British Journal of Cancer (1999) 80(Suppiement 1), 95–103

Article no. bjoc.1998.0518

EPIC-Norfolk: study design and characteristics of the cohort

N Day[1,2], S Oakes[1], R Luben[1,2], K-T Khaw[1], S Bingham[3], A Welch[1] and N Wareham[1]

[1]Department of Community Medicine, University of Cambridge, Cambridge, UK; [2]Medical Research Council, Biostatistics Unit, Cambridge UK; [3]Medical Research Council, Dunn Nutrition Unit, Cambridge UK

The European Prospective Investigation of Cancer (EPIC) was initiated in 1989 in order to define more clearly the relationship between diet and the risk of developing cancer at a range of sites (Riboli, 1992). The hypothesis that diet might be responsible for a substantial proportion of total cancer incidence was widely publicized by a report in the early 1980s (Doll, 1981) which suggested that 30% of cancer in the USA might be preventable by dietary modification. However, the uncertainty bounds around the estimate of 30% were acknowledged to be wide, and no association between a specific dietary constituent and risk for a specific cancer type was considered conclusively established. Since then, evidence has continued to accrue supporting the general association between diet and cancer, but two recently published comprehensive reviews only found sufficient evidence for a few specific relationships (World Cancer Research Fund, 1997; Department of Health, 1998). The most convincing evidence is for a protective role for fruit and vegetables for a number of cancers, including oesophagus, stomach, large bowel, lung and breast. It is not clear, however, that the same food constituents are involved for each cancer, nor even that the same type of fruit and vegetable are implicated.

The rationale for EPIC was that in order to generate data which improved the definition of diet–cancer associations, studies had to be prospective, large-scale and based on populations with wide variation both in diet and in cancer incidence. In addition, studies had to use improved methods for dietary assessment including biological markers. EPIC comprises cohorts established in France, Germany, Greece, Italy, The Netherlands, Spain and the UK, with associated cohorts in Denmark, Norway and Sweden. Recruitment of these cohorts with baseline data and biological samples is now almost complete (as shown in Table 1), achieving the target set earlier (Riboli, 1992). As an example of the range in cancer rates seen across the populations participating in EPIC, Figure 1 shows mortality rates for breast cancer.

The study in Norfolk was initially planned as a diet and cancer cohort and later joined as a collaborative component cohort of EPIC. It has now, however, broadened its scope to include end points other than cancer, and exposures other than diet. The intention now is to investigate the majority of causes of disability and death in middle and later life, and to include additional lifestyle exposures, notably exercise, physical activity and psychosocial variables.

Correspondence to: N Day, Strangeways Research Laboratories, Worts Causeway, Cambridge CB1 4RN, UK

EPIC-NORFOLK: STUDY DESIGN

Cohort definition

The intention in the Norfolk EPIC study was to identify a cohort of 25 000 men and women aged 45–74 from the general population, from a geographically circumscribed area which has little outward migration in this age group and which is mainly served by one District General Hospital. The choice of 25 000 as the target for recruitment represents a compromise between the need for large numbers to generate sufficient end points, and the need to include better defined and more discriminating instruments for assessing exposure, including biological assays. Although some cohort studies emphasize the former, very large studies usually only include imprecise measures of exposure, which can lead to uncertainty about the nature of the exposure–disease relationships. As the focus of the Norfolk EPIC study is on increasing the precision with which exposures are measured, the size of the cohort was limited to allow these more detailed measurements to be undertaken. The city of Norwich and the surrounding small towns and rural areas were chosen as the study area, and local General Practices were approached and invited to participate. The geographical area covered by the 35 participating practices (121 General Practitioners) is shown in Figure 2. All individuals in the age range in each General Practice database were invited to participate, except those marked as unsuitable by the General Practitioner, and those who consented were invited to attend for a health check. The process of recruitment, together with the number of people at each stage, is shown in Figure 3. Recruitment started in March 1993 and was completed at the end of 1997. The pattern of cumulative accrual is shown in Figure 4.

Initial instruments used for exposure definition

Non-dietary data

The Health and Lifestyle questionnaire has a common format across the EPIC cohorts and includes questions on smoking, alcohol consumption, socio-economic status, social class, occupational history, past history of disease, short family history of main disease end points, reproductive history (for women) and a short section on exercise.

Dietary data

A semi-quantitative food frequency questionnaire (FFQ) was adopted as the primary dietary instrument across the EPIC cohorts. In order to achieve finer between-individual discrimination within the Norfolk cohort, additional instruments have been included in

Table 1 Subject recruitment to the total EPIC cohort by country as of March 1998

	Subjects included in the study with:		Additional subjects to be recruited	End of subject recruitment
	Questionnaire	Blood collection		
Spain	41 552	39 792	–	Completed 1996
Italy	43 822	43 822	–	Completed Feb 1998
UK	76 041	40 901	–	Completed Feb 1998
The Netherlands	40 659	36 885	–	Completed Feb 1998
France	72 000	14 000	–	Completed 1993
Germany	44 055	44 055	9 000	Nov 1998
Greece	22 516	22 516	4 000	Nov 1998
Total 7 countries	340 645	241 971	13 000	
Associated projects: Sweden	57 932	62 122	–	Completed 1996
Denmark	57 174	56 800	–	Completed 1997
Total 9 countries	455 751	360 893	13 000	

Norfolk. Validation studies undertaken on this cohort clearly demonstrated that, for many food groups and dietary constituents, dietary diaries were superior to FFQs (Bingham et al, 1994). Furthermore, in order to study the possible importance of seasonal variation in food consumption in determining an individual's risk (Powles et al, 1997), diaries completed in different seasons would be a more direct approach. The dietary data, therefore, consisted of an FFQ and a 24-h recall completed before the health check, and a 7-day food diary given out at the health check after instruction, which was completed and returned by post (93% compliance). The first day of the diary was an interviewed 24-h recall. A second 7-day diary was posted out 18 months after the health check (75% compliance).

Health check

Table 2 gives details of the examination and samples taken at the Health Check, which included respiratory function, anthropometry, blood pressure and urine testing. Calibration was undertaken regularly to check the accuracy of both the equipment and the operators.

Sample collection and processing

A total of 42 ml of blood (not fasting, collected in EDTA, citrated and plain Monovettes) and a spot urine sample were taken at the health check. The citrated and plain blood samples and the urine were kept overnight in a refrigerator, the EDTA samples at room temperature. Early the following morning they were transported to the EPIC laboratory in Norfolk for processing. For each participant 28 (0.5 ml) straws were filled; 12 with plasma, eight with serum, four with red cells and four with buffy coat. A detailed description of the straw storage method has been given earlier (Riboli, 1992). The straws remain stored in liquid nitrogen at –196°C, half at the Norfolk laboratory and half at the International Agency for Research on Cancer, which coordinates the EPIC Study overall. The remainder of the blood samples taken on each individual were transported to Cambridge for immediate assays, as indicated in Table 3. Not all assays were performed at the start of the study, so that the number on whom each assay has been performed varies. Details of the assays have been given earlier (Ness et al, 1996).

Follow-up methods

A continuing system for ascertaining health end points from a variety of sources has been established. The entire cohort has been flagged with the NHS Central Register for death and cancer incidence. The cohort is regularly matched against the East Anglian Cancer Registry, which provides more immediate ascertainment of incident cancer cases than the Central Register. The cohort is also matched to the district database of all in-patient hospital activity. Each cohort member is also followed actively with a postal health questionnaire sent 18 months after the health check, and again at 36 months. It is planned to repeat the postal follow-up every 18 months in the future and to investigate the use of General Practice records for health end points.

DEVELOPMENT BEYOND THE INITIAL COLLECTION OF BASELINE DATA

The aim of EPIC in Norfolk is to quantify the contribution of different determinants of chronic disease in middle and later life. The emphasis has been on the development of better validated instruments, with more control of the effects of measurement error. This is as important for variables whose main role in analyses will be as potential confounders, as it is for exposures which are the principal focus of interest (Marshall, 1996). For this reason, particular effort has been put into the development of instruments for assessing physical activity and psychosocial variables, and to obtain, where feasible, repeat measures of exposure on the entire cohort, especially for the biological parameters.

SECOND HEALTH CHECK

In contrast to the majority of other cohorts in EPIC, the entire Norfolk cohort is being invited back for a second health check. This process began in January 1998, with initial compliance of 65%. In addition to the measures taken at the first health check, participants undergo ultrasound measurement of heel bone density (McCue, UK), and a bioelectrical impedance test for measuring body composition (Tanita, UK).

Psychosocial measures

Eighteen months after the first health check and immediately after the mailing of the second 7-day diary, each participant is sent a Health and Life Experience Questionnaire (HLEQ). The HLEQ includes a structured self-assessment approach to psychiatric symptoms embodying the rules of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association,

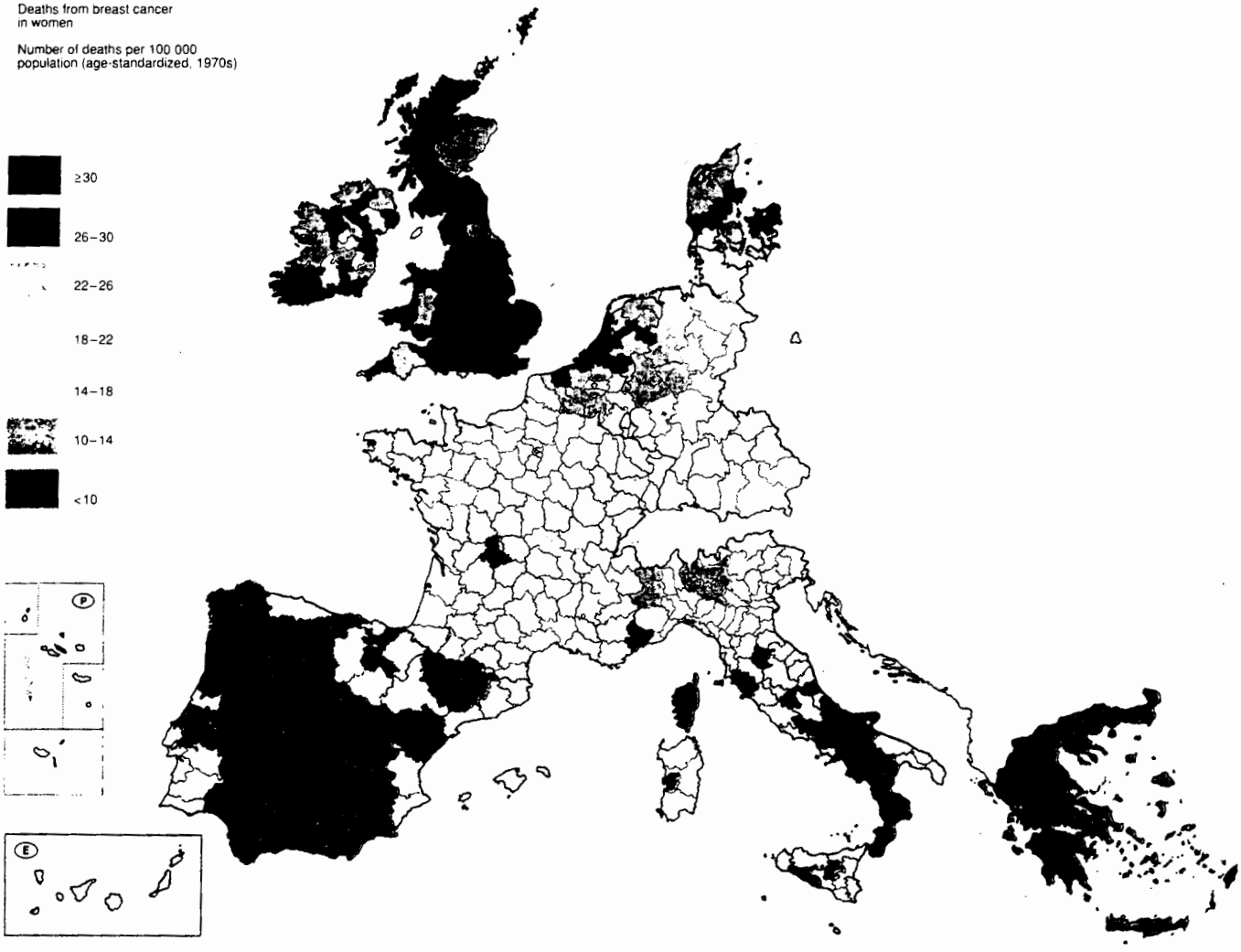

Figure 1 Mortality rates from breast cancer by administrative region in the 12 countries of the European Union (Commission of the European Communities, Directorate-General for Employment, Industrial Relations and Social Affairs (1991)

1994). However, given evidence on prevalence, this was limited to major depressive disorder and generalized anxiety disorder only. The approach is a development from the short-form scales of the formal structured assessment methods derived from the National Comorbidity Survey (Kessler et al, 1994) and from prior experience with the methods of psychiatric research diagnostic systems (Surtees et al, 1997). The novel assessment method is intended to identify those thought likely to meet putative diagnoses (anxiety or depressive) at any time in their lives and to provide evidence of chronicity.

Exercise and energy expenditure

At the second health check, participants are given a Physical Activity questionnaire (EPAQ2). This questionnaire has been developed from validation studies based on 4-day Heart Rate Monitoring calibrated against doubly labelled water measurements (Wareham et al, 1997). This questionnaire gives a substantially more comprehensive description of physical activity than has been obtained from previous studies, as it includes sections not only about occupational and leisure time sports and recreational activity, but also about activities in and around the home. It is aimed at estimating various dimensions of activity, including usual daily total energy expenditure, the period of vigorous physical activity and physical fitness.

FUTURE RESEARCH DEVELOPMENTS

Cancer genetics and the study of environment–gene interactions

Genetic modification of the risk associated with specific exposures will form an important aspect of future work, since identification of these interactions will provide clearer quantification of effect, a better understanding of mechanisms and additional approaches to prevention. In collaboration with the Department of Oncology at Cambridge, the EPIC cohort will form the basis for a range of studies to identify low penetrance genes by association studies (Dunning et al, 1998) and to estimate the frequency of high penetrance mutations in the general population (Dunning et al, 1997).

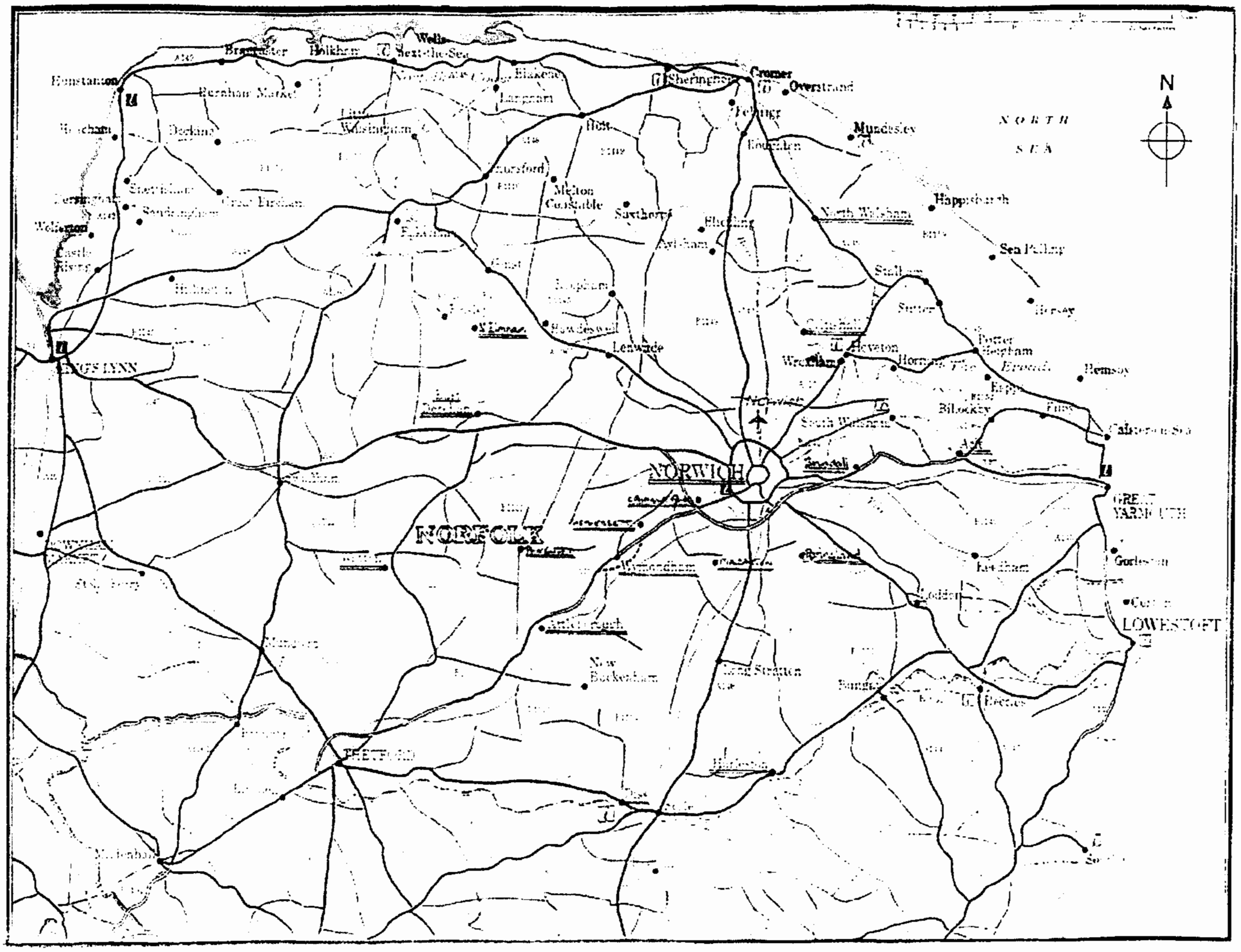

Figure 2 Map of Norfolk showing the areas covered by the General Practices participating in EPIC in Norfolk

Both the Norfolk cohort and EPIC more generally should form a powerful platform for later studies of environment–gene interaction. For example, dietary effects may be modulated by inherited polymorphisms in a number of genes governing nutrient and xenobiotic metabolism, such as genes encoding phase I and phase II enzymes in cancer (Idle, 1991; Vineis, 1997). Meat is associated with increased risk of cancer especially in relation to polymorphisms of the *N*-acetyl-transferases (NATs) (Lang et al, 1994; Roberts-Thompson et al, 1996). CYP2E1 controls *N*-nitrosamine activation and is likely to be important in interaction with individual propensity for colonic *N*-nitrosation which occurs in response to meat eating (Bingham et al, 1996). In hormone-dependent cancers, CYP11a, CYP17 and CYP19 polymorphisms control the synthesis and metabolism of androgens, cholesterol side-chain cleavage and aromatase (Gharani et al, 1997). Polymorphisms in methylenetetrahydrofolate reductase and vitamin D receptor also have been correlated with circulating levels of active metabolites and cancer risk (Taylor et al, 1996; Ma et al, 1997). The design of the Norfolk EPIC study will facilitate further studies to unravel the complex mechanisms linking dietary exposure with cancer.

The measurement of intermediate markers of disease progression

As part of this process of generating increased understanding about the pathogenesis of disease, a number of early markers of disease risk have been included in EPIC by linking the study to existing early detection programmes in the region. The majority of women aged 50–64 in the Norfolk cohort have attended the NHS Breast Screening Clinic based in Norwich. Breast density, which can be assessed mammographically, is an indicator of risk of developing breast cancer in the future (Byrne, 1997). The mammographic patterns from a subsample of the Norfolk cohort are being assessed, and will be linked to dietary and other exposure variables. Norwich is also one of the centres for the National Multi-Centre Flexible Sigmoidoscopy screening trial (Atken et al,

1993). Approximately 2000 of the EPIC cohort aged 55–64 have undergone sigmoidoscopy, and adenomas detected in slightly over 10%. Both adenomatous and normal tissue is being examined for DNA adducts.

THE ANALYTIC APPROACH IN EPIC

For those variables for which data are available on the entire cohort, Poisson regression will be the main analytic approach. However, this will only be possible for a limited number of analyses. Many of the analyses relating exposures to disease risk will use a nested case-control approach (Breslow, 1988). This is mainly because the complexity of the exposure data, particularly the dietary diaries, makes data entry for the entire cohort uneconomic. In addition, it is difficult to decide in advance which precise assays will be needed for each disease end point. Most assays of the biological samples will be done on stored material at the time of disease occurrence. The alternative to a nested case-control approach is to use a subcohort for comparison with the cases. For biological assays this approach is not appropriate since length of storage needs to be taken into account, and matched on. Some initial cross-sectional analyses will be based on a subsample of the entire cohort, and it is possible that for variables where passage of time is not a problem, e.g. dietary diaries, information from this subcohort may supplement the nested case-control analyses.

EPIC-NORFOLK: INITIAL COHORT DESCRIPTION

Non-dietary variables

The process of entering and verifying the non-dietary data has followed closely from recruitment into the study. Univariate descriptive tables are presented for the 11 996 individuals (6553 women, 5443 men) recruited and seen by the end of October 1995. Table 4 gives the age and sex distribution of this sample with the proportion of current cigarette smokers, former smokers and never smokers. Data from the Health Survey for England are included for comparison. Table 5 summarizes the anthropometric, blood lipid and blood pressure data. Blood lipid levels, particularly the high density lipoprotein cholesterol and triglycerides, and both the systolic and diastolic blood pressure increase more sharply with age in women than in men. Table 6 gives the corresponding results from the Health Survey for England 1993 (Bennett et al, 1995)

Initial end point results

Table 7 shows the number of deaths by major cause among Norfolk EPIC participants notified by the ONS up to 31 January 1998. The average follow-up time for individuals recruited by March 1997 (n = 21 453) was approximately 2 years. Table 8 shows the frequency of cancers registered with the East Anglian Cancer Registry by the end of January 1998 among those recruited by March 1997. It is noteworthy that lung cancer is only the sixth most common type of cancer, reflecting the low prevalence of current tobacco smoking in the cohort. The pattern of cancer incidence seen in Table 8 is perhaps an illustration of the pattern to be expected in Western populations after the smoking epidemic has waned, in which dietary factors may take a more dominant role.

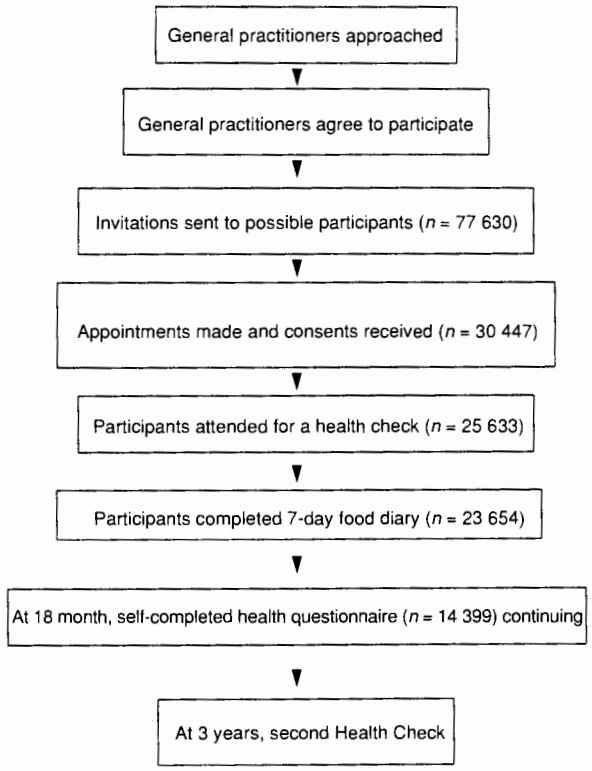

Figure 3 The process of recruitment into the Norfolk EPIC cohort as of April 1998

DISCUSSION

EPIC-Norfolk has evolved from a study aimed principally at elucidating the diet–cancer association, to a more general programme of research into the factors determining ill health in middle and later life. To assess the public health impact of lifestyle intervention based on observational epidemiological data, it is important to study the full range of health end points, including cardiovascular disease, diabetes and osteoporosis. This is increasingly becoming the central aim of EPIC in Norfolk.

In comparison with the general population of England, this cohort has fewer current smokers. It is therefore well placed to provide an unclouded view of risks associated with factors other than tobacco. In terms of anthropometric variables, blood pressure and serum lipids, this cohort is representative of the population studied in the Health Survey of England.

In the development of this cohort study, major emphasis has been placed on producing better epidemiological instruments, and on determining the real relationship between the observed exposure and the relevant underlying true exposure variable. Emphasis has also been put on collecting biological markers of exposure, on measuring early disease and disease progression, and making repeat measures where possible. The full uncertainty associated with risk estimates derived from observational epidemiology has seldom been adequately addressed, except in relatively simple situations (MacMahon et al, 1990). Failure to evaluate the full uncertainty will lead to misleading and contradictory conclusions, particularly when the exposures under study are complex, such as diet and physical activity. In these situations, much of the uncertainty is not due to random noise and, therefore, cannot simply be removed by increasing the sample size (Phillips, 1993). EPIC-Norfolk has

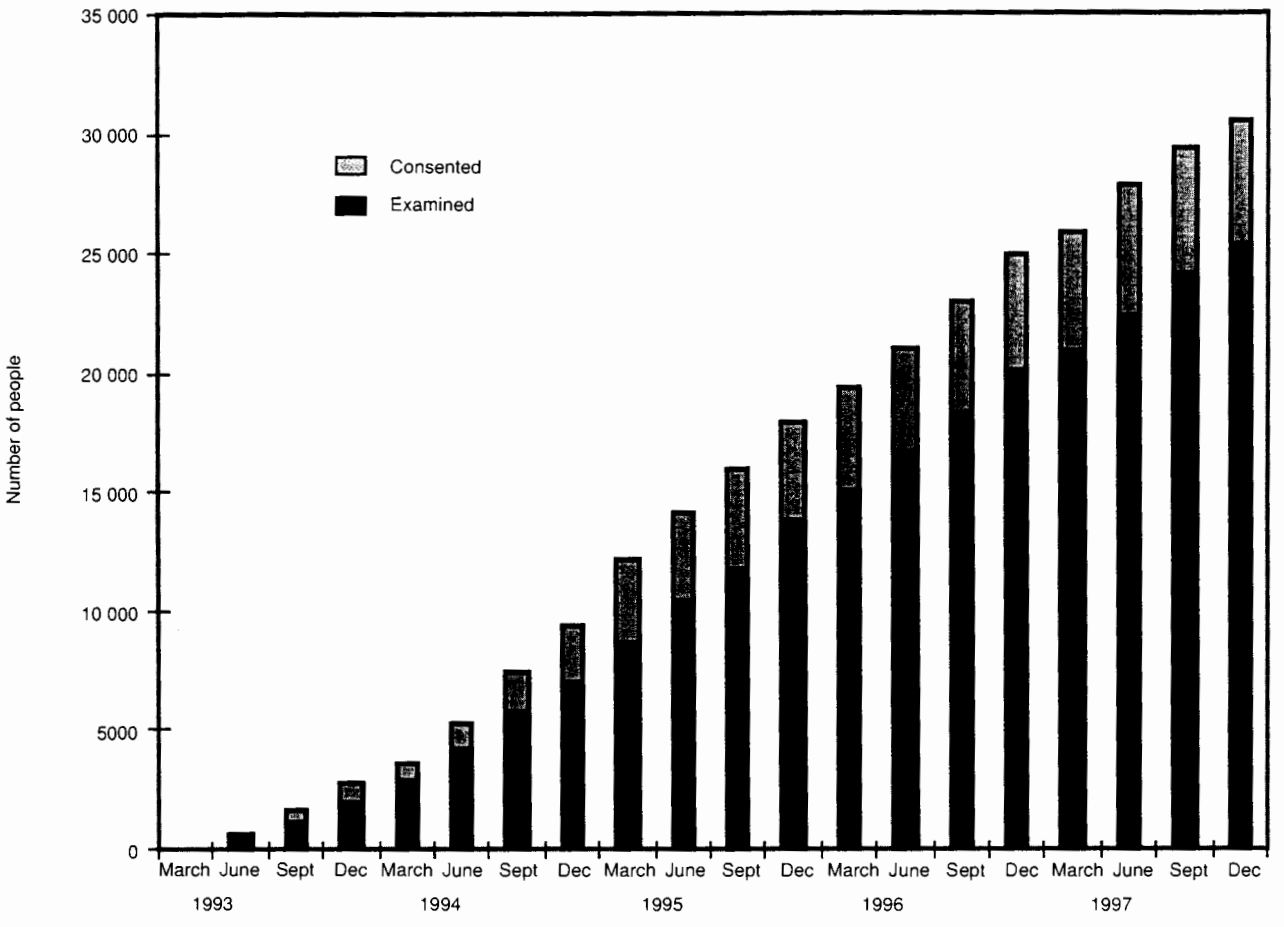

Figure 4 Cumulative recruitment into the Norfolk EPIC cohort showing the numbers who consented to take part and the numbers who presented for a health check

Table 2 Procedures for the Health Check in the Norfolk EPIC Study

Measurement	Description
Height (Lohman et al, 1991)	Measured to the nearest millimetre, without shoes using free-standing stadiometer
Weight (Lohman et al, 1991)	Measured to the nearest 0.2 kg, without shoes in light clothing using digital scales (Salter, UK)
Body circumferences (Lohman et al, 1991)	All measured to the nearest millimetre at end of expiration, in light clothing with a D loop non-stretch fibreglass tape
Chest	Defined as maximum circumference at the level of the nipple
Waist	Defined as either the minimum circumference at the natural waistline between the lower rib margin and the iliac crest (with belts moved) or, when the minimum circumference was not identifiable, as the circumference at the level of the navel
Hip	Maximum circumference between the iliac crest and the crotch (greatest of two measurements)
Blood pressure	Two measurements taken of diastolic and systolic blood pressure using Accutorr Sphygmomanometer (Datascope, UK), after participant has sat for 3 min. Arm held horizontal at the level of the mid-sternum. Medium or large cuff size used according to arm circumference
Urinalysis	Spot urine stored for further analysis. Immediate urinalysis performed with Multistix 8SG (Bayer Diagnostics, UK)
Spirometry (Cox et al, 1993)	FEV_1, FVC and PEFR measured using portable spirometer (MicroMedical, UK)

Table 3 Assays performed on blood samples from all EPIC-Norfolk participants who underwent a health check

Assay	Method
Full blood count	Measured on Coulter MD Haemolyser
Total cholesterol, HDL cholesterol, triglyceride	Measured on RA 1000 (Bayer Diagnostics, Basingstoke)
LDL cholesterol	Calculated using Friedewald formula (Friedewald et al, 1972) except where triglyceride > 4 mmol l^{-1}
HbA1c	Measured using Diamat ion exchange HPLC (Bio-Rad Laboratories, Hemel Hempstead, UK) (Standing et al, 1992)
Free T4, TSH	Measured using AutoDELFIA time resolved fluoroimmunoassay kits (Wallac, Finland)
Vitamin C	Fluorometric assay (Vuilleumier, 1989)

Table 4 Prevalence of cigarette smoking among participants recruited to the EPIC in Norfolk cohort by end October 1995 (n=11 996), by age and sex with comparison to the Health Survey for England, 1993

	Men			Women		
Age group	**45–54**	**55–64**	**65–74**	**45–54**	**55–64**	**65–74**
Norfolk EPIC Cohort						
n	1847	1760	1836	2358	2166	2029
Never smoked (%)	38.9	35.0	22.4	56.2	62.1	55.2
Ex-smoker (%)	46.3	54.0	67.7	29.7	27.2	37.0
Current smoker (%)	14.8	11.0	9.8	14.2	10.7	7.8
Health Survey for England						
Never smoked (%)	34	27	19	46	51	49
Ex-smoker (%)	39	49	61	26	24	34
Current smoker (%)	28	25	20	27	25	18

Data for the Health Survey for England from Bennett et al, 1995.

Table 5 Anthropometric, blood pressure and total cholesterol measurements for participants in the Norfolk EPIC cohort seen by the end of October 1995

	Men			Women		
	45–54	**55–64**	**65–74**	**45–54**	**55–64**	**65–74**
n	1848	1766	1853	2363	2182	2052
Height (cm)	175 (6.63)	174 (7.91)	172 (7.11)	162 (7.92)	161 (7.53)	159 (6.01)
	165–186	164–185	161–183	153–172	152–171	149–169
Weight (kg)	80.6 (11.6)	80.9 (11.3)	78.9 (11.2)	67.7 (12.0)	69.4 (12.0)	67.1 (11.2)
	63.4–102	63.6–101	62.0–97.0	52.0–91.0	53.4–90.6	51.4–87.4
Body mass index (kg m^{-2}	26.2 (3.29)	26.7 (3.75)	26.7 (3.68)	25.7 (4.55)	26.7 (4.46)	26.6 (4.19)
	21.5–32.1	21.8–32.4	21.8–32.4	20.4–34.2	21.0–34.8	20.7–34.0
WHR	0.92 (0.13)	0.93 (0.07)	0.94 (0.06)	0.77 (0.06)	0.79 (0.07)	0.82 (0.07)
	0.82–1.01	0.83–1.03	0.85–1.05	0.69–0.88	0.70–0.90	0.72–0.92
Systolic blood pressure	131 (15.4)	137 (17.9)	144 (19.5)	126 (16.7)	135 (18.0)	143 (19.5)
(mmHg)	109–157	111–170	113–177	103–156	107–166	114–178
Diastolic blood pressure	82.7 (10.5)	84.3 (11.4)	84.8 (12.1)	77.8 (10.9)	81.4 (11.1)	83.6 (12.2)
(mmHg)	67.0–102	67.0–105	67.0–106	61.5–96.5	64.5–101	66.0–103
Total cholesterol	6.02 (1.09)	6.09 (1.11)	6.11 (1.11)	5.83 (1.05)	6.53 (1.18)	6.81 (1.23)
(mmol l^{-1})	4.30–7.90	4.40–8.10	4.40–8.00	4.30–7.60	4.70–8.60	4.90–8.90
HDL cholesterol	1.22 (0.32)	1.20 (0.32)	1.21 (0.33)	1.55 (0.41)	1.52 (0.40)	1.52 (0.41)
(mmol l^{-1})	0.80–1.80	0.80–1.80	0.80–1.80	1.00–2.30	1.00–2.20	0.90–2.30
LDL cholesterol	3.97 (1.16)	4.07 (1.22)	4.03 (1.15)	3.68 (1.07)	4.24 (1.15)	4.46 (1.18)
(mmol l^{-1})	2.46–5.67	2.49–5.76	2.48–5.76	2.27–5.36	2.62–6.16	2.71–6.39
Triglycerides (mmol l^{-1})	2.01 (1.27)	2.08 (1.28)	2.03 (1.13)	1.37 (0.82)	1.72 (0.97)	1.87 (1.10)
	0.70–4.40	0.80–4.30	0.80–4.10	0.60–2.90	0.70–3.60	0.80–3.70

Data shown are mean (s.d.) and 5th and 95th percentiles. WHR, waist hip ratio; HDL, high density lipoprotein; LDL, low density lipoprotein.

Table 6 Anthropometric, blood pressure and total cholesterol measurements for the population of England in 1993, taken from the Health Survey for England 1993 (Bennett et al, 1995)

	Men			Women		
Age group	**45–54**	**55–64**	**65–74**	**45–54**	**55–64**	**65–74**
Height (cm)	175 (163–186)	173 (162–184)	171 (160–182)	162 (152–172)	160 (150–171)	158 (148–168)
Weight (kg)	82 (62–104)	81 (61–103)	78 (58–99)	70 (52–95)	70 (52–94)	68 (50–90)
BMI (kg m^{-3})	26.8 (21.3–33.0)	27.0 (21.3–33.5)	26.7 (20.8–32.6)	26.6 (20.5–36.1)	27.2 (21.1–36.5)	27.0 (20.5–36.2)
Systolic blood pressure (mmHg)	138 (116–168)	146 (119–182)	152 (120–189)	134 (110–168)	144 (113–182)	156 (123–198)
Diastolic blood pressure (mmHg)	82 (65–102)	85 (68–106)	84 (65–104)	76 (58–96)	78 (60–96)	80 (61–106)
Total cholesterol (mmol l^{-1})	6.15 (4.5–8.1)	6.25 (4.6–8.2)	6.13 (4.3–8.1)	6.03 (4.4–7.9)	6.62 (4.8–8.7)	6.94 (5.1–8.9)

Data shown are mean (5th and 95th percentiles).

Table 7 Number of deaths by major cause in EPIC participants with a Health Check recruited before 31 October 1997 (n = 21 453), flagged with ONS and notified to EPIC by 31 January 1998

Cause of death	Men	Women	Total
Cancer	82	80	162
Cardiac	104	34	138
Cerebrovascular	13	19	32
Aneurysm	10	2	12
Other	34	36	70
Unknown	3	4	7
Total	246	175	421

Table 8 Number of incident cancer diagnoses for the 10 most common sites registered by the East Anglian Cancer Registry before 31 January 1998 among all EPIC participants who had undergone a Health Check by 31 March 1997 (n=21 453)

	Men	Women	Total
Non-melanoma skin	60	53	113
Breast	–	52	52
Colon and rectum	16	20	36
Prostate	28	–	28
Corpus uteri	–	18	18
Lung	11	3	14
Oesophagus	8	3	11
Malignant melanoma	3	7	10
Non-Hodgkin's lymphoma	4	6	10
Unknown primary site	4	5	9
Total	134	167	301

focused on controlling the uncertainties due to inadequately validated instruments for measuring exposure. For disease–exposure relationships where large numbers are required, the combined EPIC cohorts provide both numbers and variation in exposure.

For the more common sites of cancer, initial results from the combined EPIC cohorts should emerge by the end of 1999. By this time, over 1000 breast cancers, 700 large bowel cancers and 500 prostate cancers are expected. These results should begin to provide a more complete and fuller picture of diet–cancer associations than currently available. As follow-up continues into the early years of the next century, corresponding results for many of the less common cancers should be generated.

ACKNOWLEDGEMENTS

The EPIC-Norfolk Study is funded by the Cancer Research Campaign, Medical Research Council, British Heart Foundation, the Ministry of Agriculture, Fisheries and Food, and the Europe against Cancer Programme of the Commission of the European Communities. We are grateful to all the participants and general practitioners who have helped with this study, and to the nurses, technicians and the staff of the EPIC coordinating centre. Undergraduate students from Coleraine, Surrey and Wageningen Universities and Leeds Polytechnic assisted with the collection and interpretation of the field work.

REFERENCES

Atken WS, Cuzick J, Northover JMA and Whynes DK (1993) Prevention of colorectal cancer by once-only sigmoidoscopy. *Lancet* **341**: 736–740

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edn. American Psychiatric Association: Washington, DC

Bennett N, Dodd T, Flatley J, Freet S and Bolling K (1995) *Health Survey for England 1993*. HMSO: London

Bingham SA, Gill C, Welch A, Day K, Cassidy A, Khaw KT, Sneyd MJ, Key TJA, Roe L and Day NE (1994) Comparison of dietary assessment methods in nutritional epidemiology: weighed records v 24h recalls, food frequency questionnaires and estimated diet records. *Br J Nutr* **74**: 619–642

Bingham SA, Pignatelli B, Pollock J, Ellul A, Mallaveille C, Gross G, Runswick S, Cummings JH and O'Neill IK (1996) Does increased formation of endogenous N Nitroso compounds in the human colon explain the association between red meat and colon cancer? *Carcinogenesis* **17**: 515–523

Bingham S, Gill C, Welch A, Cassidy A, Runswick S, Sneyd M, Thurnham D, Key TJA, Roe L, Khaw K-T and Day NE (1997) Validation of dietary assessment methods in the UK arm of EPIC. *Int J Epid* **26**: S137–S151

Breslow NE and Day NE (1988) *Statistical Methods in Cancer Research: Vol II. The Design and Analysis of Cohort Studies 1988*. IARC Scientific Publications No 82. International Agency for Research on Cancer: Lyon

Byrne C (1997) Studying mammographic density: implications for understanding breast cancer. *J Natl Cancer Inst* **89**: 531–533

Commission of the European Communities. *Directorate-General for Employment, Industrial Relations and Social Affairs (1991) Europe against Cancer*. Public Health Initiatives and texts adopted in 1990. Office for Official Publications of the European Communities: Luxembourg

Cox BD, Huppert FA and Whichelow MJ (eds) (1993) *The Health and Lifestyle survey: 7 years on*. Dartmouth Publishing Co: Aldershot

Department of Health (1998) *Report on Health and Social Subjects 48. Nutritional Aspects of the Development of Cancer; Report of the Working Group on Diet and Cancer of the Committee on Medical Aspects of Food and Nutrition Policy*. HMSO: London

Doll R and Peto R (1981) The causes of human cancer: quantitative estimates of avoidable risks of cancer in the United States. *J Natl Cancer Inst* **66**: 1192–1308

Dunning AM, Chiano M, Smith NR, Dearden J, Gore M, Oakes S, Wilson C, Stratton M, Peto J, Easton D, Clayton D and Ponder BAJ (1997) Common BRCA1 variants and susceptibility to breast and ovarian cancer in the general population. *Hum Mol Gent* **6**: 285–289

Dunning AM, Healey CS, Pharoah PDP, Poster N, Lipscombe JM, Redman KL, Easton DF, Day NE and Ponder BAJ (1998) No association between a polymorphism in the steroid metabolism gene CYP17 and risk of breast cancer. *Br J Cancer* **77**: 2045–2047

Friedewald WT, Levy RI and Fredrickson DS (1972) Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* **18**: 499–502

Gharani N, Waterworth DM, Batty S, White D, Gillingsmith C, Conway GS, McCarthy M, Franks S and Williamson RNA (1997) Association of the steroid synthesis gene CYP11a with polycystic ovary syndrome and hyperandrogenism. *Hum Mol Genet* **6**: 397–402

Idle JR (1991) Is environmental carcinogenesis modulated by host polymorphism? *Mutat Res* **247**: 259–266

Kessler RC, McGonalgle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen H-U and Kessler KS (1994) Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. *Arch Gen Psychiatry* **51**: 8–19

Lang NP, Butler MA, Massengill J, Lawson M, Stotts C, Jensen MH and Kadlubar F (1994) Rapid metabolic phenotypes for acetyltransferase and p450 and putative exposure to HAA increase the risk for colorectal cancer or polyps. *Cancer Epidemiol Biomarkers Prev* **3**: 675–682

Lohman TG, Roche AF and Martorell R (eds) (1991) *Anthropometric Standardization Reference Manual*. Human Kinetics Books: Champaign, IL

Ma J, Stampfer MJ, Giovannucci E, Artigas C, Hunter DJ, Fuchs C, Willet WC, Selhub J, Hennekens CH and Rozen R (1997) Methylenetetrahydrofolate reductase polymorphism, dietary interactions, and risk of colorectal cancer. *Cancer Res* **57**: 1098–1102

MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, Abbott R, Godwin J, Dyer A and Stamler J (1990) Blood pressure, stroke and coronary heart disease L Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* **335**: 765–774

Marshall JR and Hastrup JL (1996) Mismeasurement and the resonance of strong confounders: uncorrelated errors. *Am J Epidemiol* **143**: 1069–1078

Ness AR, Khaw K-T, Bingham S and Day NE (1996) Vitamin C status and undiagnosed angina. *J Cardiovasc Risk* **3**: 373–377

Phillips A and Davey Smith G (1993) The design of prospective epidemiological studies: more subjects or better measurements? *J Clin Epidemiol* **46**: 1203–1211

Powles J, Day NE, Sanz M and Bingham S (1996) Protective foods in winter and spring: a key to low cardiovascular mortality? *Lancet* **348**: 898–899

Riboli E (1992) Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann Oncol* **3**: 783–791

Roberts-Thompson IC, Ryan P, Khoo KK, Hart WJ, McMichael AJ and Butler RN (1996) Diet, acetylator phenotype, and risk of colorectal neoplasia. *Lancet* **347**: 1372–1374

Standing SJ and Taylor RP (1992) Glycated haemoglobin: an assessment of high liquid capacity liquid chromatographic and immunoassay methods. *Ann Clin Biochem* **29**: 494–504

Surtees PG, Wainwright NWJ, Gilks WR, Brugha TS, Meltzer HH and Jenkins R (1997) Diagnostic boundaries, reasoning and depressive disorder. II. Application of a probabilistic model to the OPCS general population survey of psychiatric morbidity in Great Britain. *Psychl Med* **27**: 847–860

Taylor JA, Hirvonen A, Watson M, Pittman G, Mohler JL and Bell DA (1996) Association of prostate cancer with vitamin D receptor gene polymorphism. *Cancer Res* **56**: 4108–4110

Vineis P (1997) Molecular epidemiology: low-dose carcinogens and genetic susceptibility. *Int J Cancer* **71**: 1–3

Vuilleumier JP and Keck E (1989) Fluorometric assay of vitamin C in biological materials using a centrifugal analyser with fluorescence attachment. *J Micronutrient Analysis* **5**: 25–34

Wareham NJ, Hennings SHJ, Prentice AM and Day NE (1997) Feasibility of heart-rate monitoring to estimate total level and pattern of energy expenditure in a population-based epidemiological study: the Ely young cohort feasibility study 1994–5. *Br J Nutr* **78**: 889–900

World Cancer Research Fund (1997) *Food Nutrition and Cancer: a Global Perspective*. American Institute for Cancer Research: Washington DC