Int. J. Epidemiol. Advance Access published May 29, 2009

Published by Oxford University Press on behalf of the International Epidemiological Association
International Journal of Epidemiology 2009;1–9
doi:10.1093/ije/dyp217

COHORT PROFILE

Cohort Profile: The EPIC-NL study

Joline W J Beulens,[1,2]* Evelyn M Monninkhof,[1] Monique W M Verschuren,[2] Yvonne T van der Schouw,[1] Jet Smit,[2] Marga C Ocke,[3] Eugene H J M Jansen,[4] Susan van Dieren,[1] Diederick E Grobbee,[1] Petra H M Peeters[1] and Bas H B Bueno-de-Mesquita[3]

Accepted 22 April 2009

How did the study come about?

A major scientific challenge for the next few decades is to understand the interaction between genetic susceptibility and environmental factors in the aetiology of chronic diseases. The most promising approach to discover these interactions requires a combined effort of epidemiology and molecular genetics and large sample sizes for sufficient power. Already in the early 90s, the European Prospective Investigation Into Cancer and Nutrition (EPIC) was initiated in 10 European countries to create a large cohort to study the aetiology of chronic diseases.[1,2] The Netherlands has contributed two cohort studies to EPIC: the Prospect cohort of 17 357 women of the Julius Center[3] in Utrecht, and the Monitoring Project on Risk Factors for Chronic Diseases (MORGEN) cohort of 22 654 men and women of the National Institute for Public Health and the Environment (RIVM) in Bilthoven.[4] In the design phase, both cohorts collaborated closely to obtain maximal synergy in the design of the questionnaires and to follow identical protocols in the collection of biological samples. Because of the efficiency gain in maintaining the cohort infrastructure and in conducting scientific analyses, the Julius Center and the RIVM decided to combine efforts to maintain and expand the cohorts and biobanks by merging them into one EPIC-Netherlands (EPIC-NL) study.

What does it cover?

Initially, the aim of the EPIC study was to investigate the role of nutrition in the aetiology of cancer.[1,2] The Prospect cohort was set up in this context to investigate the role of nutrition in the aetiology of cancer, whereas the MORGEN cohort had a broader goal to monitor risk factors for chronic diseases in the Netherlands. In addition to nutrition, both Dutch cohorts also focused on other lifestyle factors, such as smoking, alcohol and physical activity, whereas reproductive factors were more extensively assessed in the Prospect cohort and occupational factors in the MORGEN cohort. Now, the focus of the international EPIC study has broadened to include most major chronic diseases as an endpoint, such as obesity, cardiovascular diseases[5] and type 2 diabetes (www.inter-act.eu).

Who is in the sample?

The baseline measurements were performed in 40 011 respondents; 17 357 women from the Prospect cohort and 22 654 men and women from the MORGEN cohort. The recruitment procedure for both cohorts differed inherently to the aims of both studies. Prospect is a prospective cohort study among women aged 49–70, residing in the city of Utrecht or its vicinity, who participated in the nation wide Dutch breast cancer screening programme between 1993 and 1997. A general questionnaire and a food-frequency questionnaire (FFQ) were mailed to women who agreed to participate, and these were returned when women attended the screening unit.[3] The MORGEN cohort consists of a general population sample of men and women aged 20–59 years from three Dutch towns (Amsterdam, Doetinchem and Maastricht).[4] From 1993 to 1997 each year a new random sample, consisting of ~5000 subjects, was examined. A total of 50 766 persons received an invitation to participate in the MORGEN cohort. Those who returned a reply card received two questionnaires by mail (a general questionnaire on

[1] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands.
[2] Center for Prevention and Health Services Research, Bilthoven, The Netherlands.
[3] Center for Nutrition and Health, Bilthoven, The Netherlands.
[4] Laboratory for Health Protection Research, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands.

* Corresponding author. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, PO Box 85500, 3508 GA, Utrecht, The Netherlands. E-mail: J.Beulens@umcutrecht.nl

socio-demographic factors, lifestyle and health indicators, and an FFQ. They were then asked to visit the local Public Health Service for a medical examination. The baseline characteristics of the 40 011 men and women from the individual Prospect and MORGEN cohorts and the combined EPIC-NL study are shown in Table 1.

The EPIC-NL study was funded by 'Europe against Cancer' Programme of the European Commission (SANCO); the Dutch Ministry of Public Health, Welfare and Sports (formerly Ministry of Welfare, Public Health and Culture); the Dutch Cancer Society; ZonMW the Netherlands Organisation for Health Research and Development; and World Cancer Research Fund (WCRF).

What was the baseline response rate?

A total of 50 313 Dutch women were invited to take part in the Prospect cohort (1993–97), and 17 357 (34.5%) participated. Compliance was somewhat higher for younger women and for women from less urbanized areas. Participation rates were lower in the last study year in Prospect for logistic reasons. In total, 50 766 persons were approached for participation in the MORGEN cohort. Finally, 22 769 participants (45%) completed the total survey including questionnaires and a medical examination. Overall, response rates were higher among women (49%) than men (41%) and were higher among older persons (30% among people aged 20–29 and 54% among people aged 50–65). Response rates were highest in Doetinchem (68%), intermediate in Maastricht (45%) and lowest in Amsterdam (34%). A decline in response rates was seen throughout the years, from 48% in 1993 to 40% in 1997. More detailed information on characteristics is described by van Loon *et al*.[6] Overall, this resulted in an overall response rate of 40% for the EPIC-NL study.

How often have they been followed?

Participants are followed for the occurrence of cancer and cardiovascular diseases by annual linkage to several disease registries. Postal follow-up questionnaires were also sent twice with intervals of 3–5 years to participants, in order to detect changes in lifestyle and for occurrence of diseases that cannot accurately be obtained from disease registries (e.g. diabetes, fractures, etc.). About 6000 participants from the MORGEN cohort originating from the municipality of Doetinchem are repeatedly examined every 5 years[7] to obtain more detailed measurements, including blood pressure, body weight and blood sampling.

Vital status of all EPIC-NL participants is obtained through linkage with the municipal population registries. Subsequently, causes of death for the deceased persons are obtained through linkage with 'Statistics Netherlands'. To date, information on vital status and causes of death for the EPIC-NL study is complete until January 1, 2006.

Cancer cases in the EPIC-NL study are identified by annual linkage to the Dutch Cancer Registry.[8] This registry identifies incident cancer cases (hospitalization records/pathology records) and is 95% complete since 1989. The Dutch National Cancer Registry gathers its data from all eight regional cancer centres.[9] The MORGEN cohort of EPIC-NL is linked to the Dutch Cancer Registry because participants are residing in several geographical areas covered by different regional integral cancer centres. The Prospect cohort of EPIC-NL is linked to two regional cancer registries: IKMN (Integraal Kankercentrum Midden Nederland) and IKO (Integraal Kankercentrum Oost). Linkage to the Cancer Registry is based on information of birth date, gender, full name (family name + initials), first four characters of family name of spouse, place of birth and postal code and date of mutation of postal code. Follow-up for cancer incidence of the EPIC-NL study is complete until January 1, 2004 for MORGEN and until January 1, 2006 for Prospect. Prevalent cases of cancer were also identified through linkage with the cancer registry (1989–97) and by self-report using the baseline general questionnaire.

Data on morbidity other than cancer were obtained from the Dutch Hospital Association and Order of Medical Specialists, which holds a standardized computerized register of hospital discharge diagnoses. Since 1990, admission files have been stored continuously from all general and university hospitals in The Netherlands. Data on sex, date of birth, dates of admission and discharge are recorded whenever a patient is discharged from hospital. One mandatory principal diagnosis and up to nine optional additional diagnoses are reported. All diagnoses are coded according to the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9). Coding of the diagnoses is performed by qualified medical administrative personnel in the hospitals. The National Medical Registry (NMR) collects these data in the Hospital Discharge Diagnosis Database. These data are checked by the NMR, mistakes are corrected by the hospitals, and unlikely diagnoses are discussed with the hospital. The database is linked to the cohort based on information of birth date, gender, postal code and general practitioner with a validated probabilistic method.[10] Follow-up of EPIC-NL is currently complete until January 1, 2006. Prevalent cases of cardiovascular disease were also identified through linkage with the NMR (1990–97) and by self-report using the baseline general questionnaire.

The occurrence of type 2 diabetes was assessed by self-report in the follow-up questionnaires, combined

Table 1 Baseline characteristics of the EPIC-NL study population

	MORGEN		Prospect	EPIC-NL
	Male	Female		
N		22 654	17 357	40 011
Gender (m/f)	10 260/–	–/12 394	–/17 357	10 260/29 751
Age (years)	43 ± 11 (20–66)	42 ± 11 (20–66)	58 ± 6 (49–70)	49 ± 12 (20–70)
BMI (kg/m^2)	25.8 ± 3.6 (15.6–50.1)	25.2 ± 4.3 (14.9–58.3)	26.0 ± 4.1 (14.0–52.0)	25.7 ± 4.1 (14.0–58.3)
Overweight[a] (%)	56.1	43.2	49.1	49.0
Waist circumference (cm)	92.5 ± 10.9 (62.8–155.0)	81.5 ± 11.4 (55.8–150.0)	83.7 ± 10.2 (52.0–140.0)	85.3 ± 11.6 (52.0–155.0)
Hip circumference (cm)	101.4 ± 6.6 (69.0–158.0)	102.3 ± 8.6 (64.0–165.0)	105.7 ± 8.4 (76.0–152.0)	103.5 ± 8.3 (64.0–165.0)
Waist/hip ratio	0.91 ± 0.07 (0.62–1.33)	0.80 ± 0.07 (0.48–1.36)	0.79 ± 0.06 (0.59–1.06)	0.82 ± 0.08 (0.48–1.36)
Systolic blood pressure (mm Hg)	124.8 ± 15.3 (82.0–226.0)	117.7 ± 16.3 (72.0–222.0)	133.2 ± 20.1 (80.0–250.0)	126.2 ± 19.0 (72.0–250.0)
Diastolic blood pressure (mm Hg)	78.9 ± 10.5 (43.0–143.0)	74.8 ± 10.5 (42.0–130.0)	79.3 ± 10.4 (50.0–150.0)	77.8 ± 10.7 (42.0–150.0)
Hypertension[b] (%)	32.2	31.1	44.6	37.7
Physical activity (CPAI)[c]				
Inactive (%/*n*)	13.6	11.2	7.4	9.9
Moderately inactive (%/*n*)	30.9	31.7	26.4	28.9
Moderately active (%/*n*)	27.7	29.1	25.4	26.9
Active (%/*n*)	28.3	28.1	40.1	34.3
Smoking (%)				
Never	30.6	38.0	43.0	38.3
Former	30.8	26.9	34.6	31.2
Current	38.7	35.1	22.4	30.5
Alcohol consumption (g/day)	17.8 ± 20.9 (0–241.4)	7.8 ± 11.6 (0–180.6)	9.1 ± 12.6 (0–164.8)	10.9 ± 15.5 (0–241.4)
Education[d] (%)				
High	26.6	21.1	16.2	20.4
Middle	39.4	42.9	37.6	39.7
Low	34.0	36.1	46.1	39.9
Post-menopausal status[e] (%)	–	20.6	71.8	50.5
Ever used oral contraceptives[e] (%)	–	84.3	64.4	72.7
Ever hormone replacement therapy[e] (%)	–	12.2	25.8	20.9
Prevalent cancer morbidity (%)	1.4/139	3.4/421	6.5/1.108	4.2/1.668
Prevalent myocardial infarction (%)	2.6/259	0.7/78	2.2/373	1.8/710
Prevalent stroke (%)	1.0/95	0.8/100	1.8/299	1.3/494
Prevalent diabetes mellitus[f] (%)	1.4/144	1.2/151	3.0/521	2.1/816
Hyperlipidaemia (%)	15.3/1.561	6.7/831	5.6/968	8.4/3.360

[a]BMI > 25.0 kg/m^2.
[b]Defined by a physician-diagnosed self-report, measured hypertension (>140 systolic or >90 diastolic) or the use of hypertensive medication.
[c]CPAI = Cambridge Physical Activity Index.
[d]Low: primary education up to completing intermediate vocational education; middle: up to higher secondary education; high: those with higher vocational education and university.
[e]Only among women.
[f]Based on self-report only.

with the result of the use of a urinary glucose strip test for detection of glucosuria (only in Prospect) and/or linkage to the NMR as described. Participants received a urinary glucose strip test, enclosed with the first follow-up questionnaire. They were asked whether the urine strip had turned purple after 10 s, indicating glucosuria. Details are described elsewhere.[11] Prevalent cases of diabetes were identified through linkage with the NMR (1990–97) and by self-report using the general baseline questionnaire. Both prevalent and incident cases of diabetes detected by either of these methods are being verified by consulting their medical records at general practitioners and will be described elsewhere.

What is the response rate during follow-up?

Chronic disease endpoints are obtained through linkage with several existing disease registries. Linkages are performed only for those respondents who gave written informed consent (which is the case for >95% of the respondents). Therefore, linkage success rate for cardiovascular endpoints and type 2 diabetes is 97.6% ($n = 39\,055$) and for cancer endpoints it is 97.3% ($n = 38\,937$). For Prospect follow-up questionnaires, the response rates were 76.2% ($n = 13\,229$) and 68.1% ($n = 11\,812$) for the first and second follow-up questionnaire. In the MORGEN study, the response rate was 66.6% (10 284 of 15 439) for the first follow-up questionnaire in Amsterdam and Maastricht (1998–2002) and 74.7% (4916 of 6578) in Doetinchem (1998–2002), whereas this was 78.2% (4519 of 5781) for the second follow-up questionnaire in Doetinchem (2003–07) and 57.0% (7903 of 13 862) for the second follow-up questionnaire in Amsterdam and Maastricht (2003–07). Non-responders to these follow-up questionnaires were reminded by letter.

What has been measured?

At baseline, a general questionnaire and an FFQ were filled in by all participants. A physical examination including body weight, waist and hip circumference and blood pressure measurement was performed and blood samples were drawn. During follow-up, participants were followed for disease occurrence through registries and for lifestyle changes through questionnaires send with 3- to 5-year intervals.

General questionnaire and physical examination

At baseline, a general questionnaire containing questions on demographic characteristics, presence of chronic diseases and related potential risk factors was administered. Comparable information on lifestyle, disease history and diet was collected in the Prospect and MORGEN cohorts, although each project had additional questions on different topics. Coding of this information was standardized and merged into one database.

In Prospect, systolic and diastolic blood pressure was measured in supine position twice using a Boso Oscillomat (Bosch & Son, Jungingen, Germany) with a cuff of 15×52 cm on their left arm. Approximately 5–15 min after arrival, the first blood pressure assessment was performed and this was repeated after 10 min. In MORGEN, measurement of systolic and diastolic blood pressure was performed at the end of the study visit using a random zero Sphygmomanometer. The measurement was performed in supine position on the left arm using a cuff of 12×23 cm (or, if necessary, 15×33 or 9×18 cm). Systolic blood pressure was recorded when continuing rhythmic tunes were present (first-phase Korotkoff) and diastolic blood pressure was read when tunes were no longer present (fifth-phase Korotkoff). The comparability of both measurement procedures has been reported in more detail.[12] The blood pressure assessment of Prospect slightly overestimates blood pressure. For both cohorts, body weight was measured in light indoor clothing without shoes to the nearest 0.5 kg with a floor scale (Seca, Atlanta, GA, USA). Waist and hip circumference were measured identically as well. Body mass index was calculated as weight divided by height squared (kg/m^2).

Duration and types of physical activity during the year preceding study recruitment were assessed by a set of questions that was used in all EPIC cohorts.[13] By combining occupational physical activity with time spent on cycling and sporting in summer and winter, the validated Cambridge Physical Activity Index (CPAI)[14] was calculated. Based on this index, participants were divided in four physical activity categories (inactive, moderately inactive, moderately active and active).

FFQ

In both cohorts of EPIC-NL, food consumption was assessed using an identical self-administered FFQ including questions on the usual frequency of consumption of 77 main food categories during the year, preceding enrolment. Further information was collected on consumption frequency for selected subitems (e.g. skimmed, semi-skimmed or full-fat milk), preparation methods, additions (e.g. sugar), use of dietary supplements, special diets and brand names of fats used on bread and for cooking. Colour photographs were used to estimate portion size for 28 food items. Overall, the questionnaire allows the estimation of the average daily consumption of 178 foods. The FFQ has been validated before start of the study[15,16] against twelve 24-h recalls, and biomarkers in 24-h urine and serum. Table 2 shows intake of some selected nutrients of the EPIC-NL study.

Table 2 Mean ± standard deviation (SD) (range) dietary intake of food groups and nutrients of the EPIC-NL study population

Nutrients	MORGEN		Prospect	EPIC-NL
	Male	Female		
Energy (kcal)	2592 ± 728 (323–7782)	1981 ± 535 (320–6719)	1789 ± 439 (152–5697)	2055 ± 644 (152–7782)
Protein (g)	94.0 ± 26.5 (14.5–335.8)	74.7 ± 20.0 (10.3–243.1)	71.5 ± 18.1 (7.1–244.6)	78.3 ± 23.1 (7.1–335.8)
Protein (en%)	14.7 ± 2.2 (5.3–37.2)	15.3 ± 2.4 (5.7–29.5)	16.1 ± 2.4 (6.0–29.5)	15.5 ± 2.4 (5.3–37.2)
Fat (g)	103.0 ± 34.6 (6.7–392.3)	79.4 ± 26.6 (8.6–306.6)	71.3 ± 23.2 (2.1–269.3)	82.0 ± 30.4 (2.1–392.4)
Fat (en%)	35.5 ± 5.1 (9.4–59.4)	35.8 ± 5.2 (11.0–67.7)	35.5 ± 5.6 (5.1–60.6)	35.6 ± 5.3 (5.1–67.7)
Saturated fat (g)	42.3 ± 15.1 (3.1–149.1)	32.8 ± 11.6 (3.3–132.9)	30.7 ± 10.6 (1.3–128.7)	34.3 ± 13.1 (1.3–149.1)
Monounsaturated fat (g)	39.9 ± 14.1 (2.7–180.5)	30.5 ± 10.9 (3.0–143.5)	26.5 ± 9.2 (0.6–125.0)	31.2 ± 12.4 (0.6–180.5)
Polyunsaturated fat (g)	20.0 ± 7.7 (0.7–91.3)	15.4 ± 5.8 (1.2–72.9)	13.5 ± 5.2 (0.1–79.8)	15.7 ± 6.7 (0.1–91.3)
Carbohydrates (g)	291.3 ± 91.1 (21.6–905.1)	228.2 ± 68.3 (18.8–956.1)	199.3 ± 53.6 (17.8–585.7)	231.9 ± 78.7 (17.8–956.1)
Carbohydrates (en%)	45.0 ± 6.2 (12.8–75.7)	46.1 ± 6.2 (12.5–76.6)	44.8 ± 6.4 (17.7–86.3)	45.2 ± 6.3 (12.5–86.3)
Mono- and disaccharides (g)	136.4 ± 55.3 (9.5–536.5)	112.9 ± 43.9 (9.9–660.4)	107.2 ± 36.4 (4.7–418.1)	116.5 ± 45.8 (4.7–660.4)
Polysaccharides (g)	154.9 ± 51.9 (1.3–572.9)	115.1 ± 37.6 (4.0–354.9)	91.8 ± 27.3 (0.7–328.6)	115.3 ± 45.8 (0.7–572.9)
Fibre (g)	27.1 ± 8.2 (2.6–99.8)	22.8 ± 6.1 (2.1–65.9)	22.5 ± 5.6 (0.2–57.4)	23.8 ± 6.8 (0.2–99.8)
Calcium (mg)	1138.1 ± 519.6 (153.4–6229.7)	1036.0 ± 421.2 (103.9–5411.8)	1099.0 ± 396.8 (74.8–5764.8)	1089.4 ± 440.6 (74.8–6229.7)
Iron (mg)	14.0 ± 3.8 (2.7–43.5)	11.5 ± 2.8 (1.9–39.6)	10.6 ± 2.4 (0.4–32.7)	11.8 ± 3.2 (0.4–43.5)
Potassium (mg)	4214.3 ± 1092.4 (735.0–12 052.6)	3542.3 ± 874.0 (545.4–12 106.0)	3447.8 ± 775.1 (333.5–9602.0)	3674.4 ± 952.3 (333.5–12 106.0)
Retinol (μg)	1075.8 ± 835.8 (13.2–14 665.4)	708.2 ± 535.9 (5.1–7817.1)	673.7 ± 513.6 (13.6–10742.5)	787.7 ± 641.6 (5.1–14 665.4)
Beta-carotene (μg)	1531.8 ± 646.0 (73.2–9845.5)	1598.9 ± 688.3 (35.7–13 273.3)	1557.7 ± 650.3 (78.7–9869.8)	1563.9 ± 661.6 (35.7–13 273.3)
Vitamin C (mg)	100.2 ± 43.9 (7.3–488.7)	108.7 ± 44.9 (3.5–433.9)	116.3 ± 47.4 (4.8–643.9)	109.8 ± 46.2 (3.5–643.9)
Vitamin B1 (mg)	1.26 ± 0.36 (0.21–4.66)	1.02 ± 0.27 (0.10–3.69)	0.98 ± 0.24 (0.08–2.72)	1.07 ± 0.31 (0.08–4.66)
Vitamin B2 (mg)	13.4 ± 21.8 (0.3–427.4)	9.4 ± 16.1 (0.2–294.4)	8.8 ± 13.8 (0.2–285.1)	10.1 ± 17.0 (0.2–427.4)
Vitamin B6 (mg)	2.04 ± 0.59 (0.36–6.74)	1.57 ± 0.41 (0.14–4.34)	1.42 ± 0.34 (0.11–3.78)	1.62 ± 0.50 (0.11–6.74)
Vitamin B12 (μg)	5.80 ± 2.98 (0.05–52.5)	4.48 ± 2.18 (0.22–33.5)	4.4 ± 2.1 (0.2–40.8)	4.8 ± 2.5 (0.1–52.5)
Vitamin E (mg)	15.2 ± 6.0 (1.3–80.5)	12.6 ± 4.7 (1.2–57.7)	11.4 ± 4.2 (0.5–70.4)	12.7 ± 5.1 (0.5–80.5)
Vitamin D (μg)	3.7 ± 1.6 (0.1–15.5)	2.8 ± 1.2 (0.1–13.4)	2.7 ± 1.1 (0.1–15.7)	3.0 ± 1.4 (0.1–15.7)
Zinc (mg)	12.3 ± 3.6 (1.9–45.8)	9.8 ± 2.7 (1.3–34.4)	9.5 ± 2.5 (0.9–33.9)	10.3 ± 3.1 (0.9–45.8)
Magnesium (mg)	409.7 ± 113.5 (61.5–1365.1)	330.3 ± 83.4 (37.8–1158.7)	315.6 ± 71.7 (33.2–1065.0)	344.3 ± 96.0 (33.2–1365.1)

Blood sampling

In the MORGEN cohort, ethylene-diamine-tetra-acetic acid (EDTA) plasma has been collected. In both cohorts of EPIC-NL, the EPIC blood sampling protocol was identical. Participants donated a 30-ml non-fasting blood sample, using three Sarstedt safety vacuum monovettes, one dry monovette for serum and two citrated monovettes for plasma. Samples were protected against light and temporarily stored at temperatures between 4 and 10°C. Within 24 h, samples of 4 ml serum, 9 ml citrate plasma, 2 ml white blood cells and 2 ml red blood cells were fractionated into 28×0.5-ml aliquots and stored in heat-sealed plastic straws, first at –80°C and later under liquid nitrogen at –196°C, for future use. Half of all aliquots (14 straws per subject) were transported to Lyon by road under liquid nitrogen to be stored in the central European EPIC bank of biological samples at the International Agency for Research on Cancer. The other half of the aliquots were stored locally in one biobank in a single utility at the RIVM. Therefore, EPIC-NL has collected two identical biobanks: one is part of the European EPIC study and stored at IARC; the other is identical but is located at RIVM.

Biochemical measurements

A 6.5% random sample of the baseline cohort was taken for the assessment of biomarkers and genetic variations, using the efficient case–cohort design. Table 3 shows the distribution of several characteristics of the random sample compared with the total cohort. No substantial differences were detected.

Several established biochemical risk factors for cardiovascular disease and type 2 diabetes were assessed in the random sample and in all incident cases of coronary heart disease, cerebrovascular accidents and type 2 diabetes that occurred before 2006. Blood levels of established biochemical parameters were measured in EDTA or citrate plasma. We compared measurements using EDTA plasma with citrate plasma and validated these against serum in a sample of 50 participants. Correlation coefficients between EDTA or citrate plasma and serum were generally higher than 0.95 with the exception of albumin ($r = 0.82$). HbA1c was measured in erythrocytes using an immunoturbidimetric latex test. Albumin and creatinine (Jaffé method) were measured using a colorimetric method. Alanine aminotransferase, aspartate aminotransferase, gamma-glutamyltransferase, total cholesterol, triglycerides and uric acid were measured using enzymatic methods, whereas high-sensitive C-reactive protein was measured with a turbidimetric method. High-density lipoprotein (HDL)-cholesterol and low-density lipoprotein (LDL)-cholesterol were measured using a homogeneous assay with enzymatic endpoint. These assays, including the haemolytic, icteric and lipemic indices (absorbance), were all performed on an autoanalyser (LX20, Beckman Coulter, Mijdrecht, the Netherlands). Table 4 describes the biochemical measurements in the random sample, and of cases with cardiovascular disease or type 2 diabetes.

Table 3 Baseline characteristics of the 6.5% random sample compared with the entire baseline cohort

	Cohort	Random sample
N	40 011	2604
Age (years)	49 ± 12	49 ± 12
Sex (% female)	74.4	74.7
BMI (kg/m^2)	25.7 ± 4.1	25.8 ± 4.1
Systolic blood pressure (mm Hg)	126.2 ± 19.0	126.3 ± 18.9
Diastolic blood pressure (mm Hg)	77.8 ± 10.7	77.9 ± 10.7
Waist circumference (cm)	85.3 ± 11.6	85.5 ± 11.7
Hip circumference (cm)	103.5 ± 8.3	103.5 ± 8.2
Smoking		
Current (%)	37.7	38.5
Former (%)	31.2	31.6
Never (%)	38.3	37.7
Alcohol intake (g/day)	10.9 ± 15.5	11.2 ± 15.8
Cancer (%)	4.2	4.1
Myocardial infarction (%)	1.8	1.7
Diabetes (%)	2.1	1.8
Hypertension (%)	21.9	23.1
Hyperlipidaemia (%)	8.4	8.1

Disease occurrence

During follow-up, we obtained occurrence of chronic diseases through linkage with several registries. Table 5 describes the event rates of certain main endpoints obtained in our cohort.

Follow-up questionnaires

The follow-up questionnaires contained questions on occurrence of diabetes, hypertension, hyperlipidaemia, fractures, changes of body weight and lifestyle, such as physical activity, alcohol consumption, smoking and reproductive factors.

What has been found?

EPIC-NL is the Dutch contribution into the European EPIC study and as such has participated in over 150 cancer-related studies.[17–20] Studies only based on the Dutch data so far included Prospect or MORGEN data separately with a few exceptions.[21] Examples include relations between established risk factors with occurrence of chronic disease in the separate cohorts.[11,22–25] For example, a positive association between dietary glycaemic load and glycaemic index with cardiovascular disease was reported [hazard ratio (HR): 1.5], and

Table 4 Mean (± SD) plasma concentrations of established cardiovascular risk factors in the random sample, cardiovascular and type 2 diabetes cases

	CVD ($n = 1836$)	DM2 ($n = 822$)	Random sample ($n = 2389$)
Mean (SD)			
Albumin (g/l)	37.7 (4.9)	37.1 (4.8)	38.7 (4.9)
Creatinine (mg/l)	64.9 (18.6)	62.2 (14.7)	63.2 (16.3)
Cholesterol (mmol/l)	5.7 (1.1)	5.4 (1.1)	5.3 (1.1)
HDL (mmol/l)	1.2 (0.3)	1.0 (0.3)	1.3 (0.3)
LDL (mmol/l)	3.4 (0.9)	3.3 (0.9)	3.1 (0.9)
Uric acid (μmol/l)	271.2 (73.0)	285.8 (72.0)	257.8 (68.3)
Hb (g/dl)	17.2 (3.3)	17.1 (3.1)	17.3 (3.35)
Median (Q1–Q3)			
ALT (IU/l)	14.8 (12.0–19.0)	17.2 (13.6–22.9)	14.6 (11.9–18.4)
AST (IU/l)	20.1 (17.4–24.1)	20.7 (17.5–25.8)	20.0 (17.4–23.6)
GGT (IU/l)	24.5 (19.2–33.5)	29.1 (22.8–40.9)	20.5 (16.4–28.0)
hsCRP (mg/l)	1.9 (0.9–3.8)	2.6 (1.3–4.8)	1.29 (0.6–2.9)
Triglycerides (mmol/l)	1.6 (1.1–2.4)	2.0 (1.4–2.8)	1.31 (0.92–1.96)
HbA1c (%)	5.7 (5.2–6.2)	6.3 (5.6–6.9)	5.4 (5.0–5.8)

an inverse association between alcohol consumption and risk of type 2 diabetes (HR: 0.7).[11,22] In the Prospect cohort, dietary intake of phytoestrogens was not associated with risk of coronary heart disease,[25] whereas plasma phytoestrogens were associated with a reduced risk of breast cancer (HR: 0.7).[26] The Prospect cohort also showed an increased post-menopausal breast cancer risk with increased number of menstrual cycles (HR: 1.8).[23] Furthermore, post-menopausal hormone therapy was associated with a smaller decline of mammographic density.[27] In the MORGEN cohort, time-trends in risk factors showed an increase of total cholesterol whereas HDL-cholesterol remained stable[28] and a steady increase in the prevalence of obesity.[29] Another study provided evidence that overweight individuals give biased dietary information.[30] In the MORGEN cohort, smoking and coffee consumption were positively and alcohol drinking was inversely associated with plasma total homocysteine level.[31] Two studies also showed that a high intake of vitamin C or beta-carotene is protective for forced expiratory volume compared with a low intake, but not for respiratory symptoms,[32] and that flavonol and flavone intake was inversely associated with chronic cough (odds ratio = 0.8), whereas solid fruit, but not tea, intake was inversely associated with chronic obstructive pulmonary disease.[33]

Main strengths and weaknesses

Important strengths of the EPIC-NL study include inclusion of both genders, a large variation in disease determinants and its unique repository of blood samples collected before the occurrence of the chronic diseases of interest kept under liquid nitrogen for long-term use. The EPIC-NL study has a sufficient sample size to investigate obesity, cardiovascular disease and type 2 diabetes as endpoints by itself. However, the EPIC-NL study does not have sufficient power for most cancer endpoints, at least not in the coming one to two decades. The EPIC-NL study is an ongoing study with continuous and almost complete follow-up for disease occurrence. The number of endpoints is steadily increasing over the years (Table 5). The higher disease rates in Prospect are in part explained by higher ages at baseline. A limitation of the EPIC-NL study is a slight difference in some measurements of specific factors. For example, blood pressure measurements were different in MORGEN and Prospect. Therefore, for some studies analyses stratified by cohort is recommended. Finally, like many observational studies, our data are in part based on self-reported data that may be subject to misclassification.

How can I collaborate?

We welcome collaborative research on the data of the EPIC-NL study. For more information, you can check our website at: www.epicnl.eu; or mail us at: info@epicnl.eu. Otherwise, you can directly contact the National Institute for Public Health and the Environment, Division Public Health, Centre for Prevention and Health Services Research, Dr H.B. Bueno-de-Mesquita, principal investigator (Bas.Bueno.De.Mesquita@rivm.nl) or the University Medical Center Utrecht, Julius Center for Health Sciences and Primary Care, Prof. Dr P.H.M. Peeters, prinicipal investigator (p.h.m.peeters@umcutrecht.nl).

Table 5 Incidence rates of total, cardiovascular and cancer mortality, and cancer and cardiovascular morbidity

	MORGEN			Prospect		EPIC-NL	
	Cases/ person-time	Rate (per 1000 person-years)	Cases/ person-time	Rate (per 1000 person-years)	Cases/ person-time	Rate (per 1000 person-years)	Cases/ person-time	Rate (per 1000 person-years)
Cancer	287/82 734	3.5	379/99 795	3.8	1417/161 223	8.8	2083/344 237	6.1
Breast	NA	NA	156/99 795	1.6	543/161 223	3.4	699/344 237	2.0
Colon	16/82 734	0.2	21/99 795	0.2	145/161 223	0.9	182/344 237	0.5
Lung	41/82 734	0.5	30/99 795	0.3	91/161 223	0.6	162/344 237	0.5
Endometrium	NA	NA	17/99 795	0.2	77/161 223	0.5	94/344 237	0.3
Prostate	58/82 734	0.7	NA	NA	NA	NA	58/344 237	0.2
Cardiovascular disease	627/961 60	6.5	398/12 0737	3.3	1005/160 284	6.2	2030/372 860	5.4
Coronary heart disease	511/970 30	5.3	271/12 686	2.1	710/164 043	4.3	1492/380 386	3.9
Stroke	133/100 867	1.3	133/12 2498	1.1	324/166 608	1.9	590/385 623	1.5
Ischaemic stroke	104/100 643	1.0	81/122 676	0.7	234/166 862	1.4	419/386 172	1.1

NA = not applicable.

Funding

The EPIC-NL study was funded by 'Europe against Cancer' Programme of the European Commission (SANCO); the Dutch Ministry of Public Health, Welfare and Sports (formerly Ministry of Welfare, Public Health and Culture); the Dutch Cancer Society; ZonMW the Netherlands Organisation for Health Research and Development; and World Cancer Research Fund (WCRF).

Acknowledgements

We thank GBA, Statistics Netherlands, Dutch Cancer Registry and the institute PHARMO for providing data on vital status, cause of death and occurrence of cancer and other chronic diseases. Finally, we would like to thank Jan van der Laan, Anneke Blokstra, Robert Jan de Klein, Ido Toxopeus and Bernard Slotboom for their contribution to setting up this cohort.

Conflict of interest: None declared.

References

[1] Bingham S, Riboli E. Diet and cancer – the European Prospective Investigation into Cancer and Nutrition. *Nat Rev Cancer* 2004;**4:**206–15.

[2] Riboli E. Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann Oncol* 1992;**3:**783–91.

[3] Boker LK, van Noord PA, van der Schouw YT *et al*. Prospect-EPIC Utrecht: study design and characteristics of the cohort population. European Prospective Investigation into Cancer and Nutrition. *Eur J Epidemiol* 2001;**17:**1047–53.

[4] Smit HA, Verschuren W, Bueno-de-Mesquita HB, Seidell JC. Monitoring van Risicofactoren en Gezondheid in Nederland (MORGEN-project): Doelstellingen en werkwijze. Bilthoven: RIVM; 1994. Report No.: 263200001.

[5] Danesh J, Saracci R, Berglund G *et al*. EPIC-Heart: the cardiovascular component of a prospective study of nutritional, lifestyle and biological factors in 520,000 middle-aged participants from 10 European countries. *Eur J Epidemiol* 2007;**22:**129–41.

[6] Van Loon AJ, Tijhuis M, Picavet HS, Surtees PG, Ormel J. Survey non-response in the Netherlands: effects on prevalence estimates and associations. *Ann Epidemiol* 2003;**13:**105–10.

[7] Verschuren WM, Blokstra A, Picavet HS, Smit HA. Cohort profile: the Doetinchem Cohort Study. *Int J Epidemiol* 2008;**37:**1236–41.

[8] van der Sanden GA, Coebergh JW, Schouten LJ, Visser O, van Leeuwen FE. Cancer incidence in The Netherlands in 1989 and 1990: first results of the nationwide Netherlands cancer registry. Coordinating Committee for Regional Cancer Registries. *Eur J Cancer* 1995;**31A:** 1822–29.

[9] Nederlandse Kankerregistratie. http://www.ikcnet.nl/page.php?id=160 2009 Available from: URL: http://www.ikcnet.nl/page.php?id=160.

[10] Herings RM, Bakker A, Stricker BH, Nap G. Pharmacomorbidity linkage: a feasibility study comparing morbidity in two pharmacy based exposure cohorts. *J Epidemiol Community Health* 1992;**46:**136–40.

[11] Beulens JW, Stolk RP, van der Schouw YT, Grobbee DE, Hendriks HF, Bots ML. Alcohol consumption and risk of type 2 diabetes among older women. *Diabetes Care* 2005;**28:**2933–38.

[12] Schulze MB, Kroke A, Saracci R, Boeing H. The effect of differences in measurement procedure on the comparability of blood pressure estimates in multi-centre studies. *Blood Press Monit* 2002;**7:**95–104.

[13] Haftenberger M, Schuit AJ, Tormo MJ *et al*. Physical activity of subjects aged 50-64 years involved in the European Prospective Investigation into Cancer and Nutrition (EPIC). *Public Health Nutr* 2002;**5:**1163–76.

[14] Wareham NJ, Jakes RW, Rennie KL *et al*. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutr* 2003;**6:**407–13.

[15] Ocke MC, Bueno-de-Mesquita HB, Goddijn HE *et al*. The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire, and relative validity and reproducibility for food groups. *Int J Epidemiol* 1997;**26 (Suppl 1):** S37–48.

[16] Ocke MC, Bueno-de-Mesquita HB, Pols MA, Smit HA, Van Staveren WA, Kromhout D. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int J Epidemiol* 1997;**26 (Suppl 1):** S49–58.

[17] Gonzalez CA, Jakszyn P, Pera G, Agudo A *et al*. Meat intake and risk of stomach and esophageal adenocarcinoma within the European Prospective Investigation Into Cancer and Nutrition (EPIC). *J Natl Cancer Inst* 2006;**98:** 345–54.

[18] Pischon T, Boeing H, Hoffmann K *et al*. General and abdominal adiposity and risk of death in Europe. *N Engl J Med* 2008;**359:**2105–20.

[19] Trichopoulou A, Orfanos P, Norat T *et al*. Modified Mediterranean diet and survival: EPIC-elderly prospective cohort study. BMJ 2005;**330**:991.

[20] van Gils CH, Peeters PH, Bueno-de-Mesquita HB *et al*. Consumption of vegetables and fruits and risk of breast cancer. *JAMA* 2005;**293:**183–93.

[21] Waijers PM, Ocke MC, van Rossum CT *et al*. Dietary patterns and survival in older Dutch women. *Am J Clin Nutr* 2006;**83:**1170–76.

[22] Beulens JW, de Bruijne LM, Stolk RP *et al*. High dietary glycemic load and glycemic index increase risk of cardiovascular disease among middle-aged women: a population-based follow-up study. *J Am Coll Cardiol* 2007;**50:** 14–21.

[23] Chavez-MacGregor M, Elias SG *et al*. Postmenopausal breast cancer risk and cumulative number of menstrual cycles. *Cancer Epidemiol Biomarkers Prev* 2005;**14:**799–804.

[24] Van der AD, Marx JJ, Grobbee DE *et al*. Non-transferrin-bound iron and risk of coronary heart disease in postmenopausal women. *Circulation* 2006;**113:**1942–49.

[25] van der Schouw YT, Kreijkamp-Kaspers S, Peeters PH, Keinan-Boker L, Rimm EB, Grobbee DE. Prospective study on usual dietary phytoestrogen intake and cardiovascular disease risk in Western women. *Circulation* 2005;**111:**465–71.

[26] Verheus M, van Gils CH, Keinan-Boker L, Grace PB, Bingham SA, Peeters PH. Plasma phytoestrogens and subsequent breast cancer risk. *J Clin Oncol* 2007;**25:**648–55.

[27] van Duijnhoven FJ, Peeters PH, Warren RM *et al*. Postmenopausal hormone therapy and changes in mammographic density. *J Clin Oncol* 2007;**25:**1323–28.

[28] Houterman S, Verschuren WM, Oomen CM, Boersma-Cobbaert CM, Kromhout D. Trends in total and high density lipoprotein cholesterol and their determinants in The Netherlands between 1993 and 1997. *Int J Epidemiol* 2001;**30:**1063–70.

[29] Visscher TL, Kromhout D, Seidell JC. Long-term and recent time trends in the prevalence of obesity among Dutch men and women. *Int J Obes Relat Metab Disord* 2002;**26:**1218–24.

[30] Braam LA, Ocke MC, Bueno-de-Mesquita HB, Seidell JC. Determinants of obesity-related underreporting of energy intake. *Am J Epidemiol* 1998;**147:**1081–86.

[31] de Bree A, Verschuren WM, Blom HJ. Biological cardiovascular risk factors and plasma homocysteine levels in the general Dutch population. *Atherosclerosis* 2001;**154:** 513–14.

[32] Grievink L, Smit HA, Ocke MC, van't Veer P, Kromhout D. Dietary intake of antioxidant (pro)-vitamins, respiratory symptoms and pulmonary function: the MORGEN study. *Thorax* 1998;**53:**166–71.

[33] Tabak C, Arts IC, Smit HA, Heederik D, Kromhout D. Chronic obstructive pulmonary disease and intake of catechins, flavonols, and flavones: the MORGEN Study. *Am J Respir Crit Care Med* 2001;**164:**61–64.