

ARTICLE

A centenarian-only approach for assessing gene–gene interaction in human longevity

Qihua Tan¹, Giovanna De Benedictis², Svetlana V Ukraintseva¹, Claudio Franceschi³, James W Vaupel¹ and Anatoli I Yashin^{*,1,4}

¹Max-Planck Institute for Demographic Research, Rostock, Germany; ²Cell Biology Department, University of Calabria, Rende, Italy; ³Department of Experimental Pathology, University of Bologna, Italy; ⁴Center for Demographic Studies and Sanford Institute, Duke University, Durham, NC, USA

In this study, we introduce a centenarian-only approach to the assessment of gene-gene interaction that contributes to human longevity. This approach corresponds to the non-traditional case-only method in the genetic study of gene and disease associations. We first describe how the method can be implemented to screen for gene–gene interaction in human longevity. Then we apply the method to centenarian data collected from an Italian centenarian study in order to detect the interactions between the REN gene and the mitochondrial haplotypes. A significant interaction between REN gene allele 10 and the mitochondrial H haplotype, which may favour longevity, was found. Important features of the application in human longevity studies are highlighted and discussed. Since centenarians constitute a special population representing successful ageing, the centenarian-only approach will be an important tool in the search for major genes that contribute to human longevity.

European Journal of Human Genetics (2002) 10, 119–124. DOI: 10.1038/sj/ejhg/5200770

Keywords: centenarian-only approach; gene-gene interaction; longevity

Introduction

Longevity is a complex trait which is affected by both genetic and environmental factors, with potential interactions among them. The earliest confirmation of the influence of genes on longevity came from studies on related individuals such as twins. According to a well-known Danish twin study, the genetic variance component to life span was estimated to account for about 25% of total life span variation.^{1–3} These results indicate how important a role genes play in modulating individual survival. In addition, these studies also found evidence for non-additive genetic effects on life span resulting from interactions among genes either at a single locus or at different loci.^{1,2} Such results support the works of Wright⁴ and might suggest that non-additivity is an important mechanism

in maintaining individual survival. However, the molecular evidence of non-additivity in human longevity is far from clear.

There has been an ever-increasing interest in studying the association between genes and longevity due to rapid developments in the field of molecular biology.⁵ Important genes such as the ApoE have been reported to exhibit a significant influence on human longevity.^{6–10} Unfortunately, there has been very little research focusing on the coordinating effects among different genes, which constitutes the basic nature of non-additivity. One of the difficulties in this approach can be ascribed to the involvement of multiple genes or loci, especially when the genes of interest are highly polymorphic, which requires a vast amount of work in genotyping. While the genotyping problem is being solved by advanced new technology, an even more difficult task facing researchers is the search for genes that manifest coherent rather than independent effects. This issue is gaining in importance since there is soon to be an upsurge in the amount of individual genetic information available. When there are many potential gene interactions, the number of individuals for each possible gene combination

*Correspondence: AI Yashin, Max Planck Institute for Demographic Research, Doberaner Str. 114, 18057 Rostock, Germany.

Tel: +49 381 2081 106; Fax: +49 381 2081 406;

E-mail: Yashin@demogr.mpg.de

Received 23 July 2001; revised 10 December 2001; accepted 13 December 2001

will be very small. Hence the power for statistical testing will be low. In order to deal with this situation, powerful statistical methods are called for. In this paper we introduce a new approach to detecting gene–gene interactions in human longevity, the case-only design. This approach was originally derived for analysing gene-environment and gene-gene interactions in the etiology of complex diseases.^{11–13} Since we treat centenarians as cases, we refer to our method as the centenarian-only approach. After introducing the method we show, on the basis of examples, how this simple approach can be implemented to screen for gene–gene interactions in human longevity. We also discuss the basic assumptions of the model in the context of longevity studies. We highlight the advantages of our approach over other models and discuss problems that arise in the application.

Methods

In Table 1, we display the conventional case-control approach, which is aimed at assessing the influence on longevity of two gene alleles observed from two chromosomes, with an additional interest in detecting the effect of interaction between them. In Table 1, we use the subscripts *co* for controls and *ce* for centenarians. The subscripts *a* and *b* stand for the absolute numbers of observations in the centenarian and the control groups, *p*₀₀ is the frequency of subjects carrying none of the alleles, *p*₁₀ is the frequency of subjects carrying allele M but not N, *p*₀₁ is the frequency of subjects carrying allele N but not M, and *p*₁₁ is the frequency of subjects carrying both alleles. In accordance with the above notation, *OR*₁₀ is the odds ratio for the effect of allele M alone, *OR*₀₁ is the odds ratio for the effect of allele N alone, *OR*₁₁ is the odds ratio for the joint effect of the two gene alleles. The cross-product for the centenarian group (Table 2) can be calculated as

$$OR_{ce} = \frac{a_{11}a_{00}}{a_{10}a_{01}} = \frac{p_{11,ce}p_{00,ce}}{p_{10,ce}p_{01,ce}} \tag{1}$$

where *p*_{11,ce}+*p*_{10,ce}+*p*_{01,ce}+*p*_{00,ce}=1. Using the relationships between the frequencies and odds ratios in Table 1 and substituting them into (1) yields

$$OR_{ce} = \frac{OR_{11}}{OR_{10}OR_{01}} \tag{2}$$

In equation (2), *OR*_{co} is the control-only odds ratio, which can become one under two assumptions. First, the two genes are independent; second, the event (being a centenarian) related to the interaction is rare. Here independence means the two genes are not in linkage disequilibrium. Before discussing the relevant issues concerning the assumptions, we first demonstrate how the control-only odds ratio becomes one when the two assumptions are satisfied. Similarly to equation (1), the control-only odds ratio can be written as

$$OR_{co} = \frac{b_{11}b_{00}}{b_{10}b_{01}} = \frac{p_{11,co}P_{00,co}}{p_{10,co}p_{01,co}} \tag{3}$$

Again, we have in equation (3) *p*_{11,co}+*p*_{10,co}+*p*_{01,co}+*p*_{00,co}=1. If the frequencies corresponding to the four combinations in Table 1 are *p*(*M* = 1 | *N* = 1), *p*(*M* = 1 | *N* = 0), *p*(*M* = 0 | *N* = 1) and *p*(*M* = 0 | *N* = 0), the independence of the two gene alleles means that

$$\frac{p(M = 1 | N = 1)}{p(M = 0 | N = 1)} = \frac{p(M = 1 | N = 0)}{p(M = 0 | N = 0)} \tag{4}$$

In equation (4), *p*(*M* = 1 | *N* = 1) = *p*_{11,co}*p*(*co* | *N* = 1) + *p*_{11,ce}*p*(*ce* | *N* = 1). Since longevity is rare, we have *p*(*ce* | *N* = 1) ≈ 0 and *p*(*co* | *N* = 1) ≈ 1. This leads to *p*(*M* = 1 | *N* = 1) ≈ *p*_{11,co}. In the same manner, we have *p*(*M* = 1 | *N* = 0) ≈ *p*_{10,co}, *p*(*M* = 0 | *N* = 1) ≈ *p*_{01,co}, *p*(*M* = 0 | *N* = 0) ≈ *p*_{00,co}. Substituting them into equation (4) and rearranging, we get

$$\frac{p_{11,co}p_{00,co}}{p_{10,co}p_{01,co}} \approx 1.$$

Now we can rewrite equation 1 as

$$OR_{ce} = \frac{OR_{11}}{OR_{10}OR_{01}} \tag{5}$$

Equation (5) means that the centenarian-only odds ratio measures the departure from multiplicative joint effect of the two genes. In other words, it reflects the effect of gene–gene interaction. The null hypothesis for this approach is *H*₀:*OR*_{ce}=1. Any statistically significant deviation of *OR*_{ce} from one indicates that there is a gene–gene interaction that contributes to and modifies the probability of achieving longevity. By comparing variances of the MLE of the

Table 1 A 2 × 4 table for assessing gene-gene interaction based on case-control design (a, b: numbers of subjects; p: genotype frequency; OR: odds ratio)

Gene allele ^a	Gene allele	Centenarians	Controls	OR
<i>M</i>	<i>N</i>			
0	0	<i>a</i> ₀₀ (<i>p</i> _{00, ce})	<i>b</i> ₀₀ (<i>p</i> _{00, co})	<i>OR</i> ₀₀ =1
0	1	<i>a</i> ₀₁ (<i>p</i> _{01, ce})	<i>b</i> ₀₁ (<i>p</i> _{01, co})	<i>OR</i> ₀₁ = <i>a</i> ₀₁ <i>b</i> ₀₀ / <i>a</i> ₀₀ <i>b</i> ₀₁
1	0	<i>a</i> ₁₀ (<i>p</i> _{10, ce})	<i>b</i> ₁₀ (<i>p</i> _{10, co})	<i>OR</i> ₁₀ = <i>a</i> ₁₀ <i>b</i> ₀₀ / <i>a</i> ₀₀ <i>b</i> ₁₀
1	1	<i>a</i> ₁₁ (<i>p</i> _{11, ce})	<i>b</i> ₁₁ (<i>p</i> _{11, co})	<i>OR</i> ₁₁ = <i>a</i> ₁₁ <i>b</i> ₀₀ / <i>a</i> ₀₀ <i>b</i> ₁₁

^a1, present; 0, absent.

Table 2 A 2 × 2 table for assessing gene-gene interaction based on centenarian-only design (a: number of subjects; p: genotype frequency; n: sum of subjects)

Gene allele N	Gene allele ^a M		Sum
	0	1	
0	a_{00} ($p_{00, ce}$)	a_{01} ($p_{01, ce}$)	n_1
1	a_{10} ($p_{10, ce}$)	a_{11} ($p_{11, ce}$)	n_2
Sum	$n_{.1}$	$n_{.2}$	n

^a1, present; 0, absent.

logarithm of equation (1) and equation (5), Piegorsch *et al.*¹¹ concluded that the case-only approach increases precision in estimating interactions since the variance corresponding to equation (1) involves an extra component for OR_{co} , the control-only odds ratio.

A statistical test for the null hypothesis can be conducted by employing the log likelihood ratio (LLR) test. One calculates twice the difference between the log likelihoods estimated at $OR_{ce}=1$ and at OR_{ce} . When the sample size is large, the LLR is approximately distributed as χ^2 on 1 degree of freedom. Alternatively, we can apply the standard χ^2 statistic to test the null hypothesis, ie $\chi^2 = \frac{n(a_{11}a_{00} - a_{10}a_{01})^2}{n_{.1}n_{.2}n_{.1}n_{.2}}$ on one degree of freedom. Here, $n_{.1}$, $n_{.2}$, $n_{.1}$, $n_{.2}$ are marginal sums of the observations by genotype, and n is the total sum of the observations (centenarians). The χ^2 statistic with continuity correction can also be applied, but this is not recommended.¹⁴ For a significant OR_{ce} , confidence intervals can be constructed. The procedure is first to construct a confidence interval for the natural log of OR_{ce} and then exponentiate the boundaries to get the confidence intervals for OR_{ce} . The standard error of $\ln(OR_{ce})$ is given by $\sqrt{\frac{1}{a_{00}} + \frac{1}{a_{01}} + \frac{1}{a_{10}} + \frac{1}{a_{11}}}$. Since the natural log of the odds ratio is more normally distributed than the odds ratio itself, we can use the critical values of the standard normal distribution for calculating the intervals.

Application

In a multicentric longevity study conducted in 1995 in Italy,¹⁵ a sample consisting of 157 centenarians (38 males, 119 females) was collected. Participants were healthy and free from major pathologies or disabilities.¹⁶ Each individual was checked for ethnicity and geographical origin. Blood samples were genotyped for genes that one suspects to have an influence on longevity. Important genes significantly associated with longevity have been reported in previous publications.^{15,17–19} Among the genes tested, REN (renin) gene failed to show a significant contribution to individual survival. The REN gene encodes for renin, an enzyme produced in the kidney. Renin is a key element in the renin-angiotensin system (RAS) in the circulation. It is responsible for regulating blood pressure and stimulating aldosterone production. In recent years, researchers have reported the existence of renin in mitochondria.^{20,21} A

possible role of intramitochondrial renin in the regulation of steroid biosynthesis is suggested.^{20,22} Here, we want to apply the centenarian-only approach described above to check if there is any functional interaction between the REN and mtDNA variations with consideration that some polymorphisms of the two genes could have enhanced functions as a result of gene-gene interaction. Five polymorphic alleles of the REN gene (alleles 7, 8, 10, 11, 12) were detected at the locus designated by the number of short tandem repeats.¹⁵ An earlier study showed that the genotype-allele frequencies are in Hardy-Weinberg equilibrium.¹⁵ However, the rare alleles 7 and 12 were not found in the centenarians, leaving the rest with allele frequency 0.793 for REN 8 allele, 0.092 for REN 10 allele and 0.105 for allele 11. In Tables 3–5, we show the effects of gene-gene interaction for REN allele 8, 10 and 11 with the mitochondrial haplotypes, H, J, K, T, U. Given the small sample size, we combine the remaining mitochondrial haplotypes to form one group,²³ which we refer to as ‘the others’. The *P* values are calculated from Fisher’s exact test for two tails. Since we are making conclusions on interaction between each REN allele and the mitochondrial haplotypes, multiple comparison has to be taken into account. The significant level has to be adjusted using Bonferroni correction. In Tables 3–5, there are five independent mitochondrial variations in which we can adjust our significance level to $1-(1-0.05)^{1/5}=0.010$ if the original type I error is set to 0.05. While no significant result is found in Tables 3 and 5, in Table 4 the REN 10 allele shows a remarkable interaction with the mtDNAhapl-H with a *P*-value of 0.006. The estimated odds ratio is 3.274 and 95% CI is from 1.405 to 7.630. The result indicates that the presence of both genes may substantially increase the probability of achieving longevity. No additional gene–gene interaction was detected as significant in Table 4.

Simulation

By investigating the distribution of the odds ratio, we can determine the threshold for a significant odds ratio for the given sample size. The simulation was done by applying the survival model introduced by Yashin *et al.*¹⁸ and Tan *et al.*¹⁹ with proportional hazard for the interaction. Since allele frequency for REN10 is 0.092, the frequency of REN10 carriers (both homozygotes and heterozygotes) is roughly 0.2. Our preliminary analysis showed that frequency for mtDNAhapl-H is about 0.4. Using these frequency parameters and assuming no interaction between the two genes, we simulated 10 000 data sets each with sample size of 157 individuals who survived above age 100. From the simulated data, we calculated 10 000 odds ratios by applying the centenarian-only approach with a median of 0.999. In Figure 1, we show the simulated distribution of the odds ratio which peaks at 1 with a long right tail. From this distribution curve, we obtain a very low probability of observing an odds ratio bigger than that for REN10 and mtDNAhapl-H in Table 4 ($P=0.0025$). It means that, if there is no interaction between

Table 3 REN8 and mtDNAhaplotype interactions from Italian Centenarian data

Genes	REN8 Present	REN8 Absent	OR	<i>P</i> ^a	95% CI
H	63	3	0.966	1.000	0.209–4.466
J	19	1	0.870	1.000	0.099–7.630
K	9	0	0.447	1.000	0.049–4.037
T	13	1	0.569	0.487	0.064–5.098
U	17	0	0.895	1.000	0.104–7.722
Others	28	2			
Sum total	150	7			

Total sample size=157

^a*P* is calculated from Fisher’s two-tail exact test.

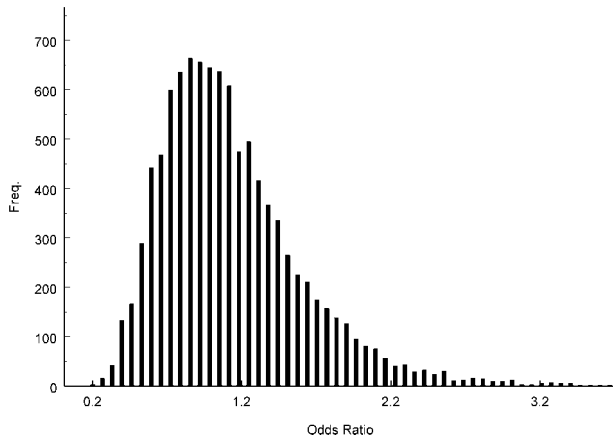


Figure 1 Simulated distribution of centenarian-only odds ratio when assuming no gene–gene interaction. 10 000 samples were simulated each with sample size of 157 individuals. The median is 0.999, probability of observing an odds ratio larger than 3.274 is 0.0025.

the two genes, the probability of having an odds ratio bigger than 3.274 is extremely low.

Discussion

The application of the centenarian-only approach has detected significant interaction between REN gene allele 10 and the mitochondrial haplotype H, which could favour longevity. Both genes had been reported to exhibit no effect on survival alone.^{17,19} However, carriers of the two genes have a significantly increased probability of achieving a long life. As specified in the method section, the model makes sense only when our primary interest is in assessing gene–gene interaction. No effect of the gene alone can be estimated by this approach. Although it is limited to interaction only, this non-traditional method has certain advantages that are unmatched by conventional case-control studies. Since the centenarian-only approach does not require controls, crucial issues in the choice of an appropriate control group, which

have complicated case-control studies, are avoided.¹² This is important since improper choice of the controls could lead to spurious conclusions that distort the study. In addition, this approach attains greater precision in estimating interactions than the traditional case-control design.^{11,24}

However, there are important assumptions that underlie the application of the model. First, before applying this method, one has to make sure that the genes in question are not in linkage disequilibrium. Second, the event associated with gene–gene interaction should be rare. The study of longevity fulfils this assumption since longevity is by definition always a rare event. It was estimated that there were only about 44 centenarians per million population in the developed countries in the year 1990.²⁵ Although the number is increasing very rapidly,²⁶ according to the United Nations’ prediction, in the year 2050 centenarians will make up only about 1% of the total population of Japan, which has the highest percentage of centenarians in the world.

Since the model only captures gene-gene interactions, additive effect of the genes can’t be measured using this approach. This means that the traditional case-control study can’t be replaced by this method when one needs to assess both the additive and non-additive effects on longevity for the genes concerned. However, the simple approach can be used as a powerful tool for screening purpose in finding important genes that work together interactively in modulating individual survival.

We must point out that, as in other association studies, the case-only approach also runs into difficulties in a situation where linkage disequilibrium exists. The detected interaction could be due to the fact that one or both of the markers are in linkage disequilibrium with the causal genes. In addition, the choice of subject in the centenarian-only study should follow the usual rules of case selection for any case-control study.²⁷ Despite these potential problems, the association approach can help to complement further work aimed at localising the specific gene loci. Since centenarians constitute the special part of the population representing successful ageing, the application of the centenarian-only approach in genetic studies on aging could help to screen important genes that contribute to human health and longevity.

Table 4 REN10 and mtDNAhaplotype interactions from Italian Centenarian data

Genes	REN10 Present	REN10 Absent	OR	<i>P</i> ^a	95% CI
H	19	47	3.274	0.006	1.405–7.630
J	3	17	0.753	1.000	0.205–2.763
K	2	7	1.280	0.672	0.252–6.509
T	1	13	0.316	0.469	0.040–2.518
U	1	16	0.250	0.201	0.032–1.966
Others	3	28			
Sum total	29	128			Total sample size=157

^a*P* is calculated from Fisher's two-tail exact test.**Table 5** REN11 and mtDNAhaplotype interactions from Italian Centenarian data

Genes	REN11 Present	REN11 Absent	OR	<i>P</i> ^a	95% CI
H	9	57	0.561	0.215	0.237–1.325
J	2	18	0.453	0.373	0.099–2.070
K	2	7	1.280	0.672	0.252–6.509
T	3	11	1.227	0.724	0.320–4.713
U	5	12	2.014	0.316	0.649–6.247
Others	8	23			
Sum total	29	128			Total sample size=157

^a*P* is calculated from Fisher's two-tail exact test.

Acknowledgments

Qihua Tan appreciates his communication with Dr Susan Clausmeyer and Professor Joer Peters at the Department of Pharmacology, University of Heidelberg. We also want to thank Dr Karl Brehmer for help in preparing this study.

References

- McGue M, Vaupel JW, Holm N, Harvald B: Longevity is moderately heritable in a sample of Danish twins born 1870–1880. *J Gerontol* 1993; **48**: B237–B244.
- Herskind AM, McGue M, Holm NV, Sorensen TI, Harvald B, Vaupel JW: The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. *Hum Genet* 1996; **97**: 319–323.
- Yashin AI, Iachine IA: Genetic analysis of durations, Correlated frailty model applied to survival Danish Twins. *Genet Epidemiol* 1995; **12**: 529–538.
- Wright, S: *Evolution and the Genetics of Populations. Vol. 3: Experimental Results and Evolutionary Deductions*. Chicago: The University of Chicago Press, 1977, pp 474–497.
- De Benedictis G, Tan Q, Jeune B *et al*: Recent advances in human gene-longevity association studies. *Mech Ageing Dev*. 2001; **1**: 909–920.
- Schachter F, Faure-Delaneff L, Guenot F *et al*: Genetic associations with human longevity at the APOE and ACE loci. *Nat Genet* 1994; **6**: 29–32.
- Kervinen K, Savolainen MJ, Salokannel J *et al*: Apolipoprotein E and B polymorphisms—longevity factors assessed in nonagenarians. *Atherosclerosis* 1994; **105**: 89–95.
- Louhija J, Miettinen HE, Kontula K, Tikkanen MJ, Miettinen TA, Tilvis RS: Aging and genetic variation of plasma apolipoproteins. Relative loss of the apolipoprotein E4 phenotype in centenarians. *Arterioscler Thromb* 1994; **14**: 1084–1089.
- Zhang JG, Ma YX, Wang CF *et al*: Apolipoprotein E and longevity among Han Chinese population. *Mech Ageing Dev* 1998; **104**: 159–167.
- Gerdes LU, Jeune B, Ranberg KA, Nybo H, Vaupel JW: Estimation of apolipoprotein E genotype-specific relative mortality risks from the distribution of genotypes in centenarians and middle-aged men: apolipoprotein E gene is a “frailty gene,” not a “longevity gene”. *Genet Epidemiol* 2000; **19**: 202–210.
- Piegorsch WW, Weinberg CR, Taylor JA: Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. *Stat Med* 1994; **13**: 153–162.
- Khoury MJ, Flanders WD: Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! *Am J Epidemiol* 1996; **144**: 207–213.
- Yang Q, Khoury MJ, Sun F, Flanders WD: Case-only design to measure gene-gene interaction. *Epidemiology* 1999; **10**: 167–170.
- Fisher LD, Van Belle G: *Biostatistics: A methodology for the health sciences*. N.Y.: John-Wiley & Sons, 1993, pp 191–200.
- De Benedictis G, Carotenuto L, Carrieri G *et al*: Gene/longevity association studies at four autosomal loci (REN, THO, PARP, SOD2). *Eur J Hum Genet* 1998; **6**: 534–541.
- Franceschi C, Motta L, Valensin S *et al*: Do men and women follow different trajectories to reach extreme longevity? Italian Multicenter Study on Centenarians (IMUSCE). *Ageing (Milano)* 2000; **12**: 77–84.
- De Benedictis G, Rose G, Carrieri G *et al*: Mitochondrial DNA inherited variants are associated with successful aging and longevity in humans. *FASEB J* 1999; **13**: 1532–1536.
- Yashin AI, De Benedictis G, Vaupel JW *et al*: Genes and longevity: Lessons from studies on centenarians. *J Gerontol* 2000; **55**: B1–B10.

- 19 Tan Q, De Benedictis G, Yashin AI *et al*: Measuring the genetic influence in modulating human life span: Gene-environment and gene-sex interactions. *Biogerontology* 2001; **2**: 141–153.
- 20 Peters J, Kranzlin B, Schaeffer S *et al*: Presence of renin within intramitochondrial dense bodies of the rat adrenal cortex. *Am J Physiol* 1996; **271**: E439–E450.
- 21 Clausmeyer S, Sturzebecher R, Peters J: An alternative transcript of the rat renin gene can result in a truncated prorenin that is transported into adrenal mitochondria. *Circ Res* 1999; **84**: 337–344.
- 22 Clausmeyer S, Reinecke A, Farrenkopf R, Unger T, Peters J: Tissue-specific expression of a rat renin transcript lacking the coding sequence for the prefragment and its stimulation by myocardial infarction. *Endocrinology* 2000; **141**: 2963–2970.
- 23 De Benedictis G, Carrieri G, Garasto S *et al*: Does a retrograde response in human aging and longevity exist? *Exp Gerontol* 2000; **35**: 795–801.
- 24 Yang Q, Khoury MJ, Flanders WD: Sample size requirements in case-only designs to detect gene-environment interaction. *Am J Epidemiol* 1997; **146**: 713–720.
- 25 Kannisto V: *Development of oldest-old mortality, 1950–1990: evidence from 28 developed countries*. Odense: Odense University Press; Odense, 1994, pp 51–54.
- 26 Vaupel JW, Carey JR, Christensen K *et al*: Biodemographic trajectories of longevity. *Science* 1998; **280**: 855–860.
- 27 Khoury MJ: Genetic epidemiology; in: Rothman KJ, Greenland S, (eds) *Modern Epidemiology*. Philadelphia: Lippincott-Raven Publisher, 1998, pp 615–617.