15)	2097 (14)
Hypertension	72 (16.6)	114 (22.1)	186 (19.6)	1320 (19.7)	2362 (28.5)	3682 (24.7)
Depressive disorder	33 (7.6)	83 (16.1)	116 (12.2)	534 (8)	2020 (24.4)	2554 (17)
Renal failure	2 (0.5)	1 (0.2)	3 (0.3)	39 (0.6)	36 (0.4)	75 (0.5)
Hyper/hypo thyroidism	5 (1.2)	74 (14.3)	79 (8.3)	135 (2)	1171 (14.1)	1306 (8.7)
Ischaemic heart disease	37 (8.5)	15 (2.9)	52 (5.5)	393 (5.9)	338 (4.1)	731 (4.9)
Chronic respiratory disorder	10 (2.3)	19 (3.7)	29 (3.0)	190 (2.8)	345 (4.2)	535 (3.6)
Hepatitis B	2 (0.5)	2 (0.4)	4 (0.4)	25 (0.4)	17 (0.2)	42 (0.3)
Hepatitis C	3 (0.7)	–	3 (0.3)	6 (0.1)	–	6 (0.0)

environmental change affects complex human diseases at a population level against a relatively stable genetic susceptibility. We present a cross-over study design on a large population (>15 000) cohort, and will compare the data on common disease-associated factors and their effects on disease frequencies, before and after the natural changes occurring at Lake Urmia.

In this report, we present the concept that applies to designing a study on a cohort exposed to a major natural change, as an innovative way to discover some of the environmental factors involved in complex human diseases. We reasoned this approach would have more scientific relevance if it were applied to a less well-studied population. To this end we described the aim and setting of AZAR cohort, a uniquely population-based study in the region of natural disaster of Lake Urmia. We argued that the AZAR cohort offers a chance to study the interactions of a rapidly changing ecological system and environmental factors, on the backbone of the relatively constant genetic susceptibility. We expect to observe new equations between genetic susceptibility and environmental factors, which will reshape the incidence of common multifactorial diseases. We present that our design resembles a case-cross-over interventional study, when the data of common disease-associated factors and their effects on disease frequencies will be compared before and after the natural intervention at Lake Urmia.

The pilot study revealed a population primarily of young adults and mature adults, with 50% of the study population already affected by at least one chronic disease. This suggests a vulnerable population that is expected to be affected by the rapidly changing environment. The relatively young age of the population is a strength of the cohort, offering the possibility of long-term follow-up, which should eventually lead to good estimates of disease frequencies and risk effect sizes.

The underlying idea of the PERSIAN cohort initiative was to provide a large sample size and cover the majority of races across the region. We would suggest that the study will be more valuable if subcohorts have specific and unique characteristics. Compared with the AZAR cohort, other parts of the PERSIAN cohort have different racial backgrounds and more stable environments. The AZAR cohort covers a population at a unique location and with precise timing of environmental changes. It thus avoids several shortcomings of other studies. The AZAR cohort includes an ethnically homogeneous population, which will likely offer more power to test associations. This cohort is also the first and largest study focusing on people with Turkish ethnicity.

In general, and even in the absence of natural disasters, there is a need to develop evidence-based health care policies in the Middle East and Western Asia which are adapted to local customs and beliefs. The AZAR cohort is expected to deliver novel results on disease outcomes in the Azeri, and will eventually lead to targeted evidence-based health and medical policies.

One of the issues in the study is participants' immigration. This undermines the recording of health events. The AZAR cohort is engaged actively with a nationwide health indexes registration switch interface board (SIB) network system, which records every Iranian's health status online with direct data entry available across the country. The software allows individuals to be tracked in the health network and facilitates the outcome measurements for AZAR cohort. Furthermore, Tabriz University of Medical Sciences (TUoMS) has available disease registry programmes and population-based registries covering common diseases in the cardiovascular, metabolic, gastrointestinal and cancer domains, which are linked to the AZAR cohort data.

One issue is the comparability of our findings with international cohorts, but we expect high quality diagnostics and international coding systems will overcome this.

Can I get hold of the data? Where can I find out more?

To further support national and international collaborations, the AZAR cohort collaborates with international experts. It uses online communication tools via its website to disseminate the information collected from the AZAR cohort, to increase awareness of the project and to attract international collaborators [http://azarcohort.com], but unfortunately the webpage is only available in Persian at the moment.

Profile in a nutshell

- The AZAR Cohort study investigates common, non-communicable disorders with a focus on gene-environmental interaction in the context of rapid environmental and social changes.
- The cohort studies Iranian Azeris, the largest minority population group in Iran. They have a Caucasian ethnic background and are the world's largest Azeri population. Azeris are a Turkic ethnic group. The study is located in Azerbaijan province in north-western Iran. The cohort population lives around Lake Urmia, a large saltwater lake which has shrunk to just 10% of its former area within the past 20 years.
- The pilot phase started in 2014 by enrolling 949 individuals; the recruitment phase in 2015 enrolled 14 557 participants aged over 35–70 years old, yielding a total of 15 006 participants.
- For assessment of risk and resilient factors, all the tests will be repeated every 5 years, for a period of 15 years after inclusion. Data will also be derived from disease and death registries.
- The main data categories are: demographics, socio-economic status, lifestyle factors (including physical activity, substance use), occupational history (including occupational exposures), past medical history, medicine use (past and present), family medical history, gynaecological and obstetrics history (for women), oral and dental health, circadian rhythm, dietary habits (including food processing and cooking methods), psychiatric evaluations (affective symptoms, personality traits, sleep pattern) and environmental factors (including water supply, housing, cellphone use, pesticide exposure). Dietary pattern is measured using a food frequency questionnaire that covers 130 food items, including bread and cereals, meat and dairy products, oils, sweets, legumes, vegetables, fruits and condiments, as well as cooking methods.
- The AZAR cohort has an informative website in Persian [http://azarcohort.com/]. Please contact the corresponding author for data requests.

Funding

This work and the study are funded by Liver and Gastrointestinal Diseases Research center, Tabriz University of Medical Sciences, East Azerbaijan, Iran and the Deputy of Technology and Research at the Ministry of Health and Medical Education (MOHME), Iran.

Acknowledgements

The study is supported by Liver and Gastrointestinal Diseases Research Centre, Nutrition Research Centre and Research Centre of Psychiatry and Behavioural Sciences, all at Tabriz University of Medical Sciences, Tabriz, Iran. We thank Jackie Senior for editing the manuscript. Authors would like to appreciate the contribution of the investigators and the staff of Azar cohort study, Shabestar health center and Persian cohort study staff for their technical support.

Author Contributions

All authors have contributed to the conception and design of the AZAR cohort. EF, SD and AO have contributed to data acquisition. NAS performed the statistical analyses. SF drafted the manuscript. MRR, MHS and BZA interpreted the results. SF, MHS and BZA revised the manuscript. All authors have read and approved the submitted version of the manuscript.

Conflict of interest: The authors declare that they have no competing interests.

References

1. Saadat S, Yousefifard M, Asady H, Moghadas Jafari A, Fayaz M, Hosseini M. The most important causes of death in Iranian population; a retrospective cohort study. *Emerg (Tehran)* 2015;**3**:16–21.
2. Kassebaum NJ, Barber RM, Bhutta ZA. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;**22**:743–800.
3. Pekkanen J, Pearce N. Environmental epidemiology: challenges and opportunities. *Environ Health Perspect* 2001;**109**:1–5.
4. Duckett B. The history of Iran: ABC-CLIO; Santa Barbara, 2012. *Ref Rev* 1998;**12**:42.
5. Taagepera R. Size and duration of empires: growth-decline curves, 600 BC to 600 AD. *Soc Sci Hist* 1979;**3**:115–38.
6. Derenko M, Malyarchuk B, Bahmanimehr A *et al*. Complete mitochondrial DNA diversity in Iranians. *PLoS One* 2013;**8**:e80673.
7. Pengra B. *The Drying of Iran's Lake Urmia and its Environmental Consequences*. Sioux Falls, SD: United Nations Environment Programme Global Environmental Alert Service (UNEP-GEAS), 2012.
8. Manolio TA, Bailey-Wilson JE, Collins FS. Genes, environment and the value of prospective cohort studies. *Nat Rev Genet* 2006;**7**:812–20.
9. Poustchi H, Eghtesad S, Kamangar F *et al*. Prospective epidemiological research studies in Iran (The PERSIAN Cohort): rationale, objectives and design. *Am J Epidemiol* 2018;**187**:647–55.

10. Asgharzadeh M, Kafil HS, Ranjbar A, Rahimipour A, Najati K, Nobar MR. Molecular diversity of mitochondrial DNA in Iranian Azeri ethnicities vis-a-vis other Azeris in Asia. *Iran J Biotech* 2011;**9**:120–25.
11. Ghirshman R, Bivar ADH. *Ancient Iran. Encyclopaedia Britannica*. Chicago, IL: Encyclopaedia Britannica Inc., 2009.
12. Zerjal T, Wells RS, Yuldasheva N, Ruzibakiev R, Tyler-Smith C. A genetic landscape reshaped by recent events: Y-chromosomal insights into Central Asia. *Am J Hum Genet* 2002;**71**:466–82.
13. Quintana-Murci L, Chaix R, Wells RS *et al*. Where West meets East: the complex mtDNA landscape of the Southwest and Central Asian Corridor. *Am J Hum Genet* 2004;**74**:827–45.
14. Somi MH, Farhang S, Mirinezhad SK, Naghashi S, Seif-Farshad M, Golzari M. Cancer in East Azerbaijan, Iran: results of a population-based cancer registry. *Asian Pac J Cancer Prev* 2008;**9**:327–30.
15. Naghavi M. *Mortality Mapping in 18 Provinces of Iran: Ministry of Health and Medical Education*. Tehran: MHME, 2004.
16. Hoseinpour M, Fakheri Fard A, Naghili R. Death of Urmia Lake, a silent disaster investigating causes, results and solutions of Urmia Lake drying. In: *Fir*st International Applied Geological Congress, Islamic Azad University-Mashad Branch*, 26-28 April, 2010*. Mashhad, Iran, 2010.
17. Zarghami M, AmirRahmani M. A system dynamics approach to simulate the restoration plans for Urmia Lake, Iran. In: Matsumoto A (ed). *Optimization and Dynamics with Their Applications*. Singapore: Springer, 2017.
18. Eghtesad S, Mohammadi Z, Shayanrad A *et al*. The PERSIAN cohort: providing the evidence needed for healthcare reform. *Arch Iran Med* 2017;**19**:291–95.