

Sampling procedure, participation rates and representativeness in the Swedish part of the European Youth Heart Study (EYHS)

Anita Hurtig Wennlöf^{1,2,*}, Agneta Yngve¹ and Michael Sjöström¹

¹Unit for Preventive Nutrition, Biosciences at NOVUM, Karolinska Institute, Huddinge, Sweden:

²Division of Biomedicine, Department of Caring Sciences, Örebro University SE-701 82 Örebro, Sweden

Submitted 4 July 2002: Accepted 23 September 2002

Abstract

Objective: The European Youth Heart Study (EYHS) is a cross-sectional, school-based population study on risk factors for future cardiovascular disease in children, with an overall participation rate in Sweden of about 50%. To study the representativeness of the participants in the Swedish part of EYHS, a comprehensive non-participant follow-up study was carried out.

Design: A structured multilevel analysis model was developed, addressing each level in the sampling procedure. The income, educational and occupational categories of the geographical regions of the study (level I), school catchment areas (level II) and parents (level III) were compared with official data. Participating and non-participating pupils (level IV) were compared through a questionnaire.

Setting: Thirty-seven state schools in two regions of Central Sweden (Örebro and southern Stockholm) were visited during the school year 1998/1999.

Subjects: Boys and girls aged 9 and 15 years were randomly sampled through a multiphase sampling procedure.

Results: Data for socio-economic status for levels I and II corresponded well to national and regional official data. At level III, non-manually working parents were slightly over-represented among parents of participating children. At level IV, non-participating subjects corresponded in most respects to participants with a few exceptions – mainly more interest in physical exercise among participants.

Conclusions: Based on the knowledge from the non-participant study, we do not foresee problems regarding interpretation of the outcomes in the EYHS, despite the low participation rate.

Keywords
Participation
Health survey
Children
Cardiovascular diseases
Risk factors

The response rates in population studies show a relationship to health status, health behaviour and socio-economic status (SES). The rates are generally higher among persons with higher SES. This has been reported in Swedish¹ and international health studies^{2–4}. Non-response is a problem in any epidemiological study in terms of the representativeness of the population being investigated. It is important to know as much as possible not only about the participants, but also about the non-participants.

The overall aim of the European Youth Heart Study (EYHS) is to investigate the nature and strength of, and the interactions between, personal, environmental and life-style influences on the risk factors for future cardiovascular disease (CVD) in children and adolescents. The test battery in the EYHS main study consisted of venous blood sampling, anthropometric measurements, assessments of fitness (maximal ergometer bicycle test), total physical activity (accelerometer registration during four consecutive days) and dietary intake (interview). Both the parents and the children completed questionnaires concerning

habits and attitudes in areas related to CVD risk factors. Data have been collected in five European countries so far (Denmark, Sweden, Estonia, Portugal and Norway). The overall participation rate in Sweden was only about 50%, which led to concern about the representativeness of the Swedish sample, and consequently about the generalisability of the results. We therefore carried out a comprehensive analysis of the non-participants, by a separate non-participant follow-up study.

The non-participant study was based on the different levels in the sampling procedure in the main study, as described below, with focus on the following questions.

- Level I (regions): were the regions chosen representative for Sweden concerning income, education and occupation?
- Level II (schools): did the collaborating schools differ from non-collaborating schools concerning the income level of their catchment areas? What were the reasons given for not participating in the study?

- Level III (parents): did the parents of participating pupils differ from the general population in the region concerning income, educational and occupational categories? Did the parents of participating pupils differ from the national population concerning familial heredity for CVD?
- Level IV (individual pupils): were there any differences between participants and non-participating peers concerning school results (written marks)? Were there any differences between these two groups in their habits and attitudes concerning physical activity and food, as assessed by a questionnaire?

Analyses of questionnaire data often result in repeated significance testing, and some significant results will then occur at random. There are statistical procedures designed to control this multiplicity effect. In the present non-participant study we have used statistical tests of questionnaire data to reflect a multiple comparison problem, for example by introducing the family error rate.

The present paper describes the sampling procedure and the participation rates in the Swedish EYHS main study, and the results of the separate non-participant follow-up study.

Material and methods

Material, EYHS main study and the non-participant study

Data were collected between the beginning of September 1998 and the end of May 1999. The sampling frame was all state schools for children aged 9–10 years (grade 3) and 15–16 years (grade 9) in  rebro and the southern region of Stockholm. The municipality of  rebro has 160 000 inhabitants, of whom there are about 1500 pupils in each of these age groups. The southern region of Stockholm, defined here as the municipalities of Huddinge, Haninge, Botkyrka, Nyn shamn, Salem, S dert lje and Tyres , has about 370 000 inhabitants, of whom there are about 5000 pupils in each of the two age groups. The non-participant study was carried out in the schools that collaborated in the EYHS.

Methods

Sampling in EYHS main study

The two geographical regions were chosen, and within those a multiphase sampling procedure was used⁵. All state schools with more than 20 children in the age groups in question were initially included. The schools were stratified with regard to school grade and the mean income level in their catchment areas (below or above the mean in their municipality). A random sampling procedure from each stratum was carried out, and the sampled schools were invited to participate in the study.

From the complete lists of pupils in the collaborating schools, groups of pupils proportional to the sizes of the respective schools were randomly selected with the aim of obtaining 250 participants in each age and gender group (totally 1000 participants). Fifty per cent over-sampling was performed in accordance with experiences from a pilot study. The sampling procedure was organised as follows.

1. Sampled pupils received a letter with information, a written informed consent form and a questionnaire to be completed by both of the parents or the legal guardians. This included questions about their income, education and occupation and the child's health. All information and questionnaires were distributed through the class teachers. The pupils were asked to return the completed form and the questionnaire to the class teacher in an enclosed envelope within a week.
2. A reminder was distributed in the same way to non-responding families, 10 days after the first letter.
3. For those who agreed to participate, a third letter with practical information was distributed one week before the day of most tests in the test battery (see Introduction).
4. Lists of participants were distributed to the teachers in order to ensure that the pupils would be allowed, and reminded, to undergo the tests during school hours.

Non-participant study

The non-participant study followed when the main study was closed at every school. One visit was scheduled to each class unit, and a questionnaire was handed out to all pupils present on that single occasion – participants, non-participants and those not sampled. The questionnaire was a shortened hard copy version of the computerised EYHS main study questionnaire, with some additional questions about reasons for participating or not. The questions included self-reported height and weight, habits and attitudes regarding physical activity, food, smoking, drugs and family life.

The analysis addressed each level in the sampling procedure as follows.

- Level I (regions): from official data, income, educational and occupational categories for the regions were compared with the same variables for Sweden.
- Level II (schools): the mean income level of each school's catchment area was compared with the mean income level in the municipality. Non-collaborating schools were also contacted by phone to find out why the school had decided not to participate in the study.
- Level III (parents): to evaluate the representativeness of the parents of participating children, we compared data on self-reported income, occupational and educational categories from the parents' questionnaire with official regional data^{6–8}. For comparison, data for the age

group 30 to 55 years was used, which best corresponded to the age of the parents (mean 43 years, range 26–72 years). Data from the parents' questionnaire concerning own or close relatives' CVD were compared with national data^{9–11}.

- Level IV (individuals): for comparisons between participating and non-participating pupils, an objective measure (written school marks) and a subjective measure (questionnaire) were used. Non-participants are defined as sampled pupils not participating in the study (i.e. sampled pupils who gave a negative answer to the question of participation, or who did not respond at all, or who gave a positive answer regarding participation but did not attend the tests).

The collaborating schools were asked to report the individual written marks from the ordinary schoolwork. No written marks are given in grade 3. In grade 9 written marks are usually given in 17 subjects, and there are four categories of credit points in each subject (0, 10, 15 and 20 credit points, respectively); thus a pupil can receive a maximum of 340 credit points. The total individual credit points and, separately, the credit points in physical education were then analysed for the participants and the rest of their peers. Restrictions in the use of the Swedish personal identification code made it impossible to create groups for all three possible alternatives (i.e. participants, non-participants and non-sampled pupils). The two latter groups had to be pooled, and the analysis had to be made for differences between participants (identified by the personal identification code) and the rest of the pupils (without a personal identification code).

Ethics

The local ethical committees approved the study protocols. The children and their families received written information about the purpose and the content of the study. Written consent was obtained from one of the parents or a legal guardian for the 9-year-olds, and from one of the parents or a legal guardian and the child for the 15-year-olds. The boards of all collaborating schools approved the use of school facilities during school hours.

Statistics

The presentations of the results on levels I–III (i.e. regions, schools and parents) are descriptive. At level IV, concerning individuals (written marks and questionnaire), Student's *t*-test and the chi-square test were used to test for differences between participants and non-participants.

The questionnaire in the non-participant study consisted of 48 questions, and the data were split into four groups (i.e. gender and grade). The large number of tests increases the risk of false significances. For this reason it is preferable, instead of using the common local significances, to use the overall significance, or family error rate

(FER). FER is defined as the probability that one or more false significances out of k tests is less than or equal to α .

Bonferroni's method is the most common approach when handling multiple comparisons, but it is very restrictive and means quite simply a change of the local significance level, α , to α divided by k , where k is the number of tests. The Bonferroni procedure guarantees that FER is less than or equal to α . Holm¹² introduced another, less restrictive method to keep FER less than α . The method of Holm proceeds as follows. Sort the P -values of the k tests in increasing order, $P_1, P_2, \dots, P_i, \dots, P_k$. If $P_1 > \alpha/k$, none of the k tests are significant, and the test procedure is finished. If $P_1 \leq \alpha/k$, test 1 is significant, and now P_2 is examined. If $P_2 > \alpha/(k-1)$, none of the $(k-1)$ remaining tests are significant, but if $P_2 \leq \alpha/(k-1)$, test 2 is significant and P_3 is examined. This procedure goes on until $P_i > \alpha/(k-i+1)$ and the procedure is interrupted.

The level of significance used was $\alpha = 5\%$. Holm's multiple test procedure for correction of mass significance effects¹² was applied to the questionnaire data. As an illustration of the method of Holm, the local P -values from the chi-square tests are reported together with modified P -values according to Holm, for one subset of data.

Results

Participation in the main study

Schools and individuals

The school boards of 42 out of 87 sampled schools gave a positive answer to participation (47% in Örebro and 51% in Stockholm).

In total, 2313 pupils were sampled. Of these, 1154 (50%) agreed to participate and gave written informed consent, 566 (24%) declined to participate and 593 (26%) did not respond at all. Of those who agreed to participate, 17 (1%) failed to keep their appointment and thus a total of 1137 (49%) of the sampled pupils participated. Split into grade and gender groups, the participation rates among 9–10-year-old boys and girls were 59% and 65%, respectively, and among 15–16-year-olds these rates were 36% and 46%, respectively (Table 1). Demographic data showed

Table 1 Participation rates (numbers of subjects) at Swedish locations in the European Youth Heart Study

	School grade 3 (9–10 years, born 1989)		School grade 9 (15–16 years, born 1983)	
	Girls	Boys	Girls	Boys
Örebro	62% (135)	61% (136)	44% (135)	35% (112)
Southern Stockholm*	68% (155)	57% (136)	48% (182)	37% (146)
Total	65% (290)	59% (272)	46% (317)	36% (258)

* Southern Stockholm – Botkyrka, Haninge, Huddinge, Nynäshamn, Salem, Södertälje and Tyresö municipalities.

that approximately 0.5% of the Swedish children in the actual age groups was assessed. Non-response items, such as registrations of low quality or missed venepuncture, will be reported separately.

Non-participant study

Level I, regions

No essential differences between regional and national population levels were found in the studied variables (Table 2). The largest differences were observed for occupation, where the Stockholm region had 4 percentage units more non-manual workers than Sweden as a whole. For all other categories of occupation and education the differences between regional and national data were smaller. The mean income level differed by 8000 SEK between the regional population of Stockholm and the national population (206 000 SEK and 198 000 SEK for Stockholm and Sweden, respectively).

Level II, schools

No special pattern in income levels of catchment areas was found in non-collaborating schools; high- and low-income areas declined collaboration to a similar extent (data not shown). The non-collaborating schools frequently referred to a shortage of facilities for the implementation of the project as the main reason for not participating in the EYHS; three to four small rooms were required during the days of tests.

Level III, parents

Data on parental income, occupation and education, in comparison with official regional and national data, are also given in Table 2. Results from the parents of participating pupils are presented for the two study

locations ( rebro and Stockholm) separately, for comparison with regional data, and the weighted averages for the two study locations (Sweden) are given for comparison with the national population. Parents of EYHS participants had a median income of 201 000 SEK, compared with a national mean income level of 198 000 SEK (Table 2). At a regional level, the parents of participants in  rebro had a lower income than the regional population (median value 191 000 SEK and mean value 198 000 SEK, respectively), while the income (median) of the parents in Stockholm equalled the regional population mean income (206 000 SEK). The distribution of the occupational and educational levels showed similar patterns in the two regions. Manual workers were under-represented by 13 and 7 percentage units in  rebro and Stockholm, respectively, as compared with official data for the regions. Non-manual workers were over-represented by 9 and 4 percentage units in  rebro and Stockholm, respectively. A low educational level was under-represented by 12 and 10 percentage units in  rebro and Stockholm, respectively, and a university level was over-represented by at most 6 percentage units compared with official regional data.

Forty-five per cent of the parents reported to have at least one close relative with diagnosed CVD, and 2% of the parents reported own CVD.

Level IV, individuals

Written marks. Information on all written marks was obtained for 1930 pupils, covering about 90% of the pupils in grade 9 in the collaborating schools. There was a significant difference between the participants and the rest of the pupils in average credit points, all 17 school subjects included (217 ± 58 vs. 183 ± 71 , respectively; $P < 0.001$).

Table 2 Income level and occupational and educational categories in Swedish families participating in the European Youth Heart Study (EYHS), compared with regional and national populations

	�rebro		Stockholm		Sweden	
	Parents of EYHS participants	Regional population	Parents of EYHS participants	Regional population	Parents of EYHS participants	National population
Income (SEK)	191 000*	198 000†	206 000*	206 000†	201 000*	198 000†
Missing (%)	6	–	5	–	6	–
Occupation (%)						
Manual worker	30	43	35	42	34	44
Non-manual worker	53	44	49	45	50	41
Other‡	13	13	12	13	12	15
Missing	4	–	4	–	4	–
Educational level (%)						
≤ 9 years	15	27	20	30	18	30
≤ 12 years	42	44	44	46	43	44
University level	32	26	26	21	28	24
Other education	8	2	6	3	7	2
Missing	3	–	4	–	4	–

* Median value.

† Arithmetic mean value.

‡ Own business or not gainfully employed (e.g. student, unemployed, retired).

Participants had significantly higher marks in physical education than their non-participating peers (13.8 ± 4.4 vs. 11.5 ± 5.5 , respectively; $P < 0.001$). The distribution of the credit point categories is shown in Fig. 1.

Questionnaire. Questionnaire data were collected for a total of 2360 pupils. Thirty-two questionnaires were excluded because of a very bad quality, and 559 came from pupils not sampled in EYHS and were not further analysed. The remaining 1769 questionnaires were from 810 non-participants and 959 participants. About 80% of the pupils were present in the classroom at the time of distribution of the questionnaire.

The most common reason given for non-participation was 'I didn't want to have a blood sample taken'; 38% of the 9-year-olds and 33% of the 15-year-olds gave this answer. The second most common answer (22%) given by the 9-year-old children was 'My parents didn't want me to participate', with similar responses from boys and girls. This answer was given as an explanation by only 4% of the non-participating 15-year-olds. In the older group, the second and third most common answers given by the girls were 'My parents didn't want me to participate' (5%) and 'I didn't want to perform a bike test' (5%), and by the boys, 'I didn't want to leave home without breakfast in the morning' (12%) and 'Because my friends were not participating' (3%).

The analysis of the self-reported height and weight showed no significant differences between participants and non-participants between genders or grade groups. However, when body mass index (BMI) was calculated from the self-reported values, significant differences in

BMI was found among the 9-year-old girls between participants and non-participants, with values of 16.7 ± 2.2 and $17.1 \pm 2.3 \text{ kg m}^{-2}$, respectively ($P < 0.05$).

Among the girls in grade 3, no significant difference between participants and non-participants was found regarding answers in the questionnaire. In the girls in grade 9, the answers to one question differed significantly, namely the question/statement: *I really like doing physical education at school* (Alternatives: yes/no). The participants showed a significantly more positive attitude to physical education.

In boys, significant differences between participants and non-participants were observed for three questions in grade 3, and for two questions in grade 9 (one of the questions was the same in both age groups); all concerning physical exercise and eating habits.

Among the boys of both school grades, the participants were more positive to the statement *If I were to exercise most days, it would get or keep me in shape* (Alternatives: yes/perhaps/no). The participating boys in grade 3 took more part in clubs, etc., as seen by the answers to the question: *How often do you take part in exercise at clubs, such as sport clubs, youth clubs, etc?* (Alternatives: hardly ever or never/once or twice a week/nearly every day/every day).

There were also differences in eating habits among the boys. According to the questionnaire, the participants in grade 3 ate fewer hamburgers, and the participants in grade 9 ate fewer pizzas, compared with the non-participants. The questions were: *How often do you eat burgers?* (Alternatives: every day/nearly every day/once or

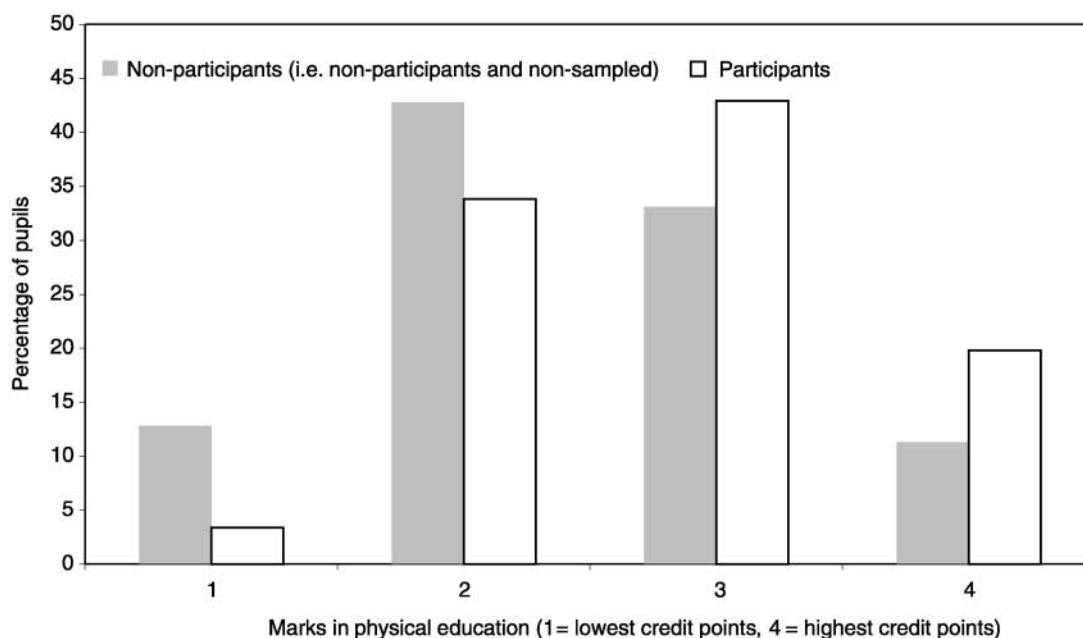


Fig. 1 Distribution of marks in physical education at school among participants in the Swedish part of the European Youth Heart Study, compared with all other pupils in the same schools. Participants had significantly higher marks than the rest of the pupils ($P < 0.001$)

twice a week/hardly ever or never) and *How often do you eat pizza?* (Alternatives: *every day/nearly every day/once or twice a week/hardly ever or never*).

The correction for FER using the method of Holm decreased the number of significant differences in the chi-square test for all four age and gender groups, from 34 to six questions. An example for third grade boys is given in Table 3.

Discussion

The individual decision to volunteer and participate in a study depends on a number of factors and is not a random process. In studies including children, the situation is complicated by the fact that other persons (e.g. parents, teachers and friends) influence the final decision to take part or not. For this reason, the present study of the representativeness of the Swedish sample in the EYHS addressed adult characteristics (i.e. characteristics of the parents) as well as the habits and attitudes of the children.

The primary finding was that participants and non-participants reported similar attitudes and behaviours in most areas covered by the questionnaire, except for interest in physical exercise and to some extent for food habits among the boys. This selection bias in voluntary amount of physical exercise may decrease the range of exposure to physical activity, but will not necessarily influence the outcome (i.e. the risk for future CVD), as the risk factors are supposed to be multi-factorial. Nevertheless it is of importance, since physical activity is one of the variables of interest. We will make use of this information in our further interpretations of the data, taking into consideration that we may have lost part of the group least interested in physical exercise, which is one component of total physical activity.

Moreover, we found the socio-economic background of the participants to be slightly above average, compared with that of the general Swedish population, and the data will be interpreted with this taken into account.

The reason for studying the SES of the participants as one aspect of representativeness is based on the known risk in epidemiological studies of under-representation of the most disadvantaged groups. Groups with lower SES are less likely to participate in risk factor screening programmes or to receive preventive services^{13–15}. Socio-economic factors have been associated with non-participation in previous Swedish health surveys, but the findings are conflicting and vary from underestimation of differences between participants and non-participants to overreporting of such differences^{2,16,17}. The impact of socio-economic conditions in childhood on future CVD is not fully understood. Health behaviour is found to be related to SES in children also, although the adult socio-economic destination seems to be the best predictor of mortality, not the childhood origin *per se*^{18–20}. Despite a

major political goal to decrease social inequalities in Sweden, it has recently been shown that the differences are still maintained or even increasing in some groups^{10,21}. Furthermore, the most deprived neighbourhood environments have been shown to have their own independent increased risks for a poor cardiovascular health profile^{22,23}. These findings correspond to those in international studies^{24,25}.

On the school level, or the level of school catchment areas, only income data were available, and these were used only as a discriminator to dichotomise schools into those with high and low income. The reason for analysing the school level was a concern that all kinds of schools should be represented, as further characteristics may follow with income, such as immigrant rates and unemployment rates, for example – data we could not analyse for every individual school.

Despite the under-representation of lower SES families in this study, we found a parental income level (median value) in agreement with the national mean value. This might be due to part-time working in well-paid families, but we have no information on part-time work. The population data are based on all inhabitants, with or without children. The restriction to a certain age group, 30–55 years, limits the error in comparing these different groups, but we do not know the magnitude or direction of that possible error.

There is no selection bias concerning the inherited risk for CVD according to the parental questionnaire data. This conclusion is based on the concordance between the parental results and national data on CVD prevalence¹⁰. An over-representation of higher SES families may result in an underestimation of risk factors, if we assume that CVD risk factors correlate inversely to SES. Griffin *et al.*²⁶ point out another possible cause of non-participation that may also result in an underestimation of the risk, namely the fear of finding one's child out of the expected normal ranges. On the other hand, the response rate is likely to be high when the subjects have a special motivation to participate, e.g. experience of a high CVD risk, and this would give an overestimation of risk factors. The decision to participate for the youngest children is probably highly influenced by the parents' attitude towards the study, which is reflected by the fact that, in the questionnaire in the non-participant study, 22% of the children stated that their parents didn't want them to participate. Families with close relatives with CVD can be defined as a high-risk subgroup. A recent Swedish, population-based case–control study of myocardial infarction, the Stockholm Heart Epidemiology Program (SHEEP)⁹, reported 58% of the referents to have at least one close relative with CVD. However, as the SHEEP referents were older than the parents in the EYHS (mean age 59 and 43 years, respectively), the difference between 58% in SHEEP and 45% in EYHS is expected. As the proportion of participating children's families with close relatives with

Table 3 Test procedure and comparisons between local significances (i.e. for each question) and the overall significances, FER adjusted as described by Holm¹². The example shows test results for boys of grade 3, and includes the 13 locally significant results from the 48 questions

	Local <i>P</i> -values, sorted in increasing order			Modified α -level			FER significance
	Local α -level	Local significance	Modified α -level	FER significance			
How often do you take part in exercise at clubs, such as sport clubs, youth clubs, etc?	0.000	yes	0.00104	yes			
If I were to exercise most days, it would get or keep me in shape.	0.000	yes	0.00106	yes			
How often do you eat burgers?	0.000	yes	0.00109	yes			
How often do you eat chips?	0.002	yes	0.00111	no			
I feel that I can easily keep up with other children of my age when we are playing games and sports.	0.005	yes	0.00114	no			
How often do you drink beer?	0.009	yes	0.00116	no			
How often do you have fizzy drinks?	0.019	yes	0.00119	no			
How often do your friends exercise or play sports?	0.020	yes	0.00122	no			
If I were to exercise most days, it would make me healthy.	0.022	yes	0.00125	no			
If I were to exercise most days, it would help me make new friends.	0.031	yes	0.00128	no			
What do you normally do at breaks (at school)?	0.038	yes	0.00132	no			
How often do you participate in playing games or sports outside school?	0.045	yes	0.00135	no			
How often do you eat pizza?	0.064	no	0.00139	no			
Number of significant tests		12		3			

FER – family error rate.

this diagnosis correspond very well to official data on CVD, we do not have to consider the possibility of an overestimation of risks due to over-representation of children from high-risk families, or of a recall bias phenomenon.

The reasons for non-participation on an individual level are not easily detected. The reported concerns about the blood sampling might partly be the easiest way to conceal lack of interest, but we did not discover any underlying explanation with the method used. The statistical analyses of the reported written marks are consistent with our impression during the study period that the most successful pupils participated to a higher degree in the EYHS. The explanation for this cannot be related exclusively to SES, as we found only minor discrepancies in the SES distribution compared with the general population. More likely, it mirrors a generally positive attitude towards all kinds of school-based activities among the participants.

Previous studies in this area have revealed a similar pattern of participation as in our study. It has been noted before that adolescent girls are more interested than boys in participating, that there is a skewed SES distribution between participants and non-participants and that adolescents are less interested in participating than younger children. Non-participant studies designed to find a few subject characteristics (e.g. interest in physical activities) could be expected to indicate that participants are generally younger, more educated and more likely to report healthier patterns^{2,3,27}.

In conclusion, we found no main objections against the choice of regions, or against the sampling results on a school level, and therefore conclude that the sampling frame was representative of the source population.

As the random sampling procedure is considered to have achieved a representative sample of families, the slight selection bias found regarding SES is most probably due to the previously described socio-economic effect on propensity to respond^{1–4}.

As the participation was quite demanding on the subjects (e.g. blood sampling and maximal ergometer bicycle test), the self-selection bias observed on an individual level was not unexpected. Based on the very few significant differences found between participants and non-participants in the questionnaire result in the non-participant study, we do not foresee problems regarding interpretation of the results in the EYHS main study, despite the low participation rate.

References

- 1 Lundström S, Särndal C-E. *Estimation in the Presence of Non-response Error and Frame Imperfections*. Örebro: Statistics Sweden (SCB-Tryck), 2001.
- 2 Berg C, Jonsson I, Conner MT, Lissner L. Sources of bias in a dietary survey of children. *Eur. J. Clin. Nutr.* 1998; **52**(9): 663–7.

- 3 Klesges R, Williamson J, Somes G, Talcott G, Lando H, Haddock C. A population comparison of participants and nonparticipants in a health survey. *Am. J. Public Health* 1999; **89**(8): 1228–31.
- 4 Twisk JW, Kemper HC, van Mechelen W. Tracking of activity and fitness and the relationship with cardiovascular disease risk factors. *Med. Sci. Sports Exerc.* 2000; **32**(8): 1455–61.
- 5 Särndal C-E, Swensson B, Wretman J. *Model Assisted Survey Sampling*. New York: Springer-Verlag, 1992.
- 6 Statistics Sweden. *Sveriges officiella statistik, befolkningsstatistik del I-II*. Örebro: SCB, 2000 [in Swedish].
- 7 Statistics Sweden. *Socio-ekonomisk indelning (SEI)*. Reports on Statistical Co-ordination 1982:4. Örebro: SCB, 1995 [in Swedish].
- 8 Statistics Sweden. *Bakgrundsutredning om befolkningens utbildning 1999:3. Utbildningsnivå i kommunerna, länen och riket 1999-01-01*. Örebro: SCB, 2001 [in Swedish].
- 9 Leander K, Hallqvist J, Reuterwall C, Ahlbom A, de Faire U. Family history of coronary heart disease, a strong risk factor for myocardial infarction interacting with other cardiovascular risk factors: results from the Stockholm Heart Epidemiology Program (SHEEP). *Epidemiology* 2001; **12**(2): 215–21.
- 10 Socialstyrelsen. *Folkhälsorapport 2001*. Stockholm: Socialstyrelsen (National Board of Health and Welfare), 2001 [in Swedish].
- 11 Socialstyrelsen. *Causes of Death Register: SoS/EpiC*. Stockholm: Socialstyrelsen (National Board of Health and Welfare), 2001 [in Swedish].
- 12 Holm S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 1979; **6**(2): 65–70.
- 13 Pirie PL, Elias WS, Wackman DB, Jacobs DR, Murray DM, Mittelmark MB, *et al.* Characteristics of participants and nonparticipants in a community cardiovascular disease risk factor screening: the Minnesota Heart Health Program. *Am. J. Prev. Med.* 1986; **2**(1): 20–5.
- 14 Lahelma E, Valkonen T. Health and social inequities in Finland and elsewhere. *Soc. Sci. Med.* 1990; **31**(3): 257–65.
- 15 Bartley M, Power C, Blane D, Smith GD, Shipley M. Birth weight and later socioeconomic disadvantage: evidence from the 1958 British cohort study. *Br. Med. J.* 1994; **309**(6967): 1475–8.
- 16 Bostrom G, Hallqvist J, Haglund BJ, Romelsjö A, Svanström L, Diderichsen F. Socioeconomic differences in smoking in an urban Swedish population. The bias introduced by non-participation in a mailed questionnaire. *Scand. J. Soc. Med.* 1993; **21**(2): 77–82.
- 17 Samuelson G, Bratteby LE, Enghardt H, Hedgren M. Food habits and energy and nutrient intake in Swedish adolescents approaching the year 2000. *Acta Paediatr. Suppl.* 1996; **415**: 1–19.
- 18 Diderichsen F. Health and social inequities in Sweden. *Soc. Sci. Med.* 1990; **31**(3): 359–67.
- 19 Lynch JW, Kaplan GA, Cohen RD, Kauhanen J, Wilson TW, Smith NL, *et al.* Childhood and adult socioeconomic status as predictors of mortality in Finland. *Lancet* 1994; **343**(8896): 524–7.
- 20 Lynch J, Kaplan G, Salonen J. Why do poor people behave poorly? Variation in adult health behaviours and psychosocial characteristics by stages of the socioeconomic lifecourse. *Soc. Sci. Med.* 1997; **44**: 809–19.
- 21 Halldorsson M, Kunst AE, Köhler L, Mackenbach JP. Socioeconomic inequalities in the health and adolescents. A comparative study of the five Nordic countries. *Eur. J. Public Health* 2000; **10**(4): 281–8.
- 22 Sundquist J, Malmström M, Johansson SE. Cardiovascular risk factors and the neighbourhood environment: a multi-level analysis. *Int. J. Epidemiol.* 1999; **28**(5): 841–5.
- 23 Wamala SP, Lynch J, Kaplan GA. Women's exposure to early and later life socioeconomic disadvantage and coronary heart disease risk: the Stockholm Female Coronary Risk Study. *Int. J. Epidemiol.* 2001; **30**(2): 275–84.
- 24 Winkleby MA, Robinson TN, Sundquist J, Kraemer HC. Ethnic variation in cardiovascular disease risk factors among children and young adults: findings from the Third National Health and Nutrition Examination Survey, 1988–1994. *J. Am. Med. Assoc.* 1999; **281**(11): 1006–13.
- 25 Bartley M, Fitzpatrick R, Firth D, Marmot M. Social distribution of cardiovascular disease risk factors: change among men in England 1984–1993. *J. Epidemiol. Community Health* 2000; **54**(11): 806–14.
- 26 Griffin TC, Christoffel KK, Binns HJ, McGuire PA. Family history evaluation as a predictive screen for childhood hypercholesterolemia. Pediatric Practice Research Group. *Pediatrics* 1989; **84**(2): 365–73.
- 27 Nathanson CA. Sex roles as variables in preventive health behavior. *J. Community Health* 1977; **3**(2): 142–55.