BMJ Open

Asylum seekers' and Refugees' Changing Health (ARCH) study protocol: an observational study in Lebanon and Denmark to assess health implications of long-distance migration on communicable and non-communicable diseases and mental health

Andreas Halgreen Eiset [iD],[1,2] Michaelangelo P Aoun,[3] Ramzi S Haddad,[3] Wadih J Naja,[3] Kurt Fuursted,[4] Henrik Vedel Nielsen,[4] Christen Rune Stensvold,[4] Monica Stougaard Nielsen,[2] Annemarie Gottlieb,[2] Morten Frydenberg,[5] Christian Wejse[1,6]

[Check for updates]

ABSTRACT

Introduction By end of 2018, the European Union countries hosted approximately 2.5 million refugees and Lebanon alone hosted more than 1 million. The majority of refugees worldwide came from Syria. The prevailing study design in published studies on asylum seekers' and refugees' health leaves a number of fundamental research questions unanswerable. In the Asylum seekers' and Refugees' Changing Health (ARCH) study, we examine the health of a homogeneous group of refugees and asylum seekers in two very different host countries with very different migration histories. We aim to study the health impact of the migration process, living conditions, access to healthcare, gene–environment interactions and the health transition.

Methods and analysis ARCH is an international multisite study of the health of adult (>18 years old) Syrian refugees and asylum seekers in Lebanon and Denmark. Using a standardised framework, we collect information on mental and physical health using validated scales and biological samples. We aim to include 220 participants in Danish asylum centres and 1100 participants in Lebanese refugee camps and settlements. We will use propensity score weights to control for confounding and multiple imputation to handle missing data.

Ethics and dissemination Ethical approval has been obtained in Lebanon and Denmark. In the short term, we will present the cross-sectional association between long-distance migration and the results of the throat and wound swab, blood and faeces samples and mental health screenings. In the longer term, we are planning to follow the refugees in Denmark with collection of dried blood spots, mental health screenings and semistructured qualitative interviews on the participant's health and

Strengths and limitations of this study

► In the Asylum seekers' and Refugees' Changing Health study, we examine the physical and mental health of a vulnerable and hard to reach population.

► We examine the health of a homogeneous group of refugees and asylum seekers in two very different host countries with very different migration histories which allows us to approach a range of research questions not possible with the traditional comparison to the autochthonous population in the host country.

► While we implement state-of-the-art epidemiological and statistical methods including 'standard-mortality-ratio-weighted' analysis on multiple imputation and comprehensive bias analysis to inform our analysis, our study is observational by nature with risk of residual confounding.

► Refugees are a difficult-to-reach population and thus the risk of selection bias is high and the sampling frame in Lebanon may be inaccurate despite efforts to secure the opposite.

► Inclusion in Lebanon and Denmark has commenced and is planned to continue until end of 2019.

access to healthcare in the time lived in Denmark. Here, we present an overview of the background for the ARCH study as well as a thorough description of the methodology.

INTRODUCTION

By end of 2018, the European Union hosted approximately 2.5 million refugees
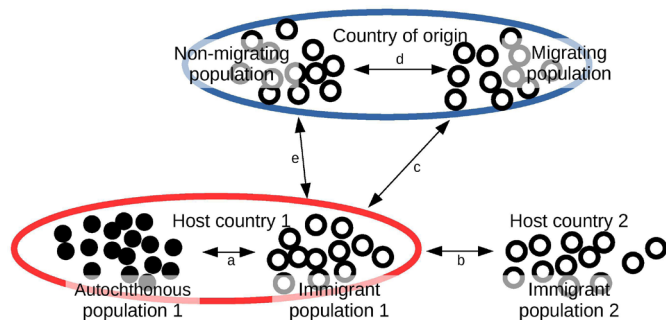
Figure 1 Possible research questions according to studied population and stage in migration. Modified from Spallek *et al*[6]. a: Autochthonous population versus immigrant pop. Differences in exposures; vulnerability of migrants; differences in access to healthcare; gene/environment effects. b: Immigrant population 1 versus immigrant population 2: differences in living conditions and healthcare facilities; probability of migrating to host country 1 versus host country 2; gene/environment effects. c: Migrating pop. versus immigrant pop: the health effects of migration; living conditions and healthcare facilities in host country; gene/environment effects. d: Non-migrating pop. versus migrating pop: healthy migrant effect. e: Non-migrating pop. versus immigrant pop: living conditions and healthcare facilities in host country; healthy immigrant effect; gene/environment effects. Comparing a country's non-migrating population with the migrating population (ie, those that has decided to migrate but has not yet immigrated to the host country; scenario a) is the most direct way of evaluating the healthy migrant effect. However, it may be very difficult to define and include these populations, particularly in refugee situations. The same is the case when comparing the non-migrating population with the immigrant population (scenario b) which allows for investigating the healthy *immigrant* effect—that is if those who actually immigrated to the host country are more healthy than those that did not migrate to begin with. Comparing the autochthonous population and the immigrant population (scenario c) is the most frequently used design in migrant research and allows for a range of possible research questions. Scenario d compares a population before and after migration, while scenario e compares immigrant populations from the same country of origin in two different host countries. The first allows for an estimate of the health effects of migration while the latter could give information on the health effects of differences in the health reception in the host country, living conditions and access to healthcare.

(excluding asylum seekers). Of these, Germany hosted 1 million and Denmark hosted 36 000.[1–3] In comparison, Lebanon hosted more than 1 million registered refugees, making Lebanon the country with the highest ratio of refugees to autochthonous population, even when disregarding the substantial number of unregistered refugees, resulting in a shut down of the asylum-seeking system for Syrians seeking refuges in Lebanon. By the end of 2018, the majority of refugees worldwide came from Syria (6.7 million), Afghanistan (2.7 million) and South Sudan (2.3 million). Most of the registered refugees in Lebanon came from Syria (944 181) and Iraq (4458); the majority of refugees in Denmark came from Syria (19 698) and Eritrea (4610).[4]

By far, the most common study design applied in migrant and refugee health research is a cross-sectional comparison, most commonly of a heterogeneous group of 'migrants' with the autochthonous population (ie, the 'host-country denominator').[5 6] This allows investigating differences in exposures between the autochthonous and migrant populations as well as health system and policy research; however, the design makes important confounders and biases difficult to disentangle and leaves many questions unanswerable (see figure 1, scenarios a, b, c, d and e). In this way, conflicting evidence has emerged: several studies have reported an apparent health advantage of migrants in general[7–16] and for asylum seekers and refugees specifically,[11 16–18] which has fuelled in particular the theory of 'the healthy migrant effect' (derived from 'the healthy worker effect'). This theory, originally put forth by Marmot *et al*[19] in a study of presumably primarily economic immigrants to England, postulates that it is the healthiest individuals who reach the host country, even though such a claim should be investigated using a 'country of origin denominator' (figure 1).[20] In contrast, other studies report that migrants are disproportionately affected by communicable, non-communicable and psychiatric diseases compared with the host population.[5 21–35 35 36] On a side note, this confounding of the effect of migration on health has only been examined for immigrants in high-income countries where other drivers of migration may be at work compared with immigration to low-income and middle-income countries.[20 37] Razum and Twardella proposed the health transition theory to try to accommodate these paradoxical findings.[38] Their theory posits an immediate decrease in maternal, child and communicable disease morbidity and mortality as the immigrants settle in the host country and a gradual rise in morbidity and mortality from non-communicable diseases prevailing in most migrant receiving countries. In short, the health transition theory proposes a 'regression to the mean' line of thought to the changing health in migrants. However, a number of biases may be underlying those findings,[39] including the indiscriminate dichotomisation of migrants versus 'non-migrants'.[40 41]

For a better understanding of asylum seekers' and refugees' health, one would need to consider the unique circumstances of this population and approach the research questions accordingly[6 42 43] using appropriate and modern epidemiological methods in line with other fields of health research.[44–50] We do not claim to have fully accomplished this task; however, we present an ongoing research project to compare newly arrived refugees in two substantially different host countries and with very different migration histories. This allows us to investigate the health impact of the migration process, differences in living conditions and access to healthcare, and gene–environment interactions (figure 1, scenarios a, b, and c). Furthermore, with follow-up in Denmark, we plan to investigate the health transition in this refugee population. A short illustrative example follows using post-traumatic stress disorder (PTSD) as the outcome of interest: PTSD

Table 1 Variables collected in the study

Dimension	Item		Instruments
Demographics	Sex; age; family; marital, educational, employment, and socioeconomic (before migration and present) status; number of children and persons in household; number of rooms in household; Syrian governorate of origin		Questionnaires and scales
Migration history	Date of emigration and immigration; means of and companions during migration; witness to or involved in violent events during and after migration		
Health behaviour	Food and drink (intake and access); smoking and vaccine status; alcohol consumption		
Healthcare	Access to healthcare professionals and medication; time since last seen by a medical doctor or registered nurse		
Mental health	PTSD; anxiety; depression; general mental health well-being	O	
Health status	Medical history; family history; current medical complaints; anthropometrics; blood pressure		Clinical interview, physical examination and biological samples
Microbiological resistance pattern	ESBL, CPO, MRSA	O	
Communicable diseases	Hepatitis B; bacteria, fungi and parasites; leishmaniases; diphtheria	O	
Non-communicable diseases	Type 2 diabetes; micronutrient status	O	

CPO, carbapenemase-producing organisms; ESBL, extended-spectrum beta-lactamase; MRSA, methicillin-resistant *Staphylococcus aureus*; O, outcome; PTSD, post-traumatic stress disorder.

is a psychopathological result of exposure to a traumatic event with core symptoms of hypervigilence, avoidance to reminders of the triggering event, altered cognition and mood, sleep disturbance and pervasive sense of immediate threat.[51] Risk factors include occupational exposure to traumatic events, female sex, childhood trauma, lower educational status and prior mental disorders. We want to investigate how migration—moving through Europe—affects PTSD prevalence in refugees. To answer this, we will compare the PTSD prevalence in Syrian asylum seekers in Denmark (exposed) with the PTSD prevalence in Syrian refugees in Lebanon (non-exposed). However, a simple comparison would be biased by confounding (eg, age) and selection (eg, survival). Our goal is expressed in the counterfactual framework,[52] to compare the PTSD prevalence in a population that *would have* migrated through Europe with the prevalence in the population that in fact did migrate through Europe—that is, two populations that are alike on all relevant variables except the exposure. This challenge is conceptually equivalent to comparing two populations with different age distributions and in this way our approach can be thought of as indirect standardisation using 'standardised mortality ratio' weights, though with complex covariate distribution that generates the weights. We have planned studies of the association of other psychiatric diseases as well as communicable and non-communicable diseases (see table 1), to give a better understanding of asylum seeker's and refugee's health.

METHODS
To study the health effects of migration and living as a refugee, we include a relatively homogenous population:

adults from the same country of origin who have migrated in the same time period and who have spent less than 12 months in the host country. This is in line with current recommendations for studies of refugees' health.[53–55]

Setting and study population
We will use the terms 'refugees' and 'asylum seekers' for participants in Lebanon and Denmark, respectively. This terminology reflects the local practice and the fact that Lebanon does not accept Syrian asylum seekers since 2015. We include participants shortly after their arrival to refugee camps in Lebanon or Danish accommodation centres.

The Lebanese settlements for Syrian refugees are to a large extent informal gatherings scattered across Lebanon. The types of shelter range from regular structures to improvised tents. In 2017, more than half of the shelters did not meet minimum humanitarian standards, with a slight deterioration in 2018.[56 57] Both the predominant shelter type and the general condition of the shelter vary according to region in Lebanon.

Danish asylum centres are under the supervision of Danish Immigration Services, while the day-to-day operation is outsourced to a number of operators.[58 59] There are different types of asylum centres in Denmark: a reception centre (the first place for an asylum seeker to stay on arriving in Denmark; here, the asylum seeker must officially apply for asylum in Denmark), accommodation centres (after applying for asylum at the reception centre, the asylum seeker is transferred to one of the accommodation centres which are scattered across Denmark. The specific centre to which a refugee is allocated is chosen at random according to vacancy), special centres (accommodation centres for unaccompanied minors and for

seriously ill asylum seekers), and departure centres (for individuals whose asylum application has been rejected). For the vast majority of asylum seekers, the larger part of their time spent in the Danish asylum system is in the accommodation centre.

Recruitment

In both Lebanon and Denmark, the inclusion criteria are (1) Syrian-born, (2) adult (>18 years old), (3) fled Syria after February 2011 (ie, after the onset of the current Syrian war), (4) arrived in the host country less than 12 months prior to inclusion and (5) resident in the settlement or centre (in Lebanon or Denmark, respectively) at the time of inclusion. We exclude individuals who are physically and/or mentally incapable of participating in the data collection. In both settings, we use a one-stage cluster sample design.[60 61]

For the inclusion in Lebanon, we divide the country into five regions: Beirut, North, South, Bekaa and Mount Lebanon. In each region, we collaborate with local operators of health services for Syrian refugees and decide on inclusion sites at random: health clinics, the settlements or at natural gatherings, although excluding sites considered unsafe at the time of inclusion. In collaboration with the local operators, we identify and invite all eligible individuals to participate. Furthermore, after inclusion, we ask each participant to spread the word (ie, network sampling).

For the inclusion in Danish asylum centres, we approach all operators, except one due to difficulties in reaching this operator's centres (on the island of Bornholm), to allow inclusion of participants at their centres. We include participants at a random sample of the accommodation asylum centres identifying and inviting all eligible residents at the chosen centres in collaboration with the asylum centre staff.

Data collection

For the inclusion, teams of two to four healthcare professionals visit the selected settlements and asylum centres in Lebanon and Denmark, respectively.

In both Lebanon and Denmark, individuals are invited to participate in the research project with oral and written information available in Arabic. At consent, the participant is handed the questionnaire and mental health scales (please see "Questionnaires" below) and instructed to complete it independently. If the participant is not able to read or write or have questions to the questionnaire or scales, one of the healthcare professionals assists the participant, with help from an interpreter if required. Afterwards, the participant is invited to a clinical interview and examination by a trained medical doctor. A venous blood sample, throat swab, and, where applicable, wound swab and biopsy are collected. Finally, the participant is provided with utensils and instructions for taking and handling three stool samples. All data are collected the same day, except for the stool samples, which are delivered within 1 week to a designated person or office (please see "Questionnaires" below and table 1 for further details). The participant is allowed to opt out for parts of the data collection and continue with the rest. If doing so, the participant is encouraged to provide a reason. Individuals who are invited but refuse to participate are asked for a brief non-structured interview on the grounds for refusal, and if we are allowed to include basic demographic data on her/him. If, during answering the questionnaire, the interview or the examination the participant withdraws his/her consent or the healthcare professional learns that the participant does not meet the inclusion criteria or meet the exclusion criteria, data collection is stopped and the collected data and biological material are destroyed.

The data collection was piloted in Denmark, including 10 participants in an accommodation centre and minor adjustments were made to the workflow accordingly.

Questionnaires

We developed a 20-item questionnaire on four constructs: the participant and his/her family, educational background, health and risk factors and migration history (see table 1 for the full set of variables collected). The questionnaire was translated by bilingual and bicultural translators: from Danish into Arabic by a translator whose mother tongue was Syrian Arabic and back translated by two translators whose mother tongue was Arabic (other dialects than Syrian). After minor corrections, each item was evaluated, and the translated questionnaire was approved in its final form by three Lebanese medical doctors (mother tongue Lebanese Arabic).

We employ a set of three validated self-rating scales on four dimensions of mental health:

WHO (Five) Well-being Index

Developed in 1998 as a condensed version of the WHO (Ten) Well-being Index, this five Likert items self-report scale is well validated for measuring subjective quality of life—with no diagnostic specificity—across settings and cultures.[62–64] The items are all positively phrased generic statements (measuring 'subjective positive well-being',[64 65]), and the respondent is asked to indicate how he or she has felt during the past 2 weeks (from 'All of the time' (item score=5) to 'At no time' (item score=0)). The item scores are summed and multiplied by 4 (to get the 'percentage score'). A cut-off percentage score ≤50 indicates a risk of poor well-being. A percentage score <13, or any one item score of 0 or 1, indicates poor well-being and in this case formal diagnostics for depression and mental health issues are recommended. We use the Arabic 1999 version[66] validated by Sibai et al.[67]

Harvard Trauma Questionnaire (HTQ)

Developed in 1992 for an Indo-Chinese refugee population,[68] the HTQ has been adapted and validated in a range of different languages and populations.[69–71] It was developed as a three-section self-report scale to assess the presence of symptoms of PTSD according to the Diagnostic

and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) criteria (part IV) and to allow a description of traumatic experiences including torture (parts I and II). A section to evaluate head injury was later added (part III).[68 69 72] We employ the initial 16 four-level Likert items of part IV corresponding to the 17 DSM-IV PTSD symptoms (giving the 'DSM-IV-score'). The respondent is asked to rate 16 statements according to how much he/she has been bothered by the symptom in question during the past week (from 'Not at all' (item score=1) to 'Extremely' (item score=4)). The mean score of the items is calculated with a recommended cut-off score ≥2.5 indicating PTSD, although debated in a cross-cultural setting.[73] We used the Arabic version developed for and validated in an Iraqi refugee population by Shoeb *et al*.[69]

Hopkins Symptom Checklist-25 (HSCL)

Developed in late 1970s for a non-refugee population from the original 90-item scale,[74 75] this 25 4-level Likert items self-report scale has been adapted and validated across multiple languages and populations, including refugees.[71 76 77] The HSCL-25 was developed to assess the presence of symptoms of anxiety (the first 10 items) and depression (the final 15 items) according to the DSM-III-R criteria. The respondent is asked to rate 25 symptoms/problems according to how much he/she has been bothered or distressed by the symptom in question during the past week (from Not at all (item score=1) to Extremely (item score=4)). Three mean scores are calculated corresponding to the items included ('total mean', 'anxiety mean' or 'depression mean'). Different cut-offs have been proposed[30] with the mean score ≥1.75 most widely used to indicate significant emotional distress (any mean) or specifically indicating anxiety or depression (anxiety mean or depression mean, respectively). The scale has been reported not to be completely robust to the setting,[78] and it has been discussed whether the proposed cut-off should be higher for Arabic speaking individuals.[71] The Arabic version used in the present study has been widely used for a number of years in clinical settings across Denmark; however, it was not possible to determine its origin.

Clinical examination

The clinical examination includes medical history, a range of measurements and collection of specimens. See table 1 for the full list of variables.

Specimens

Dried blood spot

We collect dried blood spot (DBS) specimens by applying freshly drawn venous blood to two filter paper cards (Whatman 903 Protein Saver(c)) per participant. The DBS cards are handled according to WHO guidelines.[79] In short, the cards are allowed to dry in dry racks, packed in a gas-impermeable plastic ziplock bag with four 1 g desiccant packs and one humidity indicator card (Multisorb, Humonitor(c)), and stored at −80°C. We check the humidity level in the bags regularly and change the desiccant packs and humidity indicator at signs of humidity. All cards are signed with participant identification number and date of collection.

Throat swabs

We collect cultures from both tonsils and any oral lesions using the BD ESwab (TM) collection kit regular flocked swab system for detection of methicillin-resistant *Staphylococcus aureus* (MRSA) and *Corynebacterium diphtheriae*. In Lebanon, the labelled collection tubes are stored at +5°C until being permanently stored at −20°C with added 20% glycerol (same day).[80 81] The specimens are shipped by carrier to Denmark and further stored at −20°C before final shipment by regular postal service to the laboratory for analysis. In Denmark, the specimens are stored at +5°C immediately and shipped the following day by regular postal service to the laboratory for analysis.

Wound swabs

If a participant has an infected or chronic wound, we collect wound swab using the BD ESwab (TM) collection kit regular flocked swab system for detection of MRSA and *C. diphtheriae*. The same procedures for storage and shipment as described above are implemented.

Stool samples

We collect stool by asking the participants to deliver samples from three different defecations. The participants are provided with three labelled specimen containers with three matching transportation vials. We instruct the participant on how to collect the sample and the amount necessary for each sample. The participants are given 5 days to deliver the samples that are stored in a refrigerator (+5°C). On collection, the samples are packaged and shipped by regular postal service to the laboratory for detection of pathogens and resistant bacteria. Briefely, the samples are analysed for ova and parasites by microscopy of concentrated stool. Detection of extended-spectrum beta-lactamase (ESBL)-producing *Enterobacterales* and carbapenemase-producing organisms (CPE) are done by plating 10 μL of medium on a chromogenic culture media (chromID ESBL; Biomerieux, France) for detection of ESBL and 1 μL on SSI enteric medium (SSI diagnostics, Denmark) with a meropenem Neo-Sensitabs (Rosco, Denmark) for detection of CPE (using a break-point of 27/27 mm) and incubated aerobically at 36°C. The plates are examined after 18–24 hours of incubation. Confirmation of ESBL production is performed by the ESBL +AmpC Screen Kit (Rosco, Denmark). Confirmation of CPE is performed by the ROSCO Rapid CARB Screen kit (Rosco, Denmark) according to the European Committee on Antimicrobial Susceptibility Testing standards. Identification of bacteria is done by use of Matrix-assisted laser desorption-ionisation (MALDI-TOF; Bruker, Germany). Moreover, we also perform microbiome analysis on stool samples by DNA extraction with

subsequent profiling of 16S and 18S (nuclear ribosomal genes of bacteria, fungi and parasites) as described elsewhere.[82]

Statistics

Power and sample size

We aim at including 1100 participants in Lebanon and 220 participants in Denmark (corresponding to a ratio 5:1 to represent the 5 regions in Lebanon) to be able to show a statistically significant difference between a prevalence of 5% and 12% (representing the condition being 'rare' and 'not uncommon', respectively) with a 5% statistical significance level and 80% power in an independent two-group design without continuity correction.

In Lebanon, we estimate the source population using United Nations High Commissioner for Refugees data.[83] To secure proportional allocation, we aim to include the following number of participants in each region: South, n=133, North, n=286, Beirut, n=100, Mount Lebanon, n=186, Bekaa valley, n=395.[83 84]

In Denmark, we estimate the source population from the official number of asylum applicants from Syria during the inclusion period and 1 month prior.[85] We estimate to be able to reach an average of 40 participants at each accommodation centre thus requiring visiting six different centres. Since asylum seekers are allocated to the accommodation centres at random (see the Setting section), we have not powered our study to investigate any cluster effect.

Planned analysis

To obtain an unbiased estimate of the association between exposure (eg, long-distance migration, see figure 1, scenario b and c) and outcome (eg, PTSD, see table 1) in a population with risk factor distribution that resembles that of the exposed, we will compute propensity score weights, based on confounders and potential confounders informed by directed acyclic graph,[86] to balance out the differences between the populations (ie, by a 'standard-mortality-ratio-weighted' marginal structural model).[87–89] Bias and missing data patterns will be analysed, and multiple imputation will be performed to allow for analysis of the full data set and report non-parametric bias corrected and accelerated bootstrapped CI for maximum validity.[90–96] See table 1 for the full list of outcomes and covariates.

Study data are collected and stored using REDCap electronic data capture tool hosted at Aarhus University.[97] Data management, analysis and plots will be carried out using R.[98]

Patient and public involvement

It is not possible to involve the participants in the design, conduct or dissemination of this research project. The chief investigator presents the project to collaborators at all study sites and all participants that are interested beyond the written and oral description at inclusion.

CURRENT STATUS

The Asylum seekers' and Refugees' Changing Health (ARCH) study is currently enrolling participants. As of April 2019, we have enrolled 113 participants in Denmark and 179 participants in Lebanon. Inclusion will continue through 2019 or until we have reached our target of 1100 participants in Lebanon and 220 participants in Denmark.

DISCUSSION

There is no indication that the future will bring less migration between countries and regions of the world. Thus far, the health of migrants has mainly been studied in an undefined migrant population without accounting for the heterogeneity on all relevant levels (eg, country of origin (including disease epidemiology, healthcare system and vaccine coverage), ethnic variation and related genetic disposition, reason for migration, means of migration, immigrant status in host country) leading to conflicting results and stifling clinical and academic advances. There is an increasing focus on this lack of specificity by design and some of the more recent studies have managed to specify the study population resulting in interpretable results and meaningful conclusions.[35 99 100] The vast amount of studies published on migrant health—not to mention those that actually specify a study population such as refugees—report comparisons to the autochthonous population in the host country (figure 1, scenario a). This allows for a limited number of research questions leaving more fundamental questions unanswerable (such as the healthy migrant effect and the health effects of migration, see figure 1). Advances in epidemiological methods to analyse health outcomes have the potential to inform both design and analysis of difficult to reach populations, such as many subpopulations of migrants, but they have yet to make their way into the majority of migrant health research. While we do not claim to have fully accomplished this, the ARCH study presents a comparison of a carefully specified study population in two different host countries with very different migration histories. In studying the health changes that happen during the actual migration, we are limited by the fact that we cannot provide a truly longitudinal design (not to speak of a trial), but instead are limited to a cross-sectional comparison, hindering a causal interpretation of this part of the ARCH study. By implementing a counterfactual framework—computing the average effect in the treated (ie, migrated)—we do, however, expect to be able to give a much more precise estimate of the association between migration and health status to inform clinicians and future research. Despite our efforts to handle biases, we recognise that the results of the ARCH study are vulnerable to this: for example, the cold chain and transportation of the collected specimens have to be adapted to the two different settings as described in the methods section. If this affects the measurement error of the lab results, it creates a differential measurement error of these outcomes. The information bias could potentially

be dependent (eg, if individuals suffering from PTSD consistently show higher degree of recall bias). Selection bias and missing data may arise, for example, due to self-selection into the study or if we have misspecified the sampling frame. We consider the risk of the latter very low in Denmark though not unlikely in Lebanon where the refugees regularly move around and there is little registration of the individuals. We plan to do quantitative bias analysis[101] to qualify the discussion of biases in our results.

Ethics and dissemination

All necessary permissions have been obtained prior to the beginning of inclusion; this includes from the Lebanese Ministry of Health (archives number 2018/4/38918), Lebanese University, Faculty of Medical Sciences (journal number D 144/232018), Mount Lebanon Hospital's Ethics' Committee (MLH code PSY-2018-005, PSY-2018-006 and PSY-2018-007) and the Danish Data Protection Agency (journal number 2015-41-4500).

The inclusion, analysis and reporting comply with International ethical guidelines for health-related research involving humans and the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.[102 103] The chief investigator (AHE) is formally trained in Responsible Conduct of Research and Research Integrity for Biomedical Sciences. All persons involved in data collection, handling and analysis have been thoroughly trained for their assignment. All biological samples are collected and handled according to the best practice guidelines. The R code and codebook for published results will be made freely available from the first authors GitHub account (https://github.com/eiset). Deidentified participant data will be available on reasonable request to the first author to the degree that this is preserved according to national and international law. A number of publications are planned on the results of the ARCH study in both academic and popular scientific journals: in short term, we will present the results of cross-sectional comparisons of the throat and wound swab and mental health screenings. In the longer term, we are planning to follow the refugees in Denmark with collection of DBSs, mental health screenings and semistructured qualitative interviews on the participant's health and access to healthcare in the time lived in Denmark.

Author affiliations
[1]Center for Global Health (GloHAU), Department of Public Health, Aarhus University, Aarhus, Denmark
[2]Clinic for PTSD and Anxiety, Aarhus University Hospital, Aarhus, Denmark
[3]Department of Psychiatry, Lebanese University, Beirut, Lebanon
[4]Department of Bacteria, Parasites and Fungi, Statens Serum Institut, Kobenhavn, Denmark
[5]Department of Public Health, Aarhus University, Aarhus, Denmark
[6]Department of Infectious Diseases, Aarhus University Hospital, Aarhus, Denmark

ORCID iD
Andreas Halgreen Eiset http://orcid.org/0000-0002-5487-8359

REFERENCES

1 The World Bank. World development indicators, DataBank. Available: https://databank.worldbank.org/data/reports.aspx?source=2&series=SM.POP.REFG&country=EUU# [Accessed 28 Dec 2019].
2 UNHCR. Global trends forced displacement in 2018, 2019. Available: https://www.unhcr.org/statistics/unhcrstats/5d08d7ee7/unhcr-global-trends-2018.html [Accessed 28 Dec 2019].
3 UNHCR popstats. UNHCR statistics - the world in numbers. Available: http://popstats.unhcr.org [Accessed 13 Feb 2019].
4 UNHCR. Global trends forced displacement in 2018 Annex tables 1-29, 2019. Available: https://www.unhcr.org/statistics/unhcrstats/5d08d7ee7/unhcr-global-trends-2018.html [Accessed 28 Dec 2019].
5 Eiset AH, Wejse C. Review of infectious diseases in refugees and asylum seekers-current status and going forward. *Public Health Rev* 2017;38:22.
6 Spallek J, Zeeb H, Razum O. What do we have to know from migrants' past exposures to understand their health status? A life course approach. *Emerg Themes Epidemiol* 2011;8:6.
7 Aldridge RW, Nellums LB, Bartlett S, *et al*. Global patterns of mortality in international migrants: a systematic review and meta-analysis. *Lancet* 2018;392:2553–66.
8 Cochrane A, Evlampidou I, Irish C, *et al*. Hepatitis B infection prevalence by country of birth in migrant populations in a large UK City. *J Clin Virol* 2015;68:79–82.
9 Deboosere P, Gadeyne S. Adult migrant mortality advantage in Belgium: evidence using census and register data. *Popul Engl Ed* 2005;60:655–98.
10 Fuller-Thomson E, Brennenstuhl S, Cooper R, *et al*. An investigation of the healthy migrant hypothesis: Pre-emigration characteristics of those in the British 1946 birth cohort study. *Can J Public Health* 2016;106:e502–8.
11 Greenaway C, Thu Ma A, Kloda LA, *et al*. The seroprevalence of hepatitis C antibodies in immigrants and refugees from intermediate and high endemic countries: a systematic review and meta-analysis. *PLoS One* 2015;10:e0141715.
12 Greve B. Migrants and health in the Nordic welfare states. *Public Health Rev* 2016;37:9.
13 Hamilton TG. The healthy immigrant (migrant) effect: in search of a better native-born comparison group. *Soc Sci Res* 2015;54:353–65.
14 Helgesson M, Johansson B, Nordquist T, *et al*. Healthy migrant effect in the Swedish context: a register-based, longitudinal cohort study. *BMJ Open* 2019;9:e026972.
15 Heudtlass P, Speybroeck N, Guha-Sapir D. Excess mortality in refugees, internally displaced persons and resident populations in complex humanitarian emergencies (1998-2012) - insights from operational data. *Confl Health* 2016;10:15.
16 Kearns A, Whitley E, Egan M, *et al*. Healthy migrants in an unhealthy City? the effects of time on the health of migrants living in deprived areas of Glasgow. *J Int Migr Integr* 2017;18:675–98.

17 DesMeules M, Gold J, McDermott S, *et al*. Disparities in mortality patterns among Canadian immigrants and refugees, 1980-1998: results of a national cohort study. *J Immigr Health* 2005;7:221–32.

18 Swerdlow Aj. Mortality and cancer incidence in Vietnamese refugees in England and Wales: a follow-up study. *Int J Epidemiol* 1991;20:13–19.

19 Marmot MG, Adelstein AM, Bulusu L. Lessons from the study of immigrant mortality. *Lancet* 1984;1:1455–7.

20 Acevedo-Garcia D, Sanchez-Vaznaugh EV, Viruell-Fuentes EA, *et al*. Integrating social epidemiology into immigrant health research: a cross-national framework. *Soc Sci Med* 2012;75:2060–8.

21 Aldridge RW, Yates TA, Zenner D, *et al*. Pre-entry screening programmes for tuberculosis in migrants to low-incidence countries: a systematic review and meta-analysis. *Lancet Infect Dis* 2014;14:1240–9.

22 Carlsson AC, Wändell PE, Hedlund E, *et al*. Country of birth-specific and gender differences in prevalence of diabetes in Sweden. *Diabetes Res Clin Pract* 2013;100:404–8.

23 Doocy S, Lyles E, Hanquart B, *et al*. Prevalence, care-seeking, and health service utilization for non-communicable diseases among Syrian refugees and host communities in Lebanon. *Confl Health* 2016;10:21.

24 Georgiadou E, Zbidat A, Schmitt GM, *et al*. Prevalence of mental distress among Syrian refugees with residence permission in Germany: a registry-based study. *Front Psychiatry* 2018;9:393.

25 Gerritsen AAM, Bramsen I, Devillé W, *et al*. Physical and mental health of Afghan, Iranian and Somali asylum seekers and refugees living in the Netherlands. *Soc Psychiatry Psychiatr Epidemiol* 2006;41:18–26.

26 Greenaway C, Castelli F. Infectious diseases at different stages of migration: an expert review. *J Travel Med* 2019;26:taz007.

27 Hooi JKY, Lai WY, Ng WK, *et al*. Global prevalence of Helicobacter pylori infection: systematic review and meta-analysis. *Gastroenterology* 2017;153:420–9.

28 Isenring E, Fehr J, Gültekin N, *et al*. Infectious disease profiles of Syrian and Eritrean migrants presenting in Europe: a systematic review. *Travel Med Infect Dis* 2018;25:65–76.

29 Modesti PA, Reboldi G, Cappuccio FP, *et al*. Panethnic differences in blood pressure in Europe: a systematic review and meta-analysis. *PLoS One* 2016;11:e0147601.

30 M'zah S, Lopes Cardozo B, Evans DP. Mental health status and service assessment for adult Syrian refugees resettled in metropolitan Atlanta: a cross-sectional survey. *J Immigr Minor Health* 2019;21:1019-1025.

31 van Oostrum IEA, Goosen S, Uitenbroek DG, *et al*. Mortality and causes of death among asylum seekers in the Netherlands, 2002-2005. *J Epidemiol Community Health* 2011;65:376–83.

32 Pfortmueller CA, Schwetlick M, Mueller T, *et al*. Adult asylum seekers from the middle East including Syria in central Europe: what are their health care problems? *PLoS One* 2016;11:e0148196.

33 Abdul Rahim NR, Benson J, Grocke K, *et al*. Prevalence of Helicobacter pylori infection in newly arrived refugees attending the migrant health service, South Australia. *Helicobacter* 2017;22:e12360.

34 Rehr M, Shoaib M, Ellithy S, *et al*. Prevalence of non-communicable diseases and access to care among non-cAMP Syrian refugees in northern Jordan. *Confl Health* 2018;12:33.

35 Tinghög P, Malm A, Arwidson C, *et al*. Prevalence of mental ill health, traumas and postmigration stress among refugees from Syria resettled in Sweden after 2011: a population-based survey. *BMJ Open* 2017;7:e018899.

36 Wagner J, Berthold SM, Buckley T, *et al*. Diabetes among refugee populations: what newly arriving refugees can learn from resettled Cambodians. *Curr Diab Rep* 2015;15:1–14.

37 Castelli F. Drivers of migration: why do people move? *J Travel Med* 2018;25.

38 Razum O, Twardella D. Time travel with Oliver Twist--towards an explanation foa a paradoxically low mortality among recent immigrants. *Trop Med Int Health* 2002;7:4–10.

39 Wallace M, Kulu H. Low immigrant mortality in England and Wales: a data artefact? *Soc Sci Med* 2014;120:100–9.

40 Krasnik A. Categorizations of migrants and ethnic minorities--are they useful for decisions on public health interventions? *Eur J Public Health* 2015;25:907.

41 Zimmerman C, Kiss L, Hossain M. Migration and health: a framework for 21st century policy-making. *PLoS Med* 2011;8:e1001034.

42 Ager A, Burnham G, Checchi F, *et al*. Strengthening the evidence base for health programming in humanitarian crises. *Science* 2014;345:1290–2.

43 Alwin DF. Commentary: it takes more than one to tango: life course epidemiology and related approaches. *Int J Epidemiol* 2016;45:988–93.

44 Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? *Behav Brain Sci* 2010;33:61–83.

45 Hernán MA. Beyond exchangeability: the other conditions for causal inference in medical research. *Stat Methods Med Res* 2012;21:3–5.

46 Ioannidis JPA, Greenland S, Hlatky MA, *et al*. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383:166–75.

47 Keiding N, Clayton D. Standardization and control for confounding in observational studies: a historical perspective. *Statist Sci* 2014;29:529–58.

48 Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. *J R Stat Soc Ser A Stat Soc* 2016;179:319–76.

49 Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model* 1986;7:1393–512.

50 Rubin DB. Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Serv Outcomes Res Methodol* 2001;2:169–88.

51 Shalev A, Liberzon I, Marmar C. Post-Traumatic stress disorder. *N Engl J Med* 2017;376:2459–69.

52 Pearl J. Causal inference in the health sciences: a conceptual introduction. *Health Serv Outcomes Res Methodol* 2001;2:189–220.

53 Bradby H, Humphris R, Newall D, *et al*. *Public health aspects of migrant health a review of the evidence on health status for refugees and asylum seekers in the European region*. World Health Organization, 2015. http://www.euro.who.int/en/publications/abstracts/public-health-aspects-of-migrant-health-a-review-of-the-evidence-on-health-status-for-refugees-and-asylum-seekers-in-the-european-region-2015

54 EU and UN Expert Group on refugee and Internally Displaced Persons Statistics (EGRIS). *International recommendations on refugee statistics*. Luxembourg: European Union and the United Nations, 2018.

55 World Health Organization. *Health of migrants: the way forward - Report of a global consultation*. Madrid, Spain: World Health Organization, 2010. http://apps.who.int/iris/handle/10665/44336

56 United Nations Children's Fund (UNICEF), United Nations High Commissioner for Refugees (UNHCR), United Nations World Food Programme (WFP). Vulnerability assessment for Syrian refugees in Lebanon (VASyR-2017), 2017. Available: https://www.unhcr.org/lb/wp-content/uploads/sites/16/2018/01/VASyR-2017.pdf [Accessed 25 Feb 2019].

57 United Nations Children's Fund (UNICEF), United Nations High Commissioner for Refugees (UNHCR), United Nations World Food Programme (WFP). Vulnerability assessment for Syrian refugees in Lebanon (VASyR-2018), 2018. Available: https://www.unhcr.org/lb/wp-content/uploads/sites/16/2018/01/VASyR-2017.pdf [Accessed 25 Feb 2019].

58 Bendixen MC. Asylcentrene. Refug. Asylcentrene. Available: http://refugees.dk/fakta/asylproceduren-i-danmark/asylcentrene/ [Accessed 27 Feb 2019].

59 Ny i danmark. Asylcentre. Available: http://www.nyidanmark.dk/da/Waiting/Asylum/Housing of asylum applicants/Asylum centre [Accessed 27 Feb 2019].

60 Sedgwick P. Cluster sampling. *BMJ* 2014;348:g1215.

61 Lohr SL. *Sampling: design and analysis*. 2 edn. Brooks/Cole: Cengage Learning, 2010. https://www.sharonlohr.com/sampling-design-and-analysis

62 Kvorning J. About the WHO-5. Available: https://www.psykiatri-regionh.dk/who-5/about-the-who-5/Pages/default.aspx [Accessed 15 Oct 2018].

63 Bech P, Gudex C, Johansen KS. The WHO (ten) well-being index: validation in diabetes. *Psychother Psychosom* 1996;65:183–90.

64 Topp CW, Østergaard SD, Søndergaard S, *et al*. The WHO-5 well-being index: a systematic review of the literature. *Psychother Psychosom* 2015;84:167–76.

65 Bech P. Subjective positive well-being. *World Psychiatry* 2012;11:105–6.

66 World Health Organisation. WHO-5 questionnaires. Available: https://www.psykiatri-regionh.dk/who-5/who-5-questionnaires/Pages/default.aspx [Accessed 15 Oct 2018].

67 Sibai AM, Chaaya M, Tohme RA, *et al*. Validation of the Arabic version of the 5-item who well being index in elderly population. *Int J Geriatr Psychiatry* 2009;24:106–7.

68 Mollica RF, Caspi-Yavin Y, Bollini P, *et al*. The Harvard trauma questionnaire. validating a cross-cultural instrument for measuring

torture, trauma, and posttraumatic stress disorder in Indochinese refugees. *J Nerv Ment Dis* 1992;180:111–6.

69 Shoeb M, Weinstein H, Mollica R. The Harvard trauma questionnaire: adapting a cross-cultural instrument for measuring torture, trauma and posttraumatic stress disorder in Iraqi refugees. *Int J Soc Psychiatry* 2007;53:447–63.

70 Vindbjerg E, Carlsson J, Mortensen EL, *et al*. The latent structure of post-traumatic stress disorder among Arabic-speaking refugees receiving psychiatric treatment in Denmark. *BMC Psychiatry* 2016;16:309.

71 Wind TR, van der Aa N, Knipscheer J, de la Rie S, *et al*. The assessment of psychopathology among traumatized refugees: measurement invariance of the Harvard trauma questionnaire and the Hopkins symptom Checklist-25 across five linguistic groups. *Eur J Psychotraumatol* 2017;8:1321357.

72 Sigvardsdotter E, Malm A, Tinghög P, *et al*. Refugee trauma measurement: a review of existing checklists. *Public Health Rev* 2016;37:10.

73 Ichikawa M, Nakahara S, Wakai S. Cross-Cultural use of the predetermined scale cutoff points in refugee mental health research. *Soc Psychiatry Psychiatr Epidemiol* 2006;41:248–50.

74 Derogatis LR, Lipman RS, Rickels K, *et al*. The Hopkins symptom checklist (HSCL): a self-report symptom inventory. *Behav Sci* 1974;19:1–15.

75 Winokur A, Winokur DF, Rickels K, *et al*. Symptoms of emotional distress in a family planning service: stability over a four-week period. *Br J Psychiatry* 1984;144:395–9.

76 Hollifield M, Warner TD, Lian N, *et al*. Measuring trauma and health status in refugees: a critical review. *JAMA* 2002;288:611–21.

77 Mollica RF, Wyshak G, de Marneffe D, *et al*. Indochinese versions of the Hopkins symptom Checklist-25: a screening instrument for the psychiatric care of refugees. *Am J Psychiatry* 1987;144:497–500.

78 Ventevogel P, De Vries G, Scholte WF, *et al*. Properties of the Hopkins symptom Checklist-25 (HSCL-25) and the self-reporting questionnaire (SRQ-20) as screening instruments used in primary care in Afghanistan. *Soc Psychiatry Psychiatr Epidemiol* 2007;42:328–35.

79 WHO HIV ResNet Laboratory Working Group. *WHO manual for HIV drug resistance testing using dried blood spot specimens*. Washington, DC, USA: WHO, 2010.

80 Botrugno R, Conter L, Massetti B, *et al*. A New Amies Medium +20% Glycerol Device for the Collection and Long Term Frozen Storage of Clinical Specimens for Bacteria Culture, 2016. Available: http://files.abstractsonline.com/SUPT/1/4060/AllPosterAbstracts.pdf

81 Iaffaldano L, Granata I, Pagliuca C, *et al*. Oropharyngeal microbiome evaluation highlights Neisseria abundance in active celiac patients. *Sci Rep* 2018;8:11047.

82 Ring HC, Thorsen J, Saunte DM, *et al*. The follicular skin microbiome in patients with hidradenitis suppurativa and healthy controls. *JAMA Dermatol* 2017;153:897–905.

83 UNHCR, Operational Portal, Refugee Situations. Situation Syria regional refugee response. Available: https://data2.unhcr.org/en/situations/syria/location/71 [Accessed 20 Feb 2019].

84 UNHCR, Operational Portal, Refugee Situations. Syrian refugees registered by Cadastral. Available: https://data2.unhcr.org/en/documents/details/69594 [Accessed 6 Sep 2019].

85 Udlændingestyrelsen. *Styrelsen for International Rekruttering og Integration. Tal og fakta på udlændingeområdet 2012-2016 [Figurs & facts about immigration 2012-2016*. Udlændingestyrelsen [Danish Immigration Service], 2017. http://uim.dk/publikationer/publikationer?searchable=tal+og+fakta

86 Dawid AP. Influence diagrams for causal modelling and inference. *Int Stat Rev* 2002;70:161–89.

87 Kurth T, Walker AM, Glynn RJ, *et al*. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *Am J Epidemiol* 2006;163:262–70.

88 Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550–60.

89 Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology* 2003;14:680–6.

90 Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Stat Med* 2000;19:1141–64.

91 Efron B. Bootstrap methods: another look at the jackknife. *Ann Statist* 1979;7:1–26.

92 Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 2003;14:300–6.

93 Greenland S. Multiple-bias modelling for analysis of observational data (with discussion). *J R Stat Soc Ser A Stat Soc* 2005;168:267–306.

94 Greenland S. Bayesian perspectives for epidemiologic research: III. bias analysis via missing-data methods. *Int J Epidemiol* 2009;38:1662–73.

95 Harrell F. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. 2 edn. Springer International Publishing, 2015. https://www.springer.com/la/book/9783319194240

96 Rubin DB. Inference and missing data. *Biometrika* 1976;63:581–92.

97 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.

98 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2016. http://www.R-project.org/

99 Vangay P, Johnson AJ, Ward TL, *et al*. US immigration Westernizes the human gut microbiome. *Cell* 2018;175:e10:962–72.

100 Verstraeten SPA, van den Brink CL, Mackenbach JP, *et al*. The health of Antillean migrants in the Netherlands: a comparison with the health of non-migrants in both the countries of origin and destination. *Int Health* 2018;10:258–67.

101 Lash TL, Fox MP, MacLehose RF, *et al*. Good practices for quantitative bias analysis. *Int J Epidemiol* 2014;43:1969–85.

102 Kottow Lang MH. Critical analysis of the Council for international organizations of medical sciences 2016 international guidelines for health-related research involving humans. *Medwave* 2017;17:e6956.

103 Vandenbroucke JP, von Elm E, Altman DG, *et al*. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *Epidemiology* 2007;18:805–35.