BMJ Open

The population-based Occupational and Environmental Health Prospective Cohort Study (AMIGO) in the Netherlands

Pauline Slottje,[1] C Joris Yzermans,[2] Joke C Korevaar,[2] Mariëtte Hooiveld,[2] Roel C H Vermeulen[1]

[1]Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands
[2]Netherlands Institute for Health Services Research (NIVEL), Utrecht, The Netherlands

Correspondence to
Roel C H Vermeulen;
r.c.h.vermeulen@uu.nl

ABSTRACT

Purpose: Occupational and environmental exposures remain important modifiable risk factors of public health. Existing cohort studies are often limited by the level of detail of data collected on these factors and health. It is also often assumed that the more healthy group is over-represented in cohort studies, which is of concern for their external validity. In this cohort profile, we describe how we set up the population-based Occupational and Environmental Health Cohort Study (AMIGO) to longitudinally study occupational and environmental determinants of diseases and well-being from a multidisciplinary and life course point of view. Reviewed by the Medical Ethics Research Committee of the University Medical Center Utrecht (protocol 10-268/C). All cohort members participate voluntarily and gave informed consent prior to their inclusion.

Participants: 14 829 adult cohort members (16% of those invited) consented and filled in the online baseline questionnaire. Determinants include chemical, biological, physical (eg, electromagnetic fields), and psychosocial factors. Priority health outcomes include cancer, neurological, cardiovascular and respiratory diseases and non-specific symptoms. Owing to the recruitment strategy via general practitioners of an established network, we also collect longitudinal data registered in their electronic medical records including symptoms, diagnosis and treatments. Besides the advantage of health outcomes that cannot be easily captured longitudinally by other means, this created a unique opportunity to assess health-related participation bias by comparing general practitioner-registered prevalence rates in the cohort and its source population.

Findings to date: We found no indications of such a systematic bias. The major assets of the AMIGO approach are its detailed occupational and environmental determinants in combination with the longitudinal health data registered in general practice besides linkage to cancer and mortality registries and self-reported health.

Future plans: We are now in the phase of prospective follow-up, with the aim of continuing this for as long as possible (20+ years), pending future funding.

Strengths and limitations of this study

- The major assets of the Occupational and Environmental Health Cohort Study (AMIGO) are its detailed occupational and environmental determinants in combination with the unique longitudinal health data registered in general practice besides the more usual linkages to routine cancer and mortality registries and self-reported health.
- The main limitations of the study are those known to cohort studies, in particular selection bias and loss to follow-up.
- We found no evidence for systematic health-related participation bias at study entry, based on a unique comparison of general-practitioner recorded prevalence rates of various disorders.

Findings will be disseminated through scientific conferences and peer-reviewed journals, and through newsletters and the project website to participants, stakeholders and the wider public.

INTRODUCTION

Occupational and environmental health remains important in epidemiology and public health. According to the WHO, an estimated 24% of the global disease burden (healthy life years lost) and 23% of all deaths (premature mortality) can be attributed to modifiable environmental factors, such as occupational risks, air pollution, electromagnetic fields, built environments and agricultural methods.[1] For example, for priority disease outcomes included in the present study, the WHO's global estimates of the attributable environmental fractions were 44% of asthma development and exacerbation, 42% of chronic obstructive pulmonary diseases (COPD), 19% of cancer, 16% of

cardiovascular diseases, and 13% of neuropsychiatric disorders such as Alzheimer's and Parkinson's diseases, multiple sclerosis, insomnia and migraine.

In general and to enhance this field of research, prospective epidemiological approaches are favoured for making aetiological inferences. With respect to cohort studies, occupational and environmental health has mostly been studied in (retrospective) industry-based cohort studies or specific occupational cohorts (eg, nurses) on targeted occupational exposures, or as an add-on in community-based cohort studies that originally had another focus, for example, on diet and cancer. Although such studies can be informative, they are often limited by the detail collected on occupational and environmental risk factors.[2] Therefore, we set up the population-based Occupational and Environmental Health Cohort Study (in Dutch: Arbeid, Milieu en Gezondheid Onderzoek, AMIGO), with a strong focus on occupational and environmental health from a multidisciplinary and life course point of view. Hence, we set out to assess lifetime biological, chemical and physical determinants in the home and work environment, as well as psychological and socioeconomic determinants. One of the initial research questions addressed in AMIGO concerns the health effects of exposure to electromagnetic fields, including mobile phone use as part of both a collaborative effort of multiple cohort studies in the Netherlands (pragmatically called the 'Pooled Cohort Study') and the international COSMOS study.[3]

A further challenge in occupational and environmental epidemiology is that certain priority health outcomes cannot easily be captured longitudinally, for example, Alzheimer's and Parkinson's diseases. Owing to the recruitment strategy of the AMIGO study, we are able to follow up many such health outcomes through general practitioner-recorded diagnoses, medication prescriptions and referrals. The primary health outcomes studied in AMIGO are cancer, neurological, respiratory, psychological and cardiovascular diseases, and non-specific symptoms.

In this cohort profile, we aim to describe our study protocol, present its first results and ongoing data collection, and discuss the methodological issue of potential participation bias at study entry, which is one of the concerns for the external validity of cohort studies, for example, it is often assumed that the more healthy group is over-represented. We addressed this issue by comparing the cohort members with the source population based on the general practitioner-recorded prevalence rates of various disorders. The findings of the AMIGO study will be disseminated through scientific conferences and peer-reviewed journals, and subsequently through, for example, newsletters and summaries on the project website to participants, stakeholders (eg, general practitioners, policymakers) and the wider public.

COHORT DESCRIPTION

We aimed to sample the general adult population of the Netherlands and decided to select 31–65-year-olds for various reasons, for example, working age as occupational exposure is a main determinant, and the age at onset of our main health outpoints. Our recruitment strategy was to invite subjects through 99 general practices that are part of a nationwide information and surveillance network for primary healthcare established at the Netherlands Institute for Health Services Research (NIVEL), that is, the NIVEL Primary Care Database.[4] In the Netherlands, it is compulsory to be enlisted at one particular general practice, and virtually all non-institutionalised citizens are. Since general practitioners are the first professionals to contact for health problems and they act as gatekeepers for secondary healthcare, the general practitioners have a rather complete picture of the health of those enlisted in their practice, including the healthy ones. Owing to this sampling strategy, we created the unique possibility to longitudinally study recorded health and primary healthcare use in association with determinants of the cohort members, as long as they are registered at a participating general practice. The NIVEL Primary Care Database includes an anonymised extract of the electronic medical records (EMRs) of the patients enlisted in the participating general practices. In these EMRs, the general practitioners routinely use the International Classification of Primary Care-1 (ICPC) to register their patients' health problems in term of symptoms and diagnoses.[5] The ICPC is an internationally endorsed classification system, which is compatible with the International Classification of Diseases-10 (ICD-10).[6] Prescriptions are registered according to the Anatomical Therapeutic Chemical (ATC) classification system.[7]

From the source population, that is, all 31–65-year-old subjects enlisted in one of the participating general practices, at random one adult per address in the Netherlands was selected to avoid clustering of participants within households. The AMIGO privacy code was established to protect the privacy of general practitioners and their clients. Selected subjects were invited by letter on behalf of their general practitioner in two waves in the period April 2011–July 2012. The letter was accompanied by an information brochure from the researchers to explain the study in detail. It was explicitly mentioned that participation is voluntary. All study material was in Dutch. Participation status was not reported back to the general practitioners. Prior to sending the first invitation, the general practitioners checked their list of the selected subjects to exclude recent deaths or other major objections to invite someone, for example, terminally ill or illiterate patients. After about 2 weeks, a second letter was sent to all invited subjects, which was phrased in such a way that it thanked respondents and reminded non-respondents that they could still participate until a certain date. In 2011, a second reminder letter was sent to non-respondents another 2 weeks later. In each letter, subjects received their unique participant code and password to login and get online access to the secured site run by Utrecht University, on which they

first filled in an online informed consent form and then the baseline online questionnaire. All data will be kept securely and participant confidentiality will be maintained. On registration, cohort members entered the AMIGO participant registry, in which their personal identification data from the informed consent form is kept by dedicated data managers at the Utrecht University strictly separated from the research data that are coded based on a unique participant code.

The informed consent covers prospective linkage to registries to obtain follow-up data on their addresses and vital status (Municipal Personal Records Database), and health outcomes including causes of death (Statistics Netherlands), cancer incidence (National Cancer Registry), and hospital discharge diagnoses. In addition, linkage to the EMRs of the participants' general practice is possible by using the key between the cohort participant number and the patient number that is used in the NIVEL Primary Care Database at NIVEL, that is, in such a way that the researchers cannot trace their identity. Approval of the study was given by the Institutional Research Board of Institute for Risk Assessment Sciences (IRAS) and NIVEL.

Follow-up of the cohort consists of questionnaires and linkages to the aforementioned data sources and might include health checks (eg, spirometry, blood pressure) and biological sampling depending on future funding. The follow-up questionnaires will be targeted at different occupational and environmental exposures as well as additional lifestyle characteristics such as physical exercise and nutrition.

The baseline questionnaire had two subsequent sections. The first part covered a broad range of determinants and lifestyle characteristics important for occupational and environmental health research such as sociodemographics; level of education; full occupational history (all jobs performed for at least 6 months, including job title, type of company, average number of hours per week, and screening questions on occupational exposures, eg, electromagnetic fields and shift work); current job work satisfaction and job stress (eg, decision authority, psychological, emotional and cognitive job demands and coworker support, based on the job content questionnaire);[8] full residential history and bedroom floor; current home characteristics relevant for, for example, indoor air quality, outdoor noise, air pollution and electromagnetic fields exposure; mobile and cordless phone use and WiFi use conform to the international COSMOS study;[3] smoking and alcohol use. On the basis of the full residential and job histories, lifetime time-varying exposure assessment will be performed. For example, by means of geospatial environmental exposure modelling (eg, air pollution or radiofrequency electromagnetic fields from mobile phone base stations) linked to the geocoded residential histories, and by means of job exposure matrices (eg, on various chemical and physical exposures) linked to coded job histories according to the International

Classification of Occupations (ISCO).[9] The screening questions across the job history on electromagnetic field exposures included whether participants ever worked with or near certain exposure sources, for example, electrical welding, antitheft devices in shops. The shift work screening questions addressed whether participants ever worked in shifts other than daytime shifts (eg, night shift, evening shift, etc), and, if so, when they started and stopped, and for night shift, the average number of nights per months in that period.

The second part of the baseline questionnaire addressed self-reported health, including general health, headache (Headache impact test,[10] ID Migraine[11]), sleep (Medical Outcomes Study (MOS)-Sleep[12]), memory problems, hearing problems, tinnitus, early Parkinson symptoms,[13] somatisation symptoms based on the Four-Dimensional Symptom Questionnaire,[14] respiratory symptoms based on the European Community Respiratory Health Survey (ECRHS[15][16]); doctor-diagnosed diseases and age at diagnosis, including diabetes, cardiovascular, neurological, pulmonary, gastrointestinal, musculoskeletal diseases; family history of major diseases; recent major negative life events based on the Social Readjustment Rating Scale;[17] perceived environmental exposures, risk perception and attribution of symptoms to environmental factors.

Those who provided a valid email address on registration but did not complete the first part of the questionnaire received a reminder email after approximately 2 weeks with the link to the questionnaire and a request to complete it. We consider 'respondents' as those who completed at least the informed consent form, 'participants' as those respondents who filled in at least the first part of the baseline questionnaire, and 'cohort members' as those participants who completed the consent form and at least the first part of the baseline questionnaire.

In the design of the online questionnaire, we made sure it resembles paper questionnaires as much as possible. For example, participants could read follow-up questions, but could only answer them when applicable, for example, if they ticked yes, then the applicable follow-up questions were activated and could be answered. In addition, participants could make a digital note at each question, such as notes written in the margin of paper questionnaires. The online application enabled certain logical checks, for example, questionnaire completeness and numerical checks (eg, out of range errors, such as starting date before birthdate). Only key questions were obligatory because they were necessary for the flow of the questionnaire (eg, birthdate and sex), but all other questions were not obligatory: participants were made aware of the unanswered questions on submission, but they had the opportunity to leave them blank.

We assessed potential health-related participation bias by comparing our cohort members with the source population with respect to the prevalence of various

(cardiovascular, musculoskeletal, neurological, psychological, respiratory and endocrine) disorders recorded in the EMRs in general practice. A comparison with the source population was possible, because this is based on available aggregate EMR data in the NIVEL Primary Care Database.

Statistical analysis

We performed descriptive analyses in which we calculated the participation rates by general practice and the 10th and 90th centiles of these rates. In addition, we calculated the average participation rates by the level of urbanisation (five levels from the most and the least urban areas, ie, on average more than 2500 and less than 500 addresses per km^2, respectively; Statistics Netherlands). This was carried out by first determining the level of urbanisation of the addresses of the general practices and of the baseline address of participants.

For the potential participation bias analysis, we calculated 95% CIs for the crude rates in the cohort and the source population. Upper and lower limits of the 95% CI, respectively, was calculated as follows: $(1000/n)$ $(d +(1.96 \times \sqrt{d}))$ and $(1000/n)$ $(d-(1.96 \times \sqrt{d}))$, where d is the number of events (nominator) and n is the population (denominator). If the point estimate of the cohort did not fall within the 95% CI of the prevalence rate of the source population, this is regarded as a statistically significant difference. If there was no overlap between both 95% CIs, these are qualified as the most notable differences in the text.

FINDINGS TO DATE
Response proportion

Figure 1 shows a flow chart of the recruitment process. In summary, a total of 93 550 enlisted patients were invited, of whom 20 926 (22%) responded, 14 862 (16%) were considered as participants and 14 829 (16%) as cohort members. The fast majority of the cohort members (96%) completed both parts of the questionnaire. Almost all of the cohort members (99.6%) can be linked to the unique patient number in the NIVEL Primary Care Database after verifying a match on sex and birthdate. Even though the invitations were addressed personally, it turned out that 113 cohort members were not the originally invited participants, but most likely another adult from the same household who did want to participate instead of the originally invited person. Since these 113 participants were eligible (ie, 31–65 years old, living in the Netherlands), they are treated as regular cohort members, of whom 58 can also be linked to the NIVEL Primary Care Database. We observed that most of the invitees responded on the day they received the letter or the day after that and that the (timing of the) reminder was effective (see figure 2), which is something to consider when planning the capacity of an online questionnaire and of personnel responding to questions.

Baseline results

Table 1 shows the baseline characteristics of the cohort members. We set out to recruit participants across the Netherlands to enhance contrast in environmental and occupational exposures, urbanisation level and socioeconomic factors. The mean and median age at baseline was 51 years (SD 9.4 years). Compared with the source population, the cohort members consist of more females (56%) and older subjects (50 plus years). We observed that the participation rate varied between the general practices (the 10th and 90th centiles were 9% and 23%, respectively) and also varied per level of urbanisation of the general practice location, varying on average from about 11% to 19% in the most and the least urban areas, respectively. This is also reflected in the distribution of cohort members by the level of urbanisation (less urban, more rural) compared with reference data from Statistics Netherlands (table 1). Nevertheless, we did succeed in recruiting participants across the Netherlands and with the varying level of urbanisation, as depicted in figure 3.

The majority was employed, never smoked cigarettes and did drink alcoholic beverages in the past 12 months (table 1). The fast majority was born in the Netherlands (97%). Those with intermediate levels of completed education are somewhat under-represented among cohort members compared with reference rates from Statistics Netherlands for 35–65-year-olds in 2012 (table 1). At baseline, the sex-stratified figures for overweight and obesity are similar yet somewhat lower than the national reference rates based on self-report for the age range 19–65 years (table 1).[18]
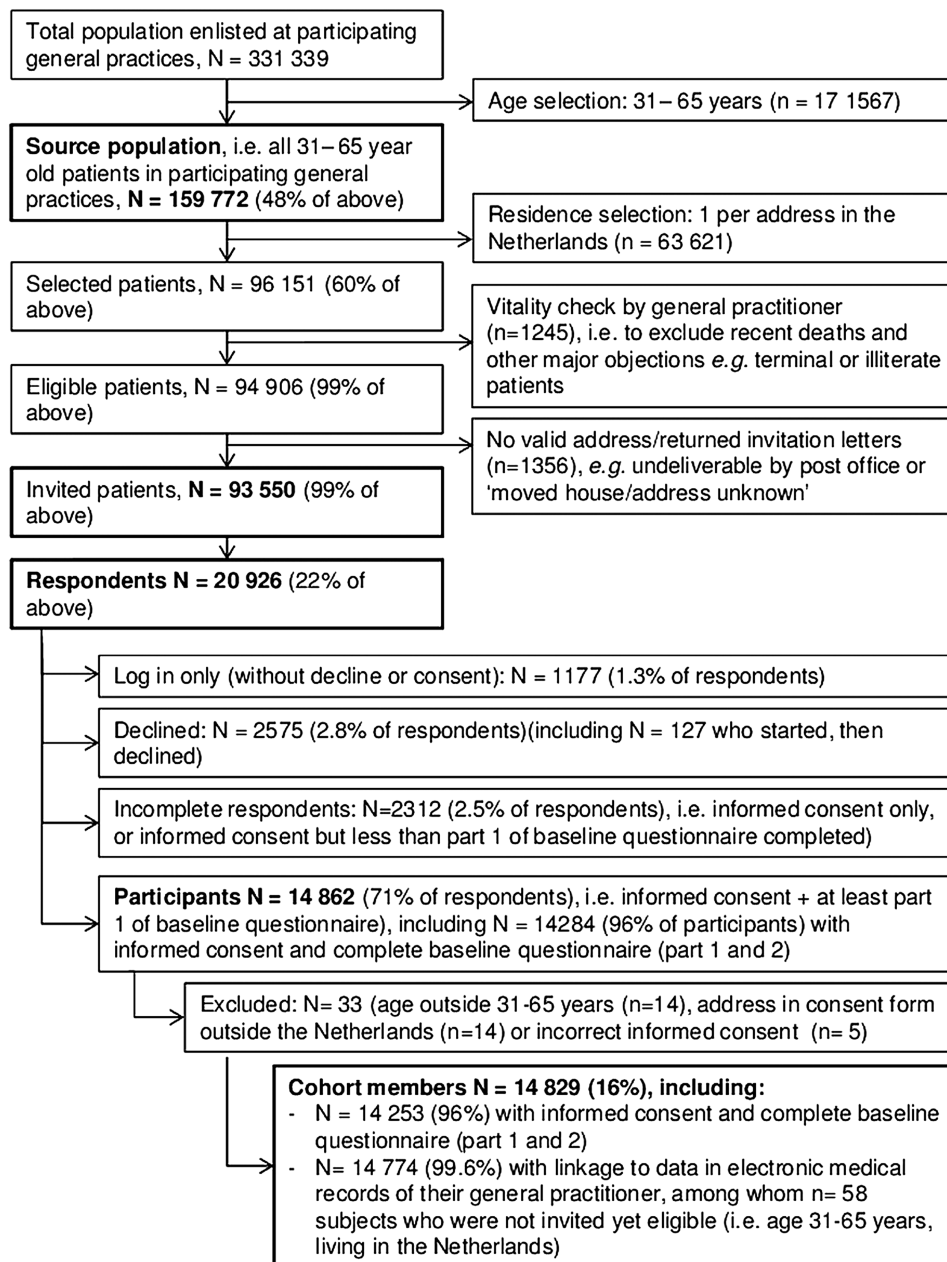
Health-related participation bias assessment

Table 2 shows that there are several small yet often statistically significant differences in the prevalence of selected disorders. The most notable differences (with no overlap between the 95% CIs of the prevalence rates) were seen for migraine and hypertension, which were more prevalent, and for diabetes and COPD (in particular among older women), which were less prevalent in the cohort compared with the source population.

STRENGTHS AND LIMITATIONS

In the AMIGO, participants will be prospectively followed through linkages to registries and follow-up measurements, such as questionnaires. A major strength of the prospective AMIGO for this field of research is its focus on environmental and occupational health from the outset, including a broad range of determinants and health outcomes. For example, baseline or current residential address or job is usually taken to model exposures. In AMIGO, we aim to extend this to health effects of exposures across the life course based on full residential and full occupational histories up to baseline supplemented with updates during prospective follow-up. While there is no reason to suspect differential recall bias by

**Figure 1** Flow chart of recruitment and participation.

Total population enlisted at participating general practices, N = 331 339

Age selection: 31 – 65 years (n = 17 1567)

Source population, i.e. all 31 – 65 year old patients in participating general practices, **N = 159 772** (48% of above)

Residence selection: 1 per address in the Netherlands (n = 63 621)

Selected patients, N = 96 151 (60% of above)

Vitality check by general practitioner (n=1245), i.e. to exclude recent deaths and other major objections *e.g.* terminal or illiterate patients

Eligible patients, N = 94 906 (99% of above)

No valid address/returned invitation letters (n=1356), *e.g.* undeliverable by post office or 'moved house/address unknown'

Invited patients, **N = 93 550** (99% of above)

Respondents N = 20 926 (22% of above)

Log in only (without decline or consent): N = 1177 (1.3% of respondents)

Declined: N = 2575 (2.8% of respondents)(including N = 127 who started, then declined)

Incomplete respondents: N=2312 (2.5% of respondents), i.e. informed consent only, or informed consent but less than part 1 of baseline questionnaire completed)

Participants N = 14 862 (71% of respondents), i.e. informed consent + at least part 1 of baseline questionnaire), including N = 14284 (96% of participants) with informed consent and complete baseline questionnaire (part 1 and 2)

Excluded: N= 33 (age outside 31-65 years (n=14), address in consent form outside the Netherlands (n=14) or incorrect informed consent (n= 5)

Cohort members N = 14 829 (16%), including:
- N = 14 253 (96%) with informed consent and complete baseline questionnaire (part 1 and 2)
- N= 14 774 (99.6%) with linkage to data in electronic medical records of their general practitioner, among whom n= 58 subjects who were not invited yet eligible (i.e. age 31-65 years, living in the Netherlands)

disease status, we will evaluate the potential cohort effects in future analyses, for example, related to differential recall or to incomplete job history because some participants were still working at the time of the questionnaires.

Another major asset of AMIGO is the availability of medical information from the EMRs of the general practitioners of the cohort members, not only at baseline but also for longitudinal follow-up because of the recruitment within an established information and surveillance network. This offers several rather unique opportunities. First, as shown here, unlike many other epidemiological studies, we were able to assess potential participation bias at baseline using aggregate data from the EMRs of the source population. The EMR data of the cohort members also enable us to assess future attrition bias in the active follow-up by means of

questionnaires. Second, besides the more usual registry linkages to obtain causes of death and cancer incidence, the additional medical data from general practitioners (diagnoses, prescriptions and referrals) enable us to study other recorded health outcomes, for which other cohort studies mostly rely on self-reported questionnaire data that are prone to reporting and recall bias and selective loss to follow-up. In particular, the main focus of prospective epidemiological studies has traditionally been on cancer, respiratory and cardiovascular health. In AMIGO, we can extend this focus to, for example, neurological diseases such as dementia and Parkinson's disease. Third, the longitudinal nature of the medical information of general practitioners enables one to study trajectories in morbidity linked to environmental and occupational determinants.

Figure 2 Timing of response (online registrations) in days after receiving the (A) invitation or (B) reminder.
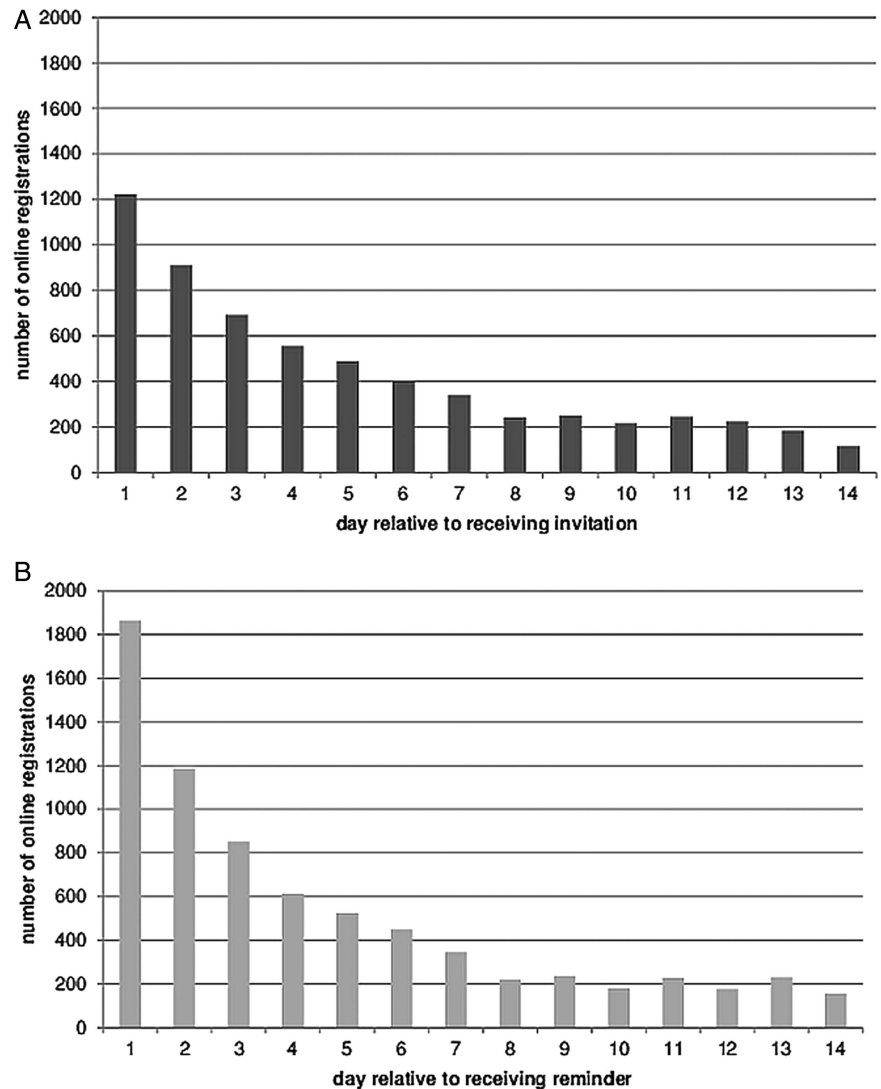
The main limitations of the study are those known to cohort studies that use active and passive follow-up, in particular selection bias and loss to follow-up related to future questionnaires, and the passive follow-up through EMRs in general practice, which will be truncated if cohort members move to another general practice. Of particular interest, related to selection bias and active follow-up is our choice to use *online* registration and (baseline) questionnaire(s). We argued that access to the internet is ubiquitous in the Netherlands, and that, owing to this online system, we could significantly cut costs such as printing and data entry, which enabled us to invite more participants. However, it requires from invitees the willingness and ability to access the internet and register and participate online. In time, as the cohort ages and to enhance the long-term participation of cohort participants, we will seek possibilities and pilot-test alternative modes such as also offering paper questionnaires on request or sending them along with reminder letters. We anticipated this possibility in the design of the online questionnaire and made sure that it resembled paper questionnaires as much as possible, as detailed in the Methods section.

With respect to selection bias at study entry, in our health-related participation bias analysis, we observed several statistically significant differences in general-practitioner recorded prevalence rates across several disorders and organ systems among cohort members compared with the source population. For example, 7 of the 10 studied disorders were statistically significantly more prevalent (most notably hypertension and migraine), while the other 3 were statistically significantly lower (most notably diabetes and COPD) in the total of cohort members compared with the source population. However, many of the statistically significant differences (in the total collective and some age and sex strata) were small. Moreover, we observed that the prevalence of one disorder in the same organ system is higher while another is less or similarly prevalent among the cohort members, which indicates that these differences are probably due to chance rather than differences in health or associated lifestyle. For example, while

Table 1 Baseline characteristics of the AMIGO cohort members (N=6561 men, and N=8268 women)

Characteristic	Per cent*	Comparative figures
Sex		The Netherlands (age 30–65 years)†
Men	44.2	50.2%
Women	55.8	49.8%
Age at baseline		The Netherlands (age 30–65 years)†
31–40 years	17.7	25.8%
41–50 years	30.4	32.1%
51–60 years	34.0	28.8%
60–65 years	17.9	13.3%
Civil state		
Married, registered partners or living together	80.3	
Divorced	6.0	
Widow(er)	2.4	
LAT relationship	2.3	
Single (no partner)	8.8	
Urbanisation level baseline address		All addresses in the Netherlands†
Very high (on average >2500 addresses per km^2)	8.4	20.5%
High (on average 1500–2500 addresses per km^2)	20.1	24.0%
Moderate (on average 1000–1500 addresses per km^2)	20.9	18.1%
Low (on average 500–1000 addresses per km^2)	24.4	18.6%
Very low (on average <500 addresses per km^2)	25.7	18.8%
Highest achieved level of education		The Netherlands (age 35–65 year)†
Low (vocational education/community college)	31	24%
Intermediate (vocational/high school)	32	55%
High (college/university or higher)	36	21%
Other/missing	0.8	0.6%
Employment status, most applicable		
Self-employed	9.9	
Employed	61.7	
(Early) retirement	9.3	
Unemployed/looking for work	2.4	
Sick leave/disability	4.6	
Housewife/men/charity/voluntary aid/study/other/social security	12.0	
Paid job		
Never	2.5	
Former	32.2	
Current	65.2	
Cigarette smoking		
Never	45.4	
Former	38.7	
Current	15.7	
Alcohol consumption		
Never	5.7	
Ever, but not in the past 12 months	7.4	
Currently on averages on 1 or 2 days per week in the past 12 months	36.0	
Currently on averages on 3 or more days per week in the past 12 months	50.7	
BMI		The Netherlands (age 19–65 years) based on self-report‡
Men		
Overweight: BMI 25–30 kg/m^2	47	40%
Obesity: BMI >30 kg/m^2	14	11%
Women		
Overweight: BMI 25–30 kg/m^2	32	27%
Obesity: BMI >30 kg/m^2	17	13%

*Percentages based on valid answers; there were 26, 60, 20, 18, 25, 28 and 28 missing values for civil state, urbanisation level baseline address, employment status, paid job, cigarette smoking, alcohol consumption and BMI, respectively.
†Statistics Netherlands (2013).
‡See ref 18.
AMIGO, Occupational and Environmental Health Cohort Study; BMI, body mass index; LAT, living-apart-together.
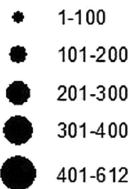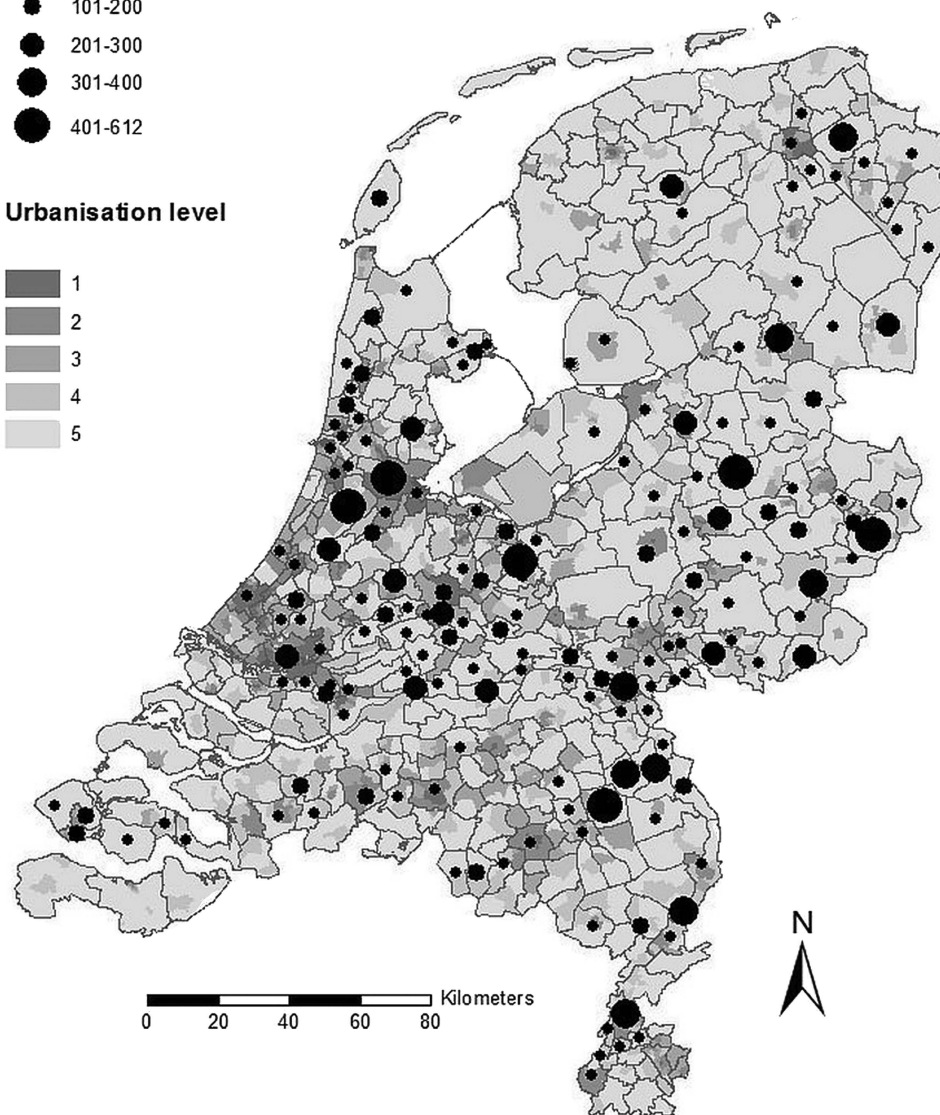
Figure 3 Geographical spread of the Occupational and Environmental Health Cohort Study (AMIGO) across the Netherlands at baseline. Legend: number of AMIGO cohort members (dots) and level of urbanisation per municipality. 1=Very high (on average >2500 addresses per km^2); 2=high (on average 1500–2500 addresses per km^2); 3=moderate (on average 1000–1500 addresses per km^2); 4=low (on average 500–1000 addresses per km^2); and 5=very low (on average <500 addresses per km^2).

hypertension and COPD were more prevalent, cerebro-vascular accidents and asthma, respectively, were less prevalent among cohort members, which does not seem to point at a participation bias based on smoking behaviour. Taken together, therefore, we found no consistent indications of systematic health-related participation bias based on these measures of morbidity or associated lifestyle such as smoking.

Comparison of the baseline characteristics to national reference rates indicated that men below 50 years of age and those with an intermediate level of education are to some extent under-represented among cohort members, while those born in the Netherlands are over-represented, probably in part due to the fact that the

questionnaire was in Dutch. The observed differences are no major concern for the internal validity of the foreseen exposure–response relationships, given the contrast we achieved in sociodemographics, geographical spread and urbanisation, and associated environmental and occupational exposures, and the results of our health-related participation bias analysis. The presented results can in the future be used for weighting purposes if generalisation to the general adult population is desired.

In conclusion, we established the AMIGO, which offers a rich research resource to enhance the knowledge base with prospective results on occupational and environmental health, including novel opportunities with general practice-based health outcomes.

Table 2 Prevalence rates (per 1000) of selected disorders among cohort members compared with the SP

Age (years)	Sex	Number of subjects		Hypertension		TCI/stroke/CVA		Back pain with radiating symptoms		Rheumatoid arthritis		Migraine		Disturbances of sleep/insomnia		Depressive disorder		COPD		Asthma		Diabetes mellitus	
		Cohort	SP	Cohort	SP	Cohort	SP	Cohort	SP	Cohort	SP	Cohort	SP	Cohort	SP	Cohort	SP	Cohort	SP	Cohort	SP	Cohort	SP
31–40	M	1041	20 855	28.8	19.5	0.96	1.63	32.7	34.8	7.68	3.84	19.2	12.0	13.4	18.3	17.3	25.5	2.88	3.45	56.7	45.5	8.65	12.0
	W	1572	21 759	27.4	21.8	0.64	1.98	37.5	37.5	7.00	5.51	55.3	50.1	24.8	26.7	41.3	52.2	1.91	2.85	55.3	58.7	5.09	8.87
	M+W	2613	42 614	27.9	20.7	0.77	1.81	35.6	36.2	7.27	4.69	40.9	31.5	20.3	22.6	31.8	39.1	2.30	3.14	55.9	52.2	6.51	10.4
41–50	M	1891	26 809	72.4	64.2	4.23	5.11	42.8	51.3	5.82	4.77	18.0	14.2	25.9	26.8	34.9	35.3	4.76	8.95	51.8	47.5	28.6	29.3
	W	2602	26 520	88.0	72.0	6.15	7.16	44.2	49.8	10.4	9.13	64.6	54.1	39.6	34.6	53.4	62.6	8.46	10.7	72.6	61.9	15.8	23.2
	M+W	4493	53 329	81.5	68.1	5.34	6.13	43.6	50.6	8.46	6.94	45.0	34.0	33.8	30.7	45.6	48.9	6.90	9.83	63.9	54.7	21.1	26.3
51–60	M	2324	22 448	177.3	161.6	20.7	19.1	64.5	63.6	10.3	9.49	24.1	14.7	32.7	34.8	37.9	43.9	24.1	28.2	52.5	43.4	65.8	73.3
	W	2694	21 723	197.1	182.2	13.4	16.6	59.8	59.2	14.5	14.8	60.5	46.9	52.3	51.1	63.5	68.6	18.6	33.1	64.6	60.2	36.7	58.0
	M+W	5018	44 171	187.9	171.8	16.7	16.7	62.0	61.4	12.6	12.1	43.6	30.5	42.3	42.8	51.6	56.1	21.1	30.6	59.0	51.7	50.2	65.0
61–65	M	1290	9895	272.1	261.4	44.2	38.6	55.8	65.2	11.6	11.7	16.3	12.3	36.4	37.3	33.3	40.6	40.3	52.2	40.3	38.6	123.3	124.4
	W	1360	9763	274.3	274.9	16.9	29.2	73.5	65.3	20.6	19.1	34.6	31.0	51.5	58.5	52.2	62.4	29.4	43.2	54.4	58.0	71.3	102.2
	M+W	2650	19 658	273.2	268.1	30.2	33.9	64.9	65.3	16.2	15.4	25.7	21.6	44.2	47.8	43.0	51.4	34.7	47.8	47.5	48.2	96.6	113.4
Total	M	6546	80 007	142.1	104.3	17.4	12.3	51.5	52.2	8.86	6.71	20.0	13.5	28.4	28.1	32.8	35.8	18.3	18.3	50.6	44.7	57.3	48.9
	W	8228	79 765	142.9	113.2	9.24	11.0	52.9	50.9	12.8	10.9	56.5	48.2	42.9	39.8	54.2	61.4	14.0	18.6	63.7	60.1	29.8	38.5
	M+W	14 774	159 772	142.5	108.7	12.9	11.6	52.3	51.5	11.0	8.80	40.3	30.8	36.5	34.0	44.7	48.6	15.9	18.5	57.9	52.4	42.0	43.7

Italic: $p < 0.05$, point estimate of the prevalence rates of the cohort members does not overlap with the 95% CI of the point estimate of the SP.

Based on the electronic medical records in general practice of n=14 774 (99.6%) of the cohort members versus the SP (ie, all participants aged 31–65 years at the time of selection and enlisted at one of the participating general practices (N=159 772)), that is, rates per 1000 regarding episodes of disorders with starting date before the date of selection for AMIGO: cardiovascular (hypertension (un)complicated; TCI/stroke/CVA); musculoskeletal (back pain with radiating symptoms; rheumatoid arthritis); neurological (migraine); psychological (depressive disorder; disturbances of sleep/insomnia); respiratory (COPD; asthma); and endocrine (diabetes mellitus) disorders.

AMIGO, Occupational and Environmental Health Cohort Study; CI, confidence interval; cohort, cohort members; COPD, chronic obstructive pulmonary disease; CVA, cerebrovascular accident; M, men; SP, source population; TCI, transient cerebral ischaemia; W, women.

Collaboration

We are now in the phase of prospective follow-up, with the aim of continuing this for as long as possible (20+ years), pending future funding. We cordially invite other researchers to propose non-commercial research based on the available data in AMIGO or requests for additional data collection with associated funding. Any such requests can be submitted to amigoproject@uu.nl or the corresponding author. Requests are reviewed by the AMIGO management committee and proposals should fulfil a number of criteria including that the work is within the bounds of consent given by participants and a data management fee.

REFERENCES

1. Prüss-Üstün A, Corvalán C. *Preventing disease through healthy environments. Towards an estimate of the environmental burden of disease*. World Health Organisation, 2006.
2. Saberi Hosnijeh F, Christopher Y, Peeters P, *et al*. Occupation and risk of lymphoid and myeloid leukaemia in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Occup Environ Med* 2013;70:464–70.
3. Schüz J, Elliott P, Auvinen A, *et al*. An international prospective cohort study of mobile phone users and health (Cosmos): design considerations and enrolment. *Cancer Epidemiol* 2011;35:37–43.
4. NIVEL Primary Care Registry. http://www.nivel.nl/en/dossier/nivel-primary-care-database (accessed 5 Mar 2014).
5. Lamberts H, Wood M. *ICPC. International classification of primary care*. Oxford: Oxford University Press, 1987.
6. World Health Organisation. *ICD-10: international statistical classification of diseases and related health problems: tenth revision*; 2004.
7. World Health Organisation. *Guidelines for ATC classification and DDD assignment 2010*. Oslo: World Health Organisation, 2009.
8. Karasek R, Brisson C, Kawakami N, *et al*. The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *J Occ Health Psychol* 1998;3:322–55.
9. International Labour Office. *International Standard Classification of Occupations (ISCO)*. Geneva: International Labour Office, 1968.
10. Kosinski M, Bayliss MS, Bjorner JB, *et al*. A six-item short-form survey for measuring headache impact: the HIT-6. *Qual Life Res* 2003;12:963–74.
11. Lipton RB, Dodick D, Sadovsky R, *et al*. A self-administered screener for migraine in primary care. The ID Migraine validation study. *Neurology* 2003;61:375–82.
12. Hays RD, Martin SA, Sesti AM, *et al*. Psychometric properties of the Medical Outcomes Study Sleep measure. *Sleep Med* 2005;6:41–4.
13. Rocca WA, Maraganore DM, McDonnell SK, *et al*. Validation of a telephone questionnaire for Parkinson's disease. *J Clin Epidemiol* 1998;51:517–23.
14. Terluin B, van Marwijk HW, Adèr HJ, *et al*. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006;6:34.
15. Burney PG, Luczynska C, Chinn S, *et al*. The European Community Respiratory Health Survey. *Eur Respir J* 1994;7:954–60.
16. Rijcken B, Kerkhof M, de Graaf A, *et al*. *Europees luchtweg onderzoek Nederland*. Groningen: Stichting drukkerij Regenboog, 1996.
17. Holmes TH, Rahe RH. The social readjustment rating scale. *J Psychosom Res* 1967;11:213–18.
18. Gezondheidsmonitor GGD'en, CBS en RIVM. http://www.nationaalkompas.nl/gezondheidsdeterminanten/persoonsgebonden/overgewicht/hoeveel-mensen-hebben-overgewicht/ (accessed 27 Nov 2013).