

## Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Pharoah PDP, Antoniou AC, Easton DF, Ponder BAJ. Polygenes, risk prediction, and targeted prevention of breast cancer. *N Engl J Med* 2008;358:2796-803.

## Pharoah 07-08739

### Appendix

For a single allele (i) of frequency  $p$ , relative risk  $R$  and log risk  $r$  the population attributable fraction is given by

$$\frac{2p(1-p)(R-1)}{2p(1-p)(R-1)+1} + \frac{(1-p)^2(R^2-1)}{(1-p)^2(R^2-1)+1}$$

The variance ( $V_i$ ) of the risk distribution due to that allele is given by

$$V_i = (1-p)^2 E^2 + 2p(1-p)(r-E)^2 + p^2(2r-E)^2$$

Where  $E$  is the expected value of  $r$  given by

$$E = 2p(1-p)r + 2p^2r$$

For multiple risk alleles the distribution of risk in the population tends towards the normal with variance

$$V = \sum V_i$$

The total genetic variance ( $V$ ) for all susceptibility alleles has been estimated to be 1.44 from twin and family studies. Thus the fraction of the genetic risk explained by a single allele is given by

$$V_i / V$$

**Supplementary table: Breast cancer incidence and all cause mortality data on which absolute risk estimates are based <sup>†</sup>.**

<b>Age</b>	<b>Breast cancer incidence (per 100,000 pop)</b>	<b>All cause mortality (per 100,000 pop)</b>
15-19	0.2	26
20-24	1.2	29
25-29	8.3	38
30-34	26	53
35-39	63	79
40-44	117	133
45-49	183	211
50-54	217	348
55-59	258	551
60-64	309	918
65-69	304	1592
70-74	309	2691
75-79	328	4310

<sup>†</sup> Office for National Statistics. Cancer Statistics registrations: registrations of cancer diagnosed in 2004, England.

Series MB1 No. 35. London, 2004