

Research article

Log odds of carrying an Ancestral Mutation in *BRCA1* or *BRCA2* for a Defined personal and family history in an Ashkenazi Jewish woman (LAMBDA)

Carmel Apicella^{1,2}, Lesley Andrews³, Shirley V Hodgson⁴, Sheila A Fisher⁴, Cathryn M Lewis⁴, Ellen Solomon⁴, Katherine Tucker³, Michael Friedlander³, Agnes Bankier⁵, Melissa C Southey^{2,6}, Deon J Venter^{2,6} and John L Hopper¹

¹Centre for Genetic Epidemiology, The University of Melbourne, Carlton, Victoria, Australia

²Peter MacCallum Cancer Institute, East Melbourne, Victoria, Australia

³Prince of Wales Hospital, Randwick, New South Wales, Australia

⁴Division of Medical and Molecular Genetics, Guy's, King's and St Thomas' School of Medicine, London, UK

⁵Victorian Clinical Genetics Service, Royal Children's Hospital, Parkville, Victoria, Australia

⁶Department of Pathology, The University of Melbourne, Parkville, Victoria, Australia

Corresponding author: John L Hopper (e-mail j.hopper@unimelb.edu.au)

Received: 30 May 2003 Revisions requested: 21 Jul 2003 Revisions received: 31 Jul 2003 Accepted: 31 Jul 2003 Published: 28 Aug 2003

Breast Cancer Res 2003, **5**:R206-R216 (DOI 10.1186/bcr644)

© 2003 Apicella *et al.*, licensee BioMed Central Ltd (Print ISSN 1465-5411; Online ISSN 1465-542X). This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Introduction: Ancestral mutations in *BRCA1* and *BRCA2* are common in people of Ashkenazi Jewish descent and are associated with a substantially increased risk of breast and ovarian cancer. Women considering mutation testing usually have several personal and family cancer characteristics, so predicting mutation status from one factor alone could be misleading. The aim of this study was to develop a simple algorithm to estimate the probability that an Ashkenazi Jewish woman carries an ancestral mutation, based on multiple predictive factors.

Methods: We studied Ashkenazi Jewish women with a personal or family history of breast or ovarian cancer and living in Melbourne or Sydney, Australia, or with a previous diagnosis of breast or ovarian cancer and living in the UK. DNA samples were tested for the germline mutations 185delAG and

5382insC in *BRCA1*, and 6174delT in *BRCA2*. Logistic regression was used to identify, and to estimate the predictive strength of, major determinants.

Results: A mutation was detected in 64 of 424 women. An algorithm was developed by combining our findings with those from similar analyses of a large study of unaffected Jewish women in Washington. Starting with a baseline score, a multiple of 0.5 (based on the logistic regression estimates) is added for each predictive feature. The sum is the estimated log odds ratio that a woman is a carrier, and is converted to a probability by using a table. There was good internal consistency.

Conclusions: This simple algorithm might be useful in the clinical and genetic counselling setting. Comparison and validation in other settings should be sought.

Keywords: Ashkenazi Jews, *BRCA1*, *BRCA2*, breast cancer, mutations, ovarian cancer

Introduction

On average, about 1 in 40 people of Ashkenazi Jewish descent carry a copy of one of the ancestral mutations 185delAG and 5382insC in *BRCA1* and 6174delT in *BRCA2* [1]. Female carriers have a high lifetime risk of

breast cancer estimated to be in the range of 30–60%, and about a 20% lifetime risk of ovarian cancer [1,2]. The increased risk of breast cancer depends on age, and perhaps gene [2], and is between 3-fold and 20-fold. In theory about 20% of all breast cancers in Ashkenazi

LAMBDA = Log odds of carrying an Ancestral Mutation in *BRCA1* or *BRCA2* for a Defined personal and family cancer history in an Ashkenazi Jewish woman.

Jewish women would be attributable to these mutations, so it is important to be able to identify the women most likely to carry this high genetic cancer risk.

Previous studies, typically of clinic-based samples, have identified several factors that individually predict a woman's mutation status. Personal factors include a previous history of breast or ovarian cancer, especially if breast cancer was diagnosed at a relatively early age, and bilateral disease or both cancers in the same woman [3–10]. The occurrence of these factors in close relatives has also been found in general to be predictive, but the strength of prediction is weaker and the conclusions are less definitive.

The aim of this paper was to try to develop a simple yet reliable model for estimating the probability that an Ashkenazi Jewish woman carries a copy of one of the ancestral mutations in *BRCA1* or *BRCA2* that will be applicable across the full spectrum of personal and family history. Note that unless the data used to derive such a model come from a random population sample, and for the Ashkenazim this is difficult to achieve, any prediction of the absolute probability of being a carrier cannot be accurate unless it takes into account all predictive features. This is because individuals concerned about their mutation status often have multiple predictive features. Predictions based on a consideration of one factor alone could be misleading.

Previous logistic regression analyses

A few studies have considered multiple predictive factors together in developing a predictive statistical model, and included Ashkenazi women [3,11–15]. The largest was of more than 5000 volunteer men and women of Ashkenazi Jewish descent living in Washington [3]. As a consequence of the sampling strategy of that study, which was designed to give population-based estimates of the cancer risks in mutation carriers, the great majority of participants had no personal or family history of breast or ovarian cancer. There were few women with the rare predictive features, especially among those who were affected.

Logistic regression models were fitted to data from the Washington study published in tabular form [13,16], and estimated the predictive strength of some characteristics of family breast cancer history while adjusting for age at diagnosis in women with breast or ovarian cancer or age at testing in unaffected women. For women with a previous diagnosis of breast cancer, the probability of being a carrier decreased successively as the age at diagnosis increased from under 40 years to 40–49 years to 50–59 years to 60 years and over. The log (to the base e) odds decreased by about 1 unit (log odds=0.97; odds ratio=2.6) per age group. For women with no previous diagnosis of breast or ovarian cancer, the probability of being a carrier decreased successively as the age at

testing increased from under 40 years to 40–49 years to 50–59 years to 60 years and over. This time the log odds decreased by about 0.5 unit (log odds=0.43; odds ratio=1.5) per age group.

An important point revealed by this modelling was that, on the logistic scale, the effects associated with breast cancer in first-degree relatives seemed to be independent of the disease status of the tested individual [13,16]. We shall exploit this feature in our LAMBDA model below. Each affected first-degree relative increased the log odds of being a carrier by about 1 unit (log odds=0.99; odds ratio=2.7). Similarly, the log odds of being a carrier increased by about 1.5 units (log odds=1.49; odds ratio=4.4) if any first-degree relative had been diagnosed before 50 years of age. Similar estimates were reported by a Canadian study [10]. The Washington data were not presented in a form that allowed concurrent modelling of the effects of both the number of affected relatives and the presence of early age at diagnosis in a relative. For ovarian cancer, having a first-degree relative with that disease increased the log odds of being a carrier by about 1.5 units (log odds=1.41; odds ratio=4.1) [13].

A similar logistic regression analysis was conducted of data on 171 Ashkenazi Jewish women living in the UK who had a previous diagnosis of breast cancer [14]. In the present paper the data used in that UK analysis, including 13 women who had a previous diagnosis of ovarian cancer only, are combined with data from Ashkenazi Jewish women living in Australia who have either a personal history or a family history of breast or ovarian cancer. The features that are most predictive of mutation status when all are considered together have been identified, and the extent to which each feature is predictive of mutation status have been quantified.

We also compared and combined parameter estimates from our new analyses with those from our logistic regression analyses of data on unaffected women from the Washington study [13], and derived a simple algorithm to estimate the probability that an Ashkenazi Jewish woman carries one of the three ancestral mutations in *BRCA1* and *BRCA2*. The algorithm therefore applies to women with or without a personal history of breast or ovarian cancer. The method of calculating the probability is straightforward: each predictive feature has an associated score, and the sum of these scores is converted to a probability by reference to a table. It should therefore be useful to a broad range of health professionals.

Methods

Ethical approval

The study was approved in Australia by the Human Research Ethics Committees of the Peter MacCallum Cancer Institute (Melbourne), The Prince of Wales Hospital

(Sydney) and The University of Melbourne, and in the UK by the Local Research Ethics Committee at Guy's and St Thomas' Trust.

Subjects

In Australia, eligible women were self-identified as being of Ashkenazi Jewish descent, living in Melbourne or Sydney, who had reported either having had a previous diagnosis of breast or ovarian cancer themselves, or having a first-degree or second-degree relative who had been diagnosed with breast or ovarian cancer. They were recruited between 1996 and 2000 through announcements in the media (local Jewish newspapers, 32%; other newspapers, 3%), approaches including brochures to general practitioners, gynaecologists, breast surgeons and oncologists (19%), the family cancer genetics clinics associated with this study (9%), information evenings (11%), and friends or relatives (26%).

In the UK, eligible women were self-identified as being of Ashkenazi Jewish descent and had a previous diagnosis of breast or ovarian cancer. Recruitment was through clinicians and Jewish agencies, lectures, broadcasts and newspaper articles [14].

For the purposes of the analyses below, the first woman in a family recruited was nominated as the proband. In Australia there were 240 probands, of whom 148 had no previous diagnosis of breast or ovarian cancer, 81 had breast cancer only, 9 had ovarian cancer only, and 2 had both breast and ovarian cancer. In UK there were 184 probands, of whom 169 had breast cancer, 13 had ovarian cancer, and 2 had both breast and ovarian cancer.

Information on personal and family cancer history

Participants in Australia were administered, face to face, the structured questionnaire developed by the Cooperative Family Registry for Breast Cancer Studies [17]. This instrument asked about past cancer history and Ashkenazi Jewish descent. Personal and family cancer histories were self-reported and not confirmed. Family history was recorded by first identifying all first-degree and second-degree adult relatives, and then asking for each about any cancer history.

Participants in the UK gave a similarly detailed family history to at least second-degree relatives, and information was collected on all cancer cases. Breast and ovarian cancer in the proband was confirmed through the relevant clinician, but no confirmation was attempted for reports in other family members [14].

Mutation testing

Blood samples were collected from all probands. In Australia, exons 2 and 20 of *BRCA1* and a fragment of exon 11 of *BRCA2* were amplified and manually sequenced.

These regions encompassed the mutations 185delAG and 5382insC in *BRCA1* and 6174delT in *BRCA2* [18]. In the UK, a heteroduplex assay was used to screen for the three ancestral mutations, as described by Hodgson and colleagues [14].

Statistical methods

Unconditional multiple linear logistic regression was used to model the probability that the proband was a mutation carrier (in other words that she carried a copy of one of the three specific mutations) as a function of her personal and family history, and either age at diagnosis of first cancer if the index person had a previous diagnosis of breast or ovarian cancer, or age at testing if unaffected (that is, not reporting a prior diagnosis of breast or ovarian cancer). The model was a linear function of the logit of the probability; ie $\log(p/[1-p]) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots$, where log signifies logarithm to the base e, p is the probability, and for each $i=1, 2, 3 \dots$ each x_i represents a potentially predictive factor that can either take two values (0 if absent or 1 if present) or is a number of affected relatives (0, 1, 2, ...), and β_i is the corresponding logistic regression coefficient. Therefore $\exp(\beta_i)$ is the ratio of the odds for being a mutation carrier comparing a woman with the i th characteristic with one without that characteristic, or having one more affected relative, given that the two women are the same for all other characteristics. If p is small, $\exp(\beta_i)$ is approximately the relative risk of being a carrier associated with factor i . Logistic regression analyses were conducted with the software package GLIM [19]. Model selection was based on forwards selection and confirmed by backwards elimination, with the fits of nested models compared by the likelihood ratio test based on asymptotic likelihood theory. Tests for interaction terms were also performed. Following convention, all nominal P values were two-tailed. In general, proportions were compared with Pearson χ^2 -test statistics, or Fisher's exact test for small cells. During modelling, trends in proportions across groups were assessed by linear logistic regression by fitting single parameters to represent linear effects (on the logit scale) with the underlying variable considered as a continuum.

Characteristics specific to the probands were having a previous diagnosis of primary breast cancer or of ovarian cancer, and of having bilateral breast cancer. To be consistent with the presentation of the Washington study [3], age at diagnosis for affected women, and age at testing for unaffected women, were both categorised as under 40 years, 40–49 years, 50–59 years, or 60 years and over (and classified as 1–4 for the assessment of trends). These age groupings were also applied to the age at onset of cancers in relatives. Family histories were categorised according to the number of affected first-degree relatives (mothers, sisters and daughters), and the number of affected second-degree relatives (maternal and paternal

grandmothers and aunts, nieces, granddaughters and half-sisters), separately for breast cancers and ovarian cancers. Bilateral breast cancer in any first-degree or second-degree relative was also included. The development of the LAMBDA model with the use of the results of the logistic regression analyses of our data and the Washington study data is described in detail in the Results section.

The internal goodness of fit of the LAMBDA model was assessed several ways, as described by Cox and Snell [20]. For each woman, let λ_i be her LAMBDA score and p_i the corresponding predicted probability. First, the observed proportions of carriers in various categories of p_i were compared with the expected proportions by using standard contingency table methods with the Pearson χ^2 test statistic. Second, the expected total number of carriers predicted by the model was calculated by summing over the predicted probabilities (that is, $E_0 = \sum p_i$). This was compared with the observed number of carriers, O_0 , by calculating the statistic $Z_0 = (O_0 - E_0) / [\text{var}(E_0)]^{1/2}$, where $\text{var}(E_0) = \sum p_i(1 - p_i)$, which has a standard normal distribution under the null hypothesis of no bias. Deviations test whether the predicted values were systematically too high or too low. Third, we let $Y_i = \log(2p_i)$ if the i th woman was a carrier and $\log[2(1 - p_i)]$ otherwise, $E_1 = n \log 2 + \sum p_i \log(p_i) + \sum (1 - p_i) \log(1 - p_i)$ and $O_1 = \sum Y_i$. The statistic $Z_1 = (O_1 - E_1) / [\text{var}(E_1)]^{1/2}$, where $\text{var}(E_1) = \sum p_i(1 - p_i) \{\log[p_i / (1 - p_i)]\}^2$ has a standard normal distribution under the null hypothesis, and deviations test whether the predicted values were too clustered or too dispersed.

Results

Descriptive statistics for the Australian and UK data

Of the 424 probands, 64 (15%) had a germline ancestral mutation; 26 had 185delAG and 10 had 5382insC in *BRCA1*, whereas 28 had 6174delT in *BRCA2*. Table 1 shows that there were 148 probands with no previous history of breast or ovarian cancer (median age 44 years). Of the remainder, 250 had breast cancer only, 22 had ovarian cancer only, and 4 had both breast and ovarian cancer. The median ages at diagnosis of these three groups were 49, 44 and 59 years, respectively.

Tables 2 and 3 show that, of the probands with no previous diagnosis of breast or ovarian cancer, 68% had at least one first-degree relative and 48% at least one second-degree relative with breast cancer, and 12% had at least one first-degree relative with ovarian cancer. Of the probands with a previous diagnosis of breast and/or ovarian cancer, the percentages with the above family histories were 29% and 34% for breast cancer and 7% for ovarian cancer, respectively. That is, as a consequence of the design, a smaller proportion of the probands with breast and/or ovarian cancer had these family history cate-

gories than that of the probands with no previous diagnosis ($P=0.001$, 0.004 and 0.04, respectively). The crude mutation frequencies presented here and below must therefore be interpreted with care, taking into account the existence and differing extents of family history that we are referring to, as well as the 'age at testing' in unaffected women and the 'age at diagnosis' in affected women.

Table 1 shows that the proportion of mutation carriers differed by the proband's cancer status, increasing from 8% in those with no previous diagnosis of breast or ovarian cancer to 17% in those with breast cancer only, to 32% in those with ovarian cancer, and to 75% in those four with both breast and ovarian cancer ($P=0.004$). Table 1 also shows that, in probands with a previous diagnosis of breast cancer, there was a strong negative effect of age at diagnosis: the proportion of carriers went from 38% for 40 years or younger, to 21% for 40–49 years, to 7% for 50 years or older ($P=0.0001$). In probands with ovarian cancer, however, no effect of age at diagnosis on mutation status was observed ($P=0.09$), although it should be noted that no carriers were observed among the eight ovarian cancer cases diagnosed before the age of 50 years. For unaffected probands, Table 1 shows that the proportion of mutation carriers was 12% in those under 40 years compared with 6% in those 40 years or older ($P=0.25$).

Table 2 shows that the proportion of mutation carriers generally increased as the number of first-degree relatives with breast cancer increased. For example, in probands with a previous diagnosis of breast cancer, it increased from 13% to 24% to 44% as the number increased from zero to one to two or more ($P=0.004$), whereas in probands with no previous diagnosis of breast or ovarian cancer it increased from 2% to 11% as the number increased from zero to one or more ($P=0.06$). It also shows that the proportion was relatively stable according to the number of second-degree relatives with breast cancer. In probands with a previous diagnosis it decreased from 18% to 14% as the number increased from zero to one or more, whereas in those with no previous diagnosis it increased from 7% to 8%, respectively.

Table 3 shows that the proportion of mutation carriers was generally greater if the proband had a family history of ovarian cancer, irrespective of whether it was in a first-degree or a second-degree relative. In probands with breast cancer only, it was 13% in those with no family history, compared with 32% in those with a first-degree or second-degree relative with ovarian cancer ($P=0.002$).

Given that age at diagnosis in probands with breast cancer was predictive of mutation status, it is logical also to examine age at diagnosis of breast cancer in affected relatives. Table 4 shows that in unaffected probands

Table 1

Proportions of probands with a germline ancestral mutation in *BRCA1* or *BRCA2* out of total tested (percentages in parentheses), by personal cancer history and age at testing in unaffected women, or age at first diagnosis in affected women

Cancer status	Age group*				Total	Median age (years)
	Under 40 years	40–49 years	50–59 years	60 years or older		
No breast or ovarian cancer	6/50 (12%)	3/48 (6%)	2/33 (6%)	1/17 (6%)	12/148 (8%)	44
Breast cancer only	15/40 (38%)	19/89 (21%)	4/68 (6%)	4/53 (8%)	42/250 (17%)	49
Ovarian cancer only	0/8 (0%)	3/6 (50%)	2/5 (40%)	2/3 (67%)	7/22 (32%)	44
Breast and ovarian cancer	0/0 (–)	1/1 (100%)	1/1 (100%)	1/2 (50%)	3/4 (75%)	59
Total	21/98 (21%)	26/144 (18%)	9/107 (8%)	8/75 (11%)	64/424 (15%)	

*Age at testing for unaffected women, age at diagnosis in affected women

Table 2

Proportions of probands with a germline ancestral mutation in *BRCA1* or *BRCA2* out of total tested (percentages in parentheses), by personal cancer history and number of first-degree relatives with breast cancer, and number of second-degree relatives with breast cancer

Cancer status	No. of first-degree relatives with breast cancer				No. of second-degree relatives with breast cancer					
	0	1	2	3	0	1	2	3	4	5
No breast or ovarian cancer	1/47 (2%)	11/94 (12%)	0/7 (0%)	–	6/77 (7%)	5/56 (9%)	1/14 (7%)	0/1 (0%)	–	–
Breast cancer only	22/175 (13%)	16/66 (24%)	4/8 (50%)	0/1 (0%)	30/165 (18%)	9/64 (14%)	1/15 (7%)	1/5 (20%)	–	1/1 (100%)
Ovarian cancer only	6/17 (35%)	1/5 (20%)	–	–	5/17 (29%)	2/4 (50%)	–	–	0/1 (0%)	–
Breast and ovarian cancer	2/3 (67%)	1/1 (100%)	–	–	1/1 (100%)	2/3 (67%)	–	–	–	–

Table 3

Proportions of probands with a germline ancestral mutation in *BRCA1* or *BRCA2* out of total tested (percentages in parentheses), by personal cancer history and number of first-degree relatives with ovarian cancer, and with number of second-degree relatives with ovarian cancer

Cancer status	No. of first-degree relatives with ovarian cancer			No. of second-degree relatives with ovarian cancer		
	0	1	2	0	1	2
No breast or ovarian cancer	10/131 (8%)	8/89 (9%)	0/7 (0%)	11/144 (8%)	1/4 (25%)	–
Breast cancer only	34/237 (14%)	13/42 (31%)	–	38/242 (16%)	4/5 (80%)	0/3 (0%)
Ovarian cancer only	4/18 (22%)	3/4 (75%)	–	5/20 (25%)	2/2 (100%)	–
Breast and ovarian cancer	3/3 (100%)	0/1 (0%)	–	3/4 (75%)	–	–

(overall proportion of carriers 8%), the proportion of mutation carriers decreased with age at diagnosis of breast cancer in a first-degree relative from 20% if diagnosed before age 40, to 9% if diagnosed at a later age; however, the difference was not significant ($P=0.4$). For second-degree relatives, these proportions were 17% and 7%,

respectively ($P=0.4$). In probands with breast cancer only (overall proportion of carriers 17%), these proportions were 50%, 30%, 33% and 17% for any first-degree relative diagnosed before 40, 40–49, 50–59 and 60 or more years, respectively (P for trend=0.09), and 38%, 25%, 13% and 10% for any second-degree relative diagnosed

Table 4

Proportions of probands with a germline ancestral mutation in *BRCA1* or *BRCA2* out of total tested (percentages in parentheses), by personal cancer history and having a relative with breast cancer with given age at diagnosis, for first-degree relatives and for second-degree relatives

Cancer status	Age at diagnosis for first-degree relatives with breast cancer				Age at diagnosis for second-degree relatives with breast cancer			
	<40 years	40–49 years	50–59 years	60 years or older	<40 years	40–49 years	50–59 years	60 years or older
No breast or ovarian cancer	3/15 (20%)	4/29 (14%)	4/34 (12%)	0/28 (0%)	1/6 (17%)	2/23 (9%)	2/19 (11%)	1/32 (3%)
Breast cancer only	3/6 (50%)	7/23 (30%)	7/21 (33%)	5/30 (17%)	3/8 (38%)	4/16 (25%)	3/24 (13%)	5/52 (10%)
Ovarian cancer only	1/1 (100%)	0/1 (0%)	0/1 (0%)	0/2 (0%)	0/0 (–)	0/2 (0%)	2/2 (100%)	1/3 (33%)
Breast and ovarian cancer	0/0 (–)	–	1/1 (100%)	–	–	–	–	–

in the same age groups (P for trend=0.03). There were too few probands with ovarian cancer to identify any influence of age at diagnosis of breast cancer in their relatives. For ovarian cancer in relatives, there was no evidence that the age at diagnosis influenced mutation status (data not shown).

Of probands with bilateral breast cancer, 34% (10 of 29) had a mutation, in comparison with 14% of those with unilateral breast cancer only ($P=0.01$). Of probands who had a first-degree or second-degree relative with bilateral breast cancer, 24% (7 of 29) had a mutation, in comparison with 14% of those who did not have a first-degree or second-degree relative with bilateral breast cancer ($P=0.2$).

Logistic regression analyses of the combined Australian and UK data

Given the possible confounding between the potential predictors of mutation status identified above, it is essential that statements about the absolute and relative probabilities – and therefore the predictive effects of factors – be based on an analysis that takes into account all factors simultaneously. Table 5 shows the parameter estimates and standard errors from a multiple linear logistic regression analysis of a model that included all the listed variables. It shows that, on the log odds scale, the strongest predictors of mutation status were breast cancer in the proband if age at diagnosis was under 50 years ($P<0.002$), breast cancer in a first-degree relative if age at diagnosis was under 60 years ($P<0.003$), and irrespective of age at diagnosis, ovarian cancer in the proband ($P=0.0001$), in a first-degree relative ($P=0.003$), or in a second-degree relative ($P=0.04$). There was marginal evidence of an additional strong effect of a bilateral breast cancer at any age ($P=0.07$). All these factors would more than double the odds of being a mutation carrier. The effects of breast cancer in the proband diagnosed when aged 50–59 years, and when 60 years or

older, were more modest but imprecise ($P=0.4$ and 0.3 , respectively), and there was marginal evidence of breast cancer diagnosed in a second-degree relative before the age of 40 years ($P=0.13$). No evidence was found for an independent effect of breast cancer in a first-degree relative diagnosed after the age of 60 years ($P=0.9$), or of bilateral breast cancer at any age in a relative ($P=0.9$). When the latter two effects were removed from the model displayed in Table 5, the estimates and standard errors of the remaining factors were virtually unchanged. There was also no evidence of interaction effects, on the logit scale, between the factors listed in Table 5. Furthermore, there was no evidence that the estimates differed by country (UK versus Australia).

There was marginal evidence within this data set for a negative effect of age at testing in probands without a previous diagnosis of breast or ovarian cancer. However, almost all had a family history of breast or ovarian cancer, and after allowing for these factors the effect of age at testing was no longer nominally significant.

The probability that a proband with a given set of personal and family characteristics is a mutation carrier can be estimated from the model fit shown in Table 5. Start with a log odds score of θ =the baseline coefficient, -3.767 . For each personal characteristic of the proband, add the respective regression coefficient to θ . For each family history characteristic, multiply the respective coefficient by the number of affected relatives and add that to θ . The final cumulative log odds score, θ , is equivalent to the probability $\exp(\theta)/[1 + \exp(\theta)]$.

Development of the LAMBDA model

Our aim was to develop a simple model for predicting whether or not an Ashkenazi woman had a germline mutation. Analyses of the Washington data [13,16] have shown that, when modelled on the log odds scale, the effects of family history were independent of the

Table 5

Logistic regression coefficients (β), standard errors and nominal significance levels for potentially predictive factors for an Ashkenazi Jewish woman having an ancestral mutation in *BRCA1* or *BRCA2*, derived from multiple linear logistic regression

Characteristic	β	Standard error	P
Baseline	-3.767	0.436	0.001
Proband breast cancer diagnosis <40 years	2.570	0.523	0.0001
Proband breast cancer diagnosis 40–49 years	1.707	0.455	0.0002
Proband breast cancer diagnosis 50–59 years	0.493	0.602	0.4
Proband breast cancer diagnosis 60+ years	0.578	0.598	0.3
Proband bilateral breast cancer at any age	0.912	0.500	0.07
Proband ovarian cancer at any age	2.621	0.570	0.0001
For each relative with cancer			
1st degree, breast cancer diagnosis <40 years	1.790	0.637	0.005
1st degree, breast cancer diagnosis 40–49 years	1.154	0.438	0.008
1st degree, breast cancer diagnosis 50–59 years	1.227	0.419	0.003
1st degree, breast cancer diagnosis 60+ years	0.033	0.482	0.9
1st degree, ovarian cancer at any age	1.339	0.454	0.003
2nd degree, breast cancer diagnosis <40 years	0.892	0.591	0.13
2nd degree, ovarian cancer at any age	1.004	0.489	0.04
1st or 2nd degree, bilateral breast cancer at any age	-0.088	0.572	0.9

proband's cancer status. Furthermore, as noted in the Introduction, the absolute magnitudes of the effects of age at diagnosis of breast cancer in the proband, and of early diagnosis of breast cancer in a first-degree relative, had a log odds ratio of about 1.0, whereas for age at testing in unaffected probands the absolute effects were about 0.5.

We therefore tried to simplify the logistic regression model fit shown in Table 5, and combine information from the Washington study analyses, to produce a model for the Log odds of the probability of carrying an Ancestral Mutation in *BRCA1* or *BRCA2* for a Defined personal and family cancer history in an Ashkenazi Jewish woman, called LAMBDA and represented by the symbol λ . The resulting model is shown in Figure 1. The total LAMBDA score is calculated on the log odds scale and is converted to a probability by the formula $\exp(\lambda)/[1 + \exp(\lambda)]$. This is easily done with the LAMBDA conversion table at the bottom of Figure 1, given that only multiples of 0.25 are valid scores for λ .

For an affected proband, the effect of each age group at diagnosis of breast cancer is 1 unit, which is consistent with the Washington study and our findings for age groups under 50 years. The effect of diagnosis of ovarian cancer is 3 units. The additional effect of a bilateral breast

cancer is 1 unit (see Table 5). Note that if a woman has had both breast and ovarian cancer, her λ score will be that due to her breast cancer, and decided by the age at diagnosis, plus 3 for the diagnosis of ovarian cancer.

Our study was of women with a personal or family history of breast or ovarian cancer; it was therefore limited in making any inference about unaffected women. For example, the effectively null result for age at testing might be a consequence of limited power. In contrast, the Washington study contained 3419 unaffected women, 80% without a family history [3]. It found a decrease in log odds ratio of about 0.5 for each successive decade [13], so this was built into the LAMBDA model.

The effect for each cancer in a first-degree relative has been taken to be one-half that of the same condition in the proband herself, which is generally consistent with Table 2 except for breast cancer at older ages. A term has been included for each breast cancer in a second-degree relative, provided that it was diagnosed before the age of 40 years, given the strong effect of this characteristic in probands and first-degree relatives and the marginal evidence of such an effect in Table 2. A term for each ovarian cancer in a second-degree relative has also been included, given the evidence in Table 3.

Figure 1

Probability that an Ashkenazi Jewish Woman Carries an Ancestral Mutation in BRCA1 or BRCA2

For each personal feature listed below, identify those relevant to the consultand and then write the corresponding coefficient in the shaded box alongside.

For each family history feature listed below, identify those relevant to the consultand. First multiply the corresponding coefficient by the number of affected relatives with that feature, and then write this in the shaded box alongside.

Add all the scores to obtain LAMBDA, then use the LAMBDA CONVERSION TABLE to give the estimated PROBABILITY that an Ashkenazi Jewish woman carries an ancestral mutation in BRCA1 or BRCA2.

	coefficient		
BASELINE	-3.75	SCORE -3.75	
CONSULTAND HAS HAD BREAST OR OVARIAN CANCER			
Personal history features	diagnosed with a primary breast cancer before age 40.....	+3	} one of
	age 40 - 49.....	+2	
	age 50 - 59.....	+1	
	diagnosed with bilateral breast cancer at any age.....	+1	
	diagnosed with a primary ovarian cancer at any age.....	+3	
CONSULTAND HAS <u>NOT</u> HAD BREAST OR OVARIAN CANCER			
	age 40 - 49.....	-0.5	} one of
	age 50 - 59.....	-1.0	
	age 60 or older.....	-1.5	
FIRST DEGREE RELATIVES no. affected relatives			
Family history features	diagnosed with breast cancer before age 40.....[].....multiplied by.....	+1.5	
	age 40 - 49.....[].....multiplied by.....	+1	
	age 50 - 59.....[].....multiplied by.....	+0.5	
	diagnosed with ovarian cancer at any age.....[].....multiplied by.....	+1.5	
	SECOND DEGREE RELATIVES (on either side of the family)		
diagnosed with breast cancer before age 40.....[].....multiplied by.....	+0.5		
diagnosed with ovarian cancer at any age... ..[].....multiplied by.....	+1.0		
LAMBDA CONVERSION TABLE			
LAMBDA.....probability	LAMBDA.....probability	LAMBDA	
2.5 or more.....>90%	-0.5.....36%	[]	
1.5.....82%	-0.75.....32%	↓	
1.25.....78%	-1.0.....27%	PROBABILITY	
1.0.....73%	-1.25.....22%	[]	
0.75.....68%	-1.5.....18%		
0.5.....62%	-2.0.....12%		
0.25.....56%	-2.5.....8%		
0.....50%	-3.0.....5%		
-0.25.....44%	-3.75.....2%		

Estimation of the probability that an Ashkenazi Jewish woman carries an ancestral mutation in *BRCA1* or *BRCA2* based on her personal and family history of breast or ovarian cancer.

Table 6

Observed and expected numbers of mutation carriers, as predicted by the LAMBDA model, for categories based on predicted range of probability of being a carrier

λ	Predicted range		Numbers of carriers (percentage)	
	Probability (%)	Number of women	Observed	Expected
<-3.5	<2	128	5 (4%)	2.1 (2%)
-3.5 to -2.25	2-10	117	12 (10%)	7.2 (6%)
-2.0 to -1.25	10-25	84	11 (13%)	13.4 (16%)
-1.0 to -0.25	25-50	59	14 (24%)	20.1 (34%)
0.0 to 0.75	50-75	30	16 (53%)	18.2 (61%)
≥ 1.0	≥ 75	6	6 (100%)	5.1 (85%)
Total		424	64 (15%)	66.2 (16%)

Tests of internal consistency of the LAMBDA model

Table 6 shows that, on the basis of each woman's predicted λ score, there were an expected 66.2 carriers in comparison with the observed 64, giving no evidence for a systematic bias ($Z_0 = -0.35$, $p = 0.3$). It also shows that the expected and observed numbers of carriers were in good agreement, especially in women with a 50% or greater predicted probability of being a carrier. There was no evidence of a deviation greater than expected by chance ($\chi_5^2 = 10.0$, $p = 0.07$). There was no evidence that the predicted scores were too clustered or too dispersed ($Z_1 = -1.26$, $p = 0.2$).

Discussion

We have attempted to create a simple algorithm that can be used easily in the clinical setting to predict the probability that an Ashkenazi Jewish woman has inherited a high and detectable genetic risk of breast and ovarian cancer. We have used logistic regression, and being aware of the usual potential problems raised by Wacholder and Hartge [21], have carefully considered the dose relationships starting with detailed descriptive analyses. We have addressed internal consistency (given that the LAMBDA model was in part developed from our data) and did not find any evidence of systematic bias or under-dispersion or over-dispersion. In developing the LAMBDA model we have considered those personal and family cancer history features that have been reported to influence the probability of being a mutation carrier when considered on their own. Given that women considering genetic testing often have two or more such factors, it is important that a multivariate model be used for predicting mutation status, as we have done here.

We have relied not only on our own analyses, which provided the framework for modelling the effects of a previous

diagnosis of breast or ovarian cancer and the effects of various extents of family histories, but also on the large Washington study in which the vast majority of women were unaffected and did not have a family history, or at most a modest family history. In other words, that data set complements ours, and the LAMBDA model has integrated the two so as to provide valid estimates for women across a range of personal and family cancer histories.

Here we have taken an empirical approach to estimating a woman's mutation status. Another way is to specify a mathematical model and use a conditional likelihood approach, such as that used by Parmigiani and colleagues to develop BRCAPRO [22]. However, that model was based on the assumption that mutations in *BRCA1* or *BRCA2* account for all the aggregation of breast cancer in the family under consideration, and this is not true even in the current setting. Epidemiological studies involving mutation testing of cases unselected for their family history are showing that perhaps less than 20% of familial aggregation on a population basis can be attributed to deleterious mutations in these two genes [23]. It is important that such approaches build into their models other familial factors, such as other major genes and/or polygenic effects [24,25].

We have used individual characteristics in our model. Other empirical models have been proposed, but they have tended to use the broad family characteristics, such as the number of affected relatives or their average age at onset [11,12]. It is not clear what was meant by 'family' in the data sets used to derive those models, and given that families vary in their size and extent, this might be problematic in using those models in other settings. Furthermore, none of those models were based on substantial numbers of Ashkenazi Jewish women.

We have pooled the three ancestral mutations in our analyses, because if a woman has a sufficiently high probability of being a carrier to initiate mutation testing, the extra work and expense to test for all three mutations is relatively small given that she will have already gone through counselling and provided a blood sample from which DNA had been extracted. However, it would be of interest to consider the specific mutations alone, given that they might be associated with different risks of breast and ovarian cancer [2]. We have also not included third-degree relatives or breast cancer in male relatives, even though these might be predictive, because these features were either not evident or not recorded in our data sets. That might explain some of the apparent underestimation of mutation carriers in women with λ scores of less than -2 . For the same reason we have not addressed the issue of whether, if there are multiple cancers in the family (including the proband), they are in genetically related women (that is, on the 'same side of the family'). It would seem prudent to base the prediction of mutation status on cancers in genetically related women only.

Bilateral breast cancer is a difficult feature to model without knowing whether an affected woman has had unilateral or bilateral mastectomy, and this is further complicated when reference is to relatives. Nevertheless, our data suggest that having breast cancer in two close relatives is more predictive of mutation status than in having one relative with two breast cancers herself.

The families that we have studied here are those who have come forward from the Jewish community, and do not necessarily have a strong cancer family history. In extrapolating to families with stronger histories, it could be important to take into account whether affected relatives were on the same side of the family.

It is tempting to consider whether this LAMBDA model could be adapted to the estimation of *BRCA1* and/or *BRCA2* mutation status in non-Ashkenazi women. It is likely that the same predictors would apply, but given that mutations in these genes are rarer the magnitudes of their effects could be smaller (because they explain less familial aggregation) and the baseline would be considerably lower, perhaps by as much as 2 units on the LAMBDA scale. Another future development would be to build into the model for affected women those tumour characteristics that recent studies have found to be associated with mutations in *BRCA1* and *BRCA2* [26]. The combination of these features might increase the sensitivity and specificity of the LAMBDA model. It is also likely that this sort of modelling could be useful in the setting of mutations in other genes, such as the mismatch repair genes, and their relationship to a personal and family history of colorectal and associated cancers, including measures of microsatellite instability and immunohistochemistry status of tumours.

Conclusions

This model for predicting the probability that an Ashkenazi Jewish woman carries a founder mutation in *BRCA1* or *BRCA2* has been derived by using data from Ashkenazi Jewish women living in Australia and the UK, as well as data from a community study performed in Washington. The risk factors and their predictive strengths seem to give predictions consistent with pooled data from other Jewish populations, specifically from North America and Israel. The LAMBDA model is simple to use and might therefore have wide clinical utility, because it seems to capture the main features that predict mutation status.

It is important that future work both should compare the performance of this LAMBDA model with other models for predicting carrier status and should test the model in other settings to assess its generality. Careful examination of the model and fits, especially in subgroups, might lead to modifications and improvements, especially in making this concept suitable in a wider setting. With these aims, large validation studies are being undertaken with the use of, for example, the Northern American resources of the NIH-funded Breast Cancer Family Registry.

Competing interests

None declared.

Acknowledgements

We wish to express our appreciation to the women and their relatives who participated in this study. The Australian component of this study was funded by the United States National Cancer Institute, National Institutes of Health, under Request for Application CA-95-003, as part of the Breast Cancer Family Registries (CFRs), and through cooperative agreements with the Fox Chase Cancer Center; the Huntsman Cancer Institute; Columbia University; the Northern California Cancer Center; Cancer Care Ontario; and The University of Melbourne. It was also supported by the Australian National Health and Medical Research Council and the Peter MacCallum Cancer Research Institute. In the UK it was supported by the Dunhill Medical Trust.

References

1. Struewing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, Timmerman MM, Brody LC, Tucker MA: **The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews.** *N Engl J Med* 1997, **336**: 1401-1408.
2. Satagopan JM, Offit K, Foulkes W, Robson ME, Wacholder S, Eng CM, Karp SE, Begg CB: **The lifetime risks of breast cancer in Ashkenazi Jewish Carriers of BRCA1 and BRCA2 mutations.** *Cancer Epidemiol Biomark Prevent* 2001, **10**:467-473.
3. Hartge P, Struewing JP, Wacholder S, Brody LC, Tucker MA: **The prevalence of common BRCA1 and BRCA2 mutations among Ashkenazi Jews.** *Am J Hum Genet* 1999, **64**:963-970.
4. Robson M, Dabney MK, Rosenthal G, Ludwig S, Seltzer MH, Gilewski T, Haas B, Osborne M, Norton L, Gilbert F, Offit K: **Prevalence of recurring BRCA mutations among Ashkenazi Jewish women with breast cancer.** *Genet Test* 1997, **1**:47-51.
5. Fodor FH, Weston A, Bleiweiss IJ, McCurdy LD, Walsh MM, Tarter PI, Brower ST, Eng CM: **Frequency and carrier risk associated with common BRCA1 and BRCA2 mutations in Ashkenazi Jewish breast cancer patients.** *Am J Hum Genet* 1998, **63**:45-51.
6. Fitzgerald MG, MacDonald DJ, Kraimer M, Hoover I, O'Neil E, Unsal H, Silva-Arrieto S, Finkelstein DM, Beer-Romero P, Englert C, Sgroi DC, Smith BL, Younger JW, Garber JE, Duda RB, Mayzel KA, Isselbacher KJ, Friend SH, Haber DA: **Germ-line BRCA1**

- mutations in Jewish and non-Jewish women with early-onset breast cancer. *N Engl J Med* 1996, **334**:143-149.
7. Moslehi R, Chu W, Karlan B, Risch H, Fields A, Smotkin D, Ben-David Y, Rosenblatt J, Russo D, Schwartz P, Tung N, Warner E, Rosen B, Friedman J, Brunet JS, Narod SA: **BRCA1 and BRCA2 mutation analysis of 208 Ashkenazi Jewish women with ovarian cancer.** *Am J Hum Genet* 2000, **66**:1259-1272.
 8. Levy-Lahad E, Catane R, Eisenberg S, Kaufman B, Hornreich G, Lishinsky E, Shohat M, Weber BL, Beller U, Lahad A, Halle D: **Founder BRCA1 and BRCA2 mutations in Ashkenazi Jews in Israel: frequency and differential penetrance in ovarian cancer and in breast-ovarian cancer families.** *Am J Hum Genet* 1997, **60**:1059-1067.
 9. Abeliovich D, Kaduri L, Lerer I, Weinberg N, Amir G, Sagi M, Zlotogora J, Heching N, Peretz T: **The founder mutations 185delAG and 5382insC in BRCA1 and 6174delT in BRCA2 appear in 60% of ovarian cancer and 30% of early-onset breast cancer patients among Ashkenazi women.** *Am J Hum Genet* 1997, **60**:505-514.
 10. Warner E, Foulkes W, Goodwin P, Meschino W, Blondal J, Paterson C, Ozcelik H, Goss P, Allingham-Hawkins D, Hamel N, Di Prospero L, Contiga V, Serruya C, Klein M, Moslehi R, Honeyford J, Liede A, Glendon G, Brunet JS, Narod S: **Prevalence and penetrance of BRCA1 and BRCA2 gene mutations in unselected Ashkenazi Jewish women with breast cancer.** *J Natl Cancer Inst* 1999, **91**:1241-1247.
 11. Shattuck-Eidens D, Oliphant A, McClure M, McBride C, Gupte J, Rubano T, Pruss D, Tavtigian SV, Teng DH, Adey N, Staebell M, Gumpfer K, Lundstrom R, Hulick M, Kelly M, Holmen J, Lingenfelter B, Manley S, Fujimura F, Luce M, Ward B, Cannon-Albright L, Steele L, Offit K, Thomas A: **BRCA1 sequence analysis in women at high risk for susceptibility mutations. Risk factor analysis and implications for genetic testing.** *JAMA* 1997, **278**:1242-1250.
 12. Couch FJ, DeShano ML, Blackwood A, Calzone K, Stopfer J, Campeau L, Ganguly A, Rebbeck T, Weber BL: **BRCA1 mutations in women attending clinics that evaluate the risk of breast cancer.** *N Engl J Med* 1997, **336**:1409-1415.
 13. Hopper JL, Jenkins MA: **Modeling the probability that Ashkenazi Jewish women carry a founder mutation in BRCA1 or BRCA2.** *Am J Hum Genet* 1999, **65**:1771-1775.
 14. Hodgson SV, Heap E, Cameron J, Ellis D, Mathew CG, Eeles RA, Solomon E, Lewis CM: **Risk factors for detecting germline BRCA1 and BRCA2 founder mutations in Ashkenazi Jewish women with breast or ovarian cancer.** *J Med Genet* 1999, **36**:369-373.
 15. Martin AM, Blackwood MA, Antin-Ozerkis D, Shih HA, Calzone K, Colligon TA, Seal S, Collins N, Stratton MR, Weber BL, Nathanson KL: **Germline mutations in BRCA1 and BRCA2 in breast-ovarian families from a breast cancer risk evaluation clinic.** *J Clin Oncol* 2001, **19**:2247-2253.
 16. Foulkes WD, Brunet JS, Warner E, Goodwin PJ, Meschino W, Narod SA, Goss PE, Glendon G: **The importance of a family history of breast cancer in predicting the presence of a BRCA mutation.** *Am J Hum Genet* 1999, **65**:1776-1779.
 17. Hopper JL, Chenevix-Trench G, Jolley DJ, Dite GS, Jenkins MA, Venter DJ, McCredie MR, Giles GG: **Design and analysis issues in a population-based, case-control-family study of the genetic epidemiology of breast cancer and the Co-operative Family Registry for Breast Cancer Studies (CFRBCS).** *J Natl Cancer Inst Monogr* 1999, **26**:95-100.
 18. Southey MC, Tesoriero AA, Anderson CR, Jennings KM, Brown SM, Dite GS, Jenkins MA, Osborne RH, Maskiell JA, Porter L, Giles GG, McCredie MR, Hopper JL, Venter DJ: **BRCA1 mutations and other sequence variants in a population-based sample of Australian women with breast cancer.** *Br J Cancer* 1999, **79**:34-39.
 19. Baker RJ, Nelder JA: *The GLIM System, Release 3: Generalized Linear Interactive Modelling.* Oxford: Numerical Algorithms Group; 1978.
 20. Cox DR, Snell EJ: *Analysis of Binary Data. 2nd edition.* London: Chapman & Hall; 1989:41-43.
 21. Wacholder S, Hartge P: **Response to the letters from Hopper and Jenkins and Foulkes et al.** *Am J Hum Genet* 1999, **65**:1775-1776.
 22. Parmigiani G, Berry D, Aguilar O: **Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2.** *Am J Hum Genet* 1998, **62**:145-158.
 23. Peto J, Collins N, Barfoot R, Warren W, Rahman N, Easton DF, Evans C, Deacon J, Stratton MR: **Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer.** *J Natl Cancer Inst* 1999, **91**:943-949.
 24. Antoniou AC, Pharoah PDP, McMullan G, Day NE, Stratton MR, Peto J, Ponder BJ, Easton DFA: **Comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes.** *Br J Cancer* 2002, **86**:76-83.
 25. Cui J, Antoniou AC, Dite GS, Southey MC, Venter DJ, Easton DF, Giles GG, McCredie MRE, Hopper JL: **After BRCA1 and BRCA2 – what next? Multifactorial segregation analyses of three-generation, population-based Australian families affected by female breast cancer.** *Am J Hum Genet* 2001, **68**:420-431.
 26. Armes JE, Trute L, White D, Southey MC, Hammet F, Tesoriero A, Hutchins AM, Dite GS, McCredie MR, Giles GG, Hopper JL, Venter DJ: **Distinct molecular pathogenesis of early-onset breast cancers in BRCA1 and BRCA2 mutation carriers: a population-based study.** *Cancer Res* 1999, **59**:2011-2017.

Correspondence

John L Hopper, The University of Melbourne, Centre for Genetic Epidemiology, 723 Swanston Street, Carlton, Victoria 3053, Australia. Tel: +61 3 8344 0697; fax: +61 3 9347 6929; e-mail: j.hopper@unimelb.edu.au