

## Reliability of self-reported diagnostic radiation history in BRCA1/2 mutation carriers

Anouk Pijpe · Peggy Manders · Renée L. Mulder ·  
HEBON · Flora E. van Leeuwen · Matti A. Rookus

Received: 9 July 2009 / Accepted: 4 December 2009 / Published online: 12 January 2010  
© Springer Science+Business Media B.V. 2010

**Abstract** We assessed reliability of self-reported diagnostic radiation history in BRCA1/2 mutation carriers with and without breast cancer. Within the framework of the HEBON study, 401 BRCA1/2 mutation carriers completed a baseline (1999–2004) and a follow-up questionnaire (2006–2007). Test–retest reliability of self-reported exposure to chest X-rays, fluoroscopies and mammograms was assessed for the entire study population and by case status. Overall proportion agreement on reporting ever/never exposure was good ( $> 75\%$ ), while the corresponding kappa coefficients were between 0.40 and 0.75, indicating at least moderate reliability beyond chance. Reliability of number of exposures was also good ( $> 75\%$ ). Proportion agreement on reporting age at first mammogram was low (40%) for exact consistency and moderate (60%) for consistency  $\pm 1$  year. Reliability of age at first mammogram was higher for cases than for unaffected carriers ( $P < 0.001$ ) but this difference disappeared when excluding diagnostic mammograms ( $P = 0.60$ ). In unaffected carriers proportion agreement on age at last mammogram was 50%. In general, the direction of disagreement on all items was equally distributed. More consistent reporting was mainly determined by a younger age at questionnaire

completion. In conclusion, inconsistent self-report of diagnostic radiation by BRCA1/2 mutation carriers was mainly non-differential by disease status.

**Keywords** Reliability · Questionnaire · Diagnostic imaging · BRCA · Breast cancer

### Abbreviations

BPM	Bilateral prophylactic mastectomy
BPO	Bilateral prophylactic (salpingo-) oophorectomy
Carriers	BRCA1/2 mutation carriers
CI	Confidence interval
HEBON	Hereditary Breast and Ovarian cancer study the Netherlands
MRI	Magnetic resonance imaging
NCR	Netherlands Cancer Registry
OR	Odds ratio
PALGA	Netherlands Pathology Database
TB	Tuberculosis

A. Pijpe · P. Manders · R. L. Mulder · F. E. van Leeuwen ·  
M. A. Rookus (✉)  
Department of Epidemiology, Netherlands Cancer Institute,  
Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands  
e-mail: m.rookus@nki.nl

P. Manders  
Department of Human Genetics, Radboud University Nijmegen  
Medical Center, Nijmegen, The Netherlands

R. L. Mulder  
Department of Pediatric Oncology, Emma Children's Hospital/  
Academic Medical Center, Amsterdam, The Netherlands

### Introduction

Exposure to low dose ionizing (or: diagnostic) radiation may increase the risk of breast cancer in the general population, especially when exposure occurred at young ages [1, 2]. Because BRCA mutations are associated with a reduction in DNA repair efficiency [3], it has been hypothesized that BRCA1/2 mutation carriers might experience greater risk of radiation-induced breast cancer. In two recent studies, exposure to chest X-rays, especially at young ages, was associated with an increased breast

cancer risk among carriers [4, 5], but in two other studies on exposure to mammograms no association was observed [6, 7]. However, these results must be interpreted with caution since the studies relied on self-reported diagnostic radiation history and had a retrospective design (with potential recall and/or survival bias). In the Netherlands, carriers are currently recommended to start breast cancer screening by MRI from age 25 and mammography from age 30 onwards [8], but in the past carriers may have started mammographic screening at younger ages [6, 7].

Both retrospective and prospective studies rely on self-reports rather than medical record review because of the time and cost constraints of medical record abstraction as well as the difficulties in assessing diagnostic radiation history from medical records, especially for exposures that occurred during childhood or in the distant past [9]. It is important to investigate the accuracy of self-reported diagnostic radiation histories. In retrospective studies inaccuracy may lead to non-differential and/or differential misclassification which would affect the risk estimates in an epidemiological study.

A number of studies on reliability [10–16] and validity [9, 17–19] of self-reported diagnostic radiation exposure in the general population have been conducted. Test–retest reliability varies within the measure of self-reported mammography [10–16] and consistency of reporting lifetime number of mammograms decreases with increasing number of exposures [11, 12]. Differences in accuracy of self-reported diagnostic radiation history between affected and unaffected women were so far only investigated in a few validation studies [17, 18, 20]. Although these studies showed a certain amount of disagreement between self-reports and medical records, there were no differences in disagreement between cases and controls (non-differential misclassification).

This is the first study investigating reliability of self-reported diagnostic radiation exposure at young ages in

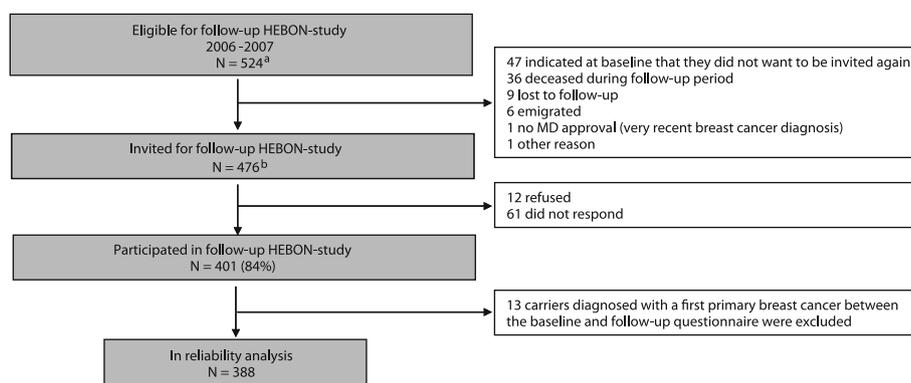
BRCA1/2 mutation carriers. Based on their positive family history, associated cancer screens, and DNA testing in a clinical setting, these women are probably more health conscious than the general population, regardless of whether they had breast cancer or not. Furthermore, because of their high-risk status, their exposure to diagnostic radiation is relatively high since they are screened more intensively from a relatively young age onwards.

## Materials and methods

### Study population

This study on reliability of self-reported diagnostic radiation history among BRCA1/2 mutation carriers was conducted within the framework of the HEBON study, of which the design was described earlier [21]. In brief, the HEBON study is an ongoing nationwide retrospective cohort study with prospective follow-up among members of BRCA1/2 families in the Netherlands. Female family members were eligible if they (a) carried a BRCA1 or BRCA2 mutation; (b) were alive and had no personal history of breast and/or ovarian cancer on January 1, 1960, or born after 1960; and (c) were at age 18 or older at study entry. Figure 1 depicts the identification of the study population for the present study. Between 1999 and 2004, 524 BRCA1/2 mutation carriers completed a baseline risk factor questionnaire and were eligible for a follow-up questionnaire in 2006–2007. Hundred carriers could not be invited for the follow-up questionnaire, mainly because they had died during follow-up period (7%) or indicated at baseline that they did not want to be invited again (9%). Of the 476 carriers (45% affected with breast cancer) invited for the follow-up study, 401 responded (84%). Thirteen carriers were diagnosed with a primary invasive breast cancer in the period between baseline and follow-up. Since

**Fig. 1** Identification of the study population ( $N = 401$ )



<sup>a</sup> Carriers registered at 7/8 Clinical Genetic Centers

<sup>b</sup> Includes 52 carriers from 1 center invited through other Clinical Genetic Centers

this event might have influenced reporting in the follow-up questionnaire, and the number of incident cases was too small for a subgroup analysis, we excluded them from the analysis. Thus, information from 388 carriers (167 prevalent breast cancer cases) was used for the reliability analysis.

#### Data collection

In the baseline questionnaire, assessment of diagnostic radiation exposure was based on the following questions: ever/never having had a chest X-ray before age 20, fluoroscopy for TB, and/or mammogram. For fluoroscopies and mammograms the age at first exposure was asked (and for mammograms also the age at last exposure). Number of exposures was reported in categories (fluoroscopy for TB and chest X-rays before age 20: 1–3, 4–6, 7–9, 10 or more; mammograms before age 30: 1–2, 3–4, more than 4; lifetime number of mammograms: 1–4, 5–10, more than 10).

The follow-up questionnaire contained indication-based questions on lifetime diagnostic and therapeutic radiation exposure (fluoroscopies, chest X-rays, CT-scans, mammograms, radiotherapy, and other radiographic procedures using ionizing radiation like DXA scans). For fluoroscopies, chest X-rays, and mammograms the age at first and last exposure were reported, as well as the exact number of exposures before age 20, between age 20 and 29, and between age 30 and 39. For mammograms, the lifetime number was also asked in categories (1–3, 4–6, 7–9, 10 or more).

Informed consent was obtained from each subject. Information on cancer history and prophylactic surgeries was either collected through PALGA [22] and NCR until August 2007 and/or self-reported for the period not covered by the registries (<1989). Ninety-six percent of breast cancers diagnosed after 1988 were confirmed by PALGA/NCR.

#### Statistical analysis

Reliability was assessed by proportion agreement and Cohen's kappa coefficient ( $\kappa$ ), which accounts for the effect of chance agreement. Reliability was considered poor if  $\kappa \leq 0.40$ , moderate if  $0.40 \geq \kappa \leq 0.75$ , and excellent if  $\kappa > 0.75$  [23]. We also investigated the direction of disagreement, i.e. under- and/or overreporting of exposure in one questionnaire versus the other.

Self-reported exposure to chest X-rays before age 20, fluoroscopies before age 30, mammograms before age 30, and mammograms lifetime was compared between the baseline and the follow-up questionnaire in the entire study population and in prevalent cases and unaffected carriers separately. First, reliability of ever/never exposure was assessed. Then, among women who reported to have been exposed in both questionnaires, reliability of age at first exposure (fluoroscopies before age 30, mammograms

before age 30, and mammograms lifetime) and number of exposures (chest X-rays before age 20, fluoroscopies before age 30, and mammograms before age 30) was assessed. In the baseline questionnaire, number of exposures was reported in categories, while in the follow-up questionnaire the exact number was reported. We examined if the exact number fitted the indicated category. For reliability analysis of mammogram exposure, we excluded 13 carriers who had received their first mammogram after the baseline questionnaire.

In affected carriers, lifetime exposure to diagnostic radiation was for a large part due to their breast cancer diagnosis. Such exposures are not taken into account in an epidemiological study of diagnostic radiation exposure and breast cancer risk. Therefore, we also assessed reliability of prediagnostic mammograms by excluding all mammograms that occurred in the year before breast cancer diagnosis and thereafter. For number of mammograms before age 40 and age at last mammogram, the pre- and postdiagnostic period could not be distinguished. Therefore, reliability of these measures was assessed in unaffected carriers only. For the reliability analysis of age at last mammogram, women who had their last mammogram (as reported in follow-up questionnaire) after baseline questionnaire completion were excluded.

Multivariate logistic regression was used to find predictors of inconsistent self-reporting of diagnostic radiation exposures. The dependent variable in this model was disagreement (no/yes) with agreement as the reference category. The following potential determinants were examined: case status, age at questionnaire completion, number of years between baseline and follow-up questionnaire completion, educational level, nulliparity, having had prophylactic surgery (RRM and/or BPO), and menopausal status. For disagreement on the age at first mammogram we also examined the reason for the first mammogram, lifetime number of mammograms, and length of recall (age at questionnaire minus age at first mammogram).

Differences in proportion agreement and other discrete variables between prevalent cases and unaffected carriers were examined by Pearson's  $\chi^2$  test. Differences in continuous variables were examined by Student's *t*-test. Two-sided *P*-values  $\leq 0.05$  were considered statistically significant. Missing values were excluded from reliability analysis. All analyses were performed using STATA/SE 10.0 (StataCorp LP).

#### Results

The general characteristics of the study population are summarized in Table 1. The age (mean  $\pm$  standard deviation) at baseline questionnaire completion was  $43.9 \pm 12.2$  years.

**Table 1** Characteristics of the study population ( $N = 388$ )

Characteristic	Total ( $N = 388$ )		Prevalent cases ( $N = 167$ )		Unaffected ( $N = 221$ )	
	$N$	%	$N$	%	$N$	%
Nulliparous	100	26	38	23	62	28
Postmenopausal						
At baseline questionnaire	248	64	140	84	108	49
At follow-up questionnaire	301	78	155	93	146	66
Ever BPM	203	52	89 <sup>a</sup>	53	114	52
Ever BPO	260	67	130 <sup>b</sup>	78	130	59
Educational level <sup>c</sup>						
Low	157	41	83	50	74	34
Medium	133	34	44	27	89	40
High	97	25	39	23	58	26

<sup>a</sup> All after breast cancer diagnosis

<sup>b</sup> 5/130 Before breast cancer diagnosis

<sup>c</sup> Low, primary school; medium, secondary school; high, college or university

Prevalent cases were older at questionnaire completion than unaffected carriers ( $48.5 \pm 10.4$  and  $40.5 \pm 12.4$  years, respectively;  $P < 0.001$ ) and they had a lower educational level ( $P < 0.05$ ). Mean time between questionnaires was  $5.4 \pm 0.8$  years, and there was no difference between cases and unaffected carriers ( $P = 0.731$ ). The current study population did not differ substantially from the eligible group in general characteristics like age, parity, oral contraceptive use, and menopausal status (data not shown).

Tables 2 and 3 show the results of the reliability analysis on reporting chest X-rays before age 20, and fluoroscopies and mammograms before age 30. Overall proportions agreement on ever/never chest X-rays exposure before age 20 and fluoroscopies and mammograms before age 30 were 79, 77 and 90%, respectively. Disagreement tended to involve a change from 'ever' in the baseline questionnaire to 'never' in the follow-up questionnaire rather than vice versa for chest X-rays and fluoroscopies, but not for mammograms. Prevalent cases had a significantly lower agreement (of  $\sim 10\%$ ) for ever/never fluoroscopy for TB before age 30 than unaffected carriers ( $P = 0.043$ ), while there were no differences in the direction of disagreement ( $P = 0.512$ ). For all items on ever/never exposure the kappa was moderate, except for ever/never chest X-rays before age 20 in unaffected carriers where it was poor (0.34). Agreement on age at first fluoroscopy exposure was poor (43%) and kappa was just below moderate (0.39), and this was similar for prevalent cases and unaffected carriers ( $P = 0.637$ ). The mean age at first fluoroscopy was  $16.4 \pm 5.3$  years in the baseline and  $16.0 \pm 5.5$  years in the follow-up questionnaire, respectively; there were no differences between prevalent cases and unaffected carriers ( $P = 0.444$  and  $P = 0.771$  for baseline and follow-up questionnaire comparison, respectively). Reliability of age

at first mammogram before age 30 was lower for unaffected carriers than for prevalent cases ( $P = 0.030$ ), but for cases this analysis included diagnostic mammograms. For consistency within 1 year, agreement improved to 81% and the difference in agreement between prevalent cases and unaffected disappeared ( $P = 0.375$ ). For both age at first fluoroscopy and age at first mammogram before age 30, the direction of disagreement was equally distributed and not different between prevalent cases and unaffected carriers ( $P = 0.340$  and  $P = 0.650$ , respectively). Agreement on number of exposures for all three diagnostic procedures was good ( $>70\%$ ) and no differences between prevalent cases and unaffected carriers were found. However, for number of chest X-rays and mammograms kappa was poor (0.26 and 0.35), and for number of fluoroscopies kappa was moderate (0.46). For number of fluoroscopies and mammograms before age 30 kappa was higher in prevalent cases than in unaffected carriers. For all items on number of exposures, the direction of disagreement was equally distributed and not different between prevalent cases and unaffected carriers. The size of disagreement on number of exposures was generally small, e.g. for all women who disagreed on number of mammograms before age 30, the difference in number was between 1 and 4 exposures (Table 4).

Reliability of lifetime exposure to mammograms was assessed for all mammograms lifetime (including diagnostic) and for prediagnostic mammograms only, separately (Table 5). Reliability of ever/never mammograms was excellent, however, the categories were hardly discriminant. Among women who reported to have ever had a mammogram in both questionnaires, the mean age at first mammogram was  $35.4 \pm 9.3$  years in the baseline and  $36.0 \pm 9.6$  years in the follow-up questionnaire, respectively (data

**Table 2** Comparison of self-reported exposure to chest X-rays before age 20 and fluoroscopies before age 30 between baseline and follow-up questionnaire for the entire study population ( $N = 388$ ) and for prevalent cases ( $N = 167$ ) and unaffected carriers ( $N = 221$ ) separately

	Total ( $N = 388$ )		Prevalent cases ( $N = 167$ )		Unaffected ( $N = 221$ )		$P^*$
	$N$	%	$N$	%	$N$	%	
<i>Chest X-rays &lt;20</i>							
Ever/never (baseline/follow-up)							
Never/never	250	67	101	64	149	70	0.855
Never/ever	27	7	10	6	17	8	
Ever/never	50	13	22	14	28	13	
Ever/ever	45	12	25	16	20	9	
Agreement (%)	79.3		79.8		79.0		
Kappa	0.41		0.48		0.34		
No. of exposures <sup>a</sup> (baseline/follow-up)							
1–3/1–3	28	65	15	65	13	65	0.692
1–3/>3	6	14	3	13	3	15	
>3/1–3	5	12	3	13	2	10	
>3/>3	4	9	2	9	2	10	
Agreement (%)	74.4		73.9		75.0		
Kappa	0.26		0.23		0.29		
<i>Fluoroscopy for TB &lt;30</i>							
Ever/never (baseline/follow-up)							
Never/never	185	48	58	38	127	61	0.043
Never/ever	12	4	8	5	4	2	
Ever/never	71	20	35	23	36	17	
Ever/ever	90	28	50	33	40	19	
Agreement (%)	76.8		71.5		80.7		
Kappa	0.52		0.45		0.54		
Age at first <sup>a</sup>							
Equal	23	43	13	46	10	40	0.637
Baseline < follow-up	13	25	8	29	5	20	
Baseline > follow-up	17	32	7	25	10	40	
Agreement (%)	43.4		46.4		40.0		
Kappa	0.39		0.42		0.35		
No. of exposures <sup>a,b</sup> (baseline/follow-up)							
1–3/1–3	32	46	14	36	18	60	0.907
1–3/>3	9	13	4	10	5	17	
>3/1–3	9	13	6	15	3	10	
>3/>3	19	28	15	38	4	13	
Agreement (%)	73.9		74.4		73.3		
Kappa	0.46		0.49		0.32		

\*  $P$ -value of Pearson's  $\chi^2$  test difference in proportion agreement between prevalent cases and unaffected carriers

<sup>a</sup> Reliability analysis in women who reported to ever have been exposed in both baseline and follow-up questionnaire. The numbers in the categories of age at first exposure and number of exposures do not always add up to 100% of ever/ever group due to missing values

<sup>b</sup> The number of exposures before age 30 reported in baseline questionnaire was based on the reported lifetime number of exposures, assuming that the majority had occurred before the age of 30

not shown). Reliability of age at first mammogram lifetime was significantly poorer for unaffected carriers than for prevalent cases ( $P < 0.001$ ). However, when diagnostic mammograms were excluded, reliability of age at first prediagnostic mammogram was poor in prevalent cases as

well, and no longer significantly different from unaffected carriers (agreement 39% and kappa 0.36,  $P = 0.599$ ). The difference in age at first prediagnostic mammogram (age reported in follow-up questionnaire minus age reported in baseline questionnaire) for all women ranged from  $-11$  to

**Table 3** Comparison of self-reported *mammograms before age 30* between baseline and follow-up questionnaire for the entire study population ( $N = 388$ ) and for prevalent cases ( $N = 167$ ) and unaffected carriers ( $N = 221$ ) separately

	Total ( $N = 388$ )		Prevalent cases ( $N = 167$ )		Unaffected ( $N = 221$ )		$P^*$
	$N$	%	$N$	%	$N$	%	
<i>All mammograms before age 30</i>							
Ever/never (baseline/follow-up)							
Never/never	251	67	132	80	119	59	0.368
Never/ever	13	4	7	4	6	3	
Ever/never	24	7	7	4	17	8	
Ever/ever	78	21	18 <sup>a</sup>	11	60	30	
Agreement (%)	89.9		91.5		88.6		
Kappa	0.74		0.67		0.75		
Age at first <sup>b</sup>							
Equal	39	53	13	76	26	46	0.030
Baseline < follow-up	13	18	3	18	10	18	
Baseline > follow-up	21	29	1	6	20	36	
Agreement (%)	53.4		76.5		46.4		
Kappa	0.48		0.71		0.40		
Agreement within 1 year (%)	80.8		88.2		78.6		0.375
No. of exposures—ever/ever group <sup>b,c</sup> (baseline/follow-up)							
1–2/1–2	9	20	6	38	3	11	0.617
1–2/>2	6	14	2	13	4	14	
>2/1–2	7	16	2	13	5	18	
>2/>2	22	50	6	38	16	57	
Agreement (%)	70.5		75.0		67.9		
Kappa	0.35		0.50		0.18		

\*  $P$ -value of Pearson's  $\chi^2$  test difference in proportion agreement between prevalent cases and unaffected carriers

<sup>a</sup> Includes 13 women with breast cancer diagnosis before age 30

<sup>b</sup> Reliability analysis in women who reported to ever had a mammogram before age 30 in baseline and follow-up questionnaire. The numbers in the categories of age at first mammogram and number of mammograms do not always add up to 100% of ever/ever group due to missing values

<sup>c</sup> Women who were < 30 years at baseline questionnaire completion were excluded ( $N = 35$ )

**Table 4** Comparison of self-reported number of mammograms before age 30 between baseline and follow-up questionnaire ( $N = 49$ )

	Follow-up questionnaire											Total	
	1	2	3	4	5	6	7	8	9	10	>10		Missing
Baseline questionnaire													
1–2	<b>6</b>	<b>3</b>	3	1	2	0	0	0	0	0	0	3	18
3–4	2	3	<b>2</b>	<b>6</b>	2	2	0	1	0	0	0	1	19
>4	0	2	3	0	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>	<b>1</b>	<b>0</b>	<b>3</b>	0	11
Missing	0	0	1	0	0	0	0	0	0	0	0	0	1
Total	8	8	9	7	4	3	0	2	1	0	3	4	49

Bold values represent carriers who reported the same number of mammograms in the baseline and the follow-up questionnaire

+24 years (median –1 year; mean down  $2.6 \pm 2.1$  years; mean up  $4.2 \pm 4.7$  years). In 53/147 (36%) carriers who reported an inconsistent age at first prediagnostic mammogram, the difference in age was only 1 year (43% of carriers differed >2 years and 14% >5 years). For consistency

within 1 year, reliability of age at first prediagnostic mammogram increased to moderate (59%) for both prevalent cases and unaffected carriers ( $P = 0.767$ ), and for all mammograms reliability increased to moderate for unaffected carriers and good for prevalent cases ( $P = 0.004$ ). The direction of disagreement on age at first mammogram lifetime was equally distributed, although prevalent cases reported more often a younger age at first (prediagnostic) mammogram in the follow-up questionnaire than in the baseline questionnaire; however, this was not significantly different from unaffected carriers ( $P = 0.326$ ).

Among unaffected carriers (data not shown), agreement on number of mammograms before age 40 was good (81%) and kappa was moderate (0.57). Agreement on age at last mammogram before baseline was 50% and kappa was moderate (0.48). For both items, there were no differences in the direction of disagreement.

We examined determinants of inconsistent reporting of age at first prediagnostic mammogram lifetime

**Table 5** Comparison of self-reported *mammograms lifetime* between baseline and follow-up questionnaire for the entire study population ( $N = 388$ ) and for prevalent cases ( $N = 167$ ) and unaffected carriers ( $N = 221$ ) separately

	Total ( $N = 388$ )		Prevalent cases ( $N = 167$ )		Unaffected ( $N = 221$ )		$P^*$
	$N$	%	$N$	%	$N$	%	
<b>All mammograms</b>							
Ever/never (baseline/follow-up)							
Never/never	6	2	0	0	6	3	0.218
Never/ever	2	<1	2	1	0	0	
Ever/never	2	<1	1	<1	1	<1	
Ever/ever	365	97	164	98	201	97	
Agreement (%)	98.9		98.2		99.5		
Kappa	0.74		–		0.92		
Age at first <sup>a</sup>							
<i>Exact agreement</i>							
Equal	148	44	84	55	64	35	<0.001
Baseline < follow-up	91	27	32	21	59	32	
Baseline > follow-up	96	29	36	24	60	33	
Agreement (%)	44.2		55.3		35.0		
Kappa	0.42		0.53		0.33		
Age at first <sup>a</sup>							
<i>Agreement within 1 year</i>							
Equal	219	65	112	74	107	58	0.004
Baseline < follow-up	63	19	22	14	41	22	
Baseline > follow-up	53	16	18	12	35	19	
Agreement (%)	65.4		73.7		58.5		
Kappa	0.64		0.73		0.57		
<b>Prediagnostic mammograms</b>							
Ever/never (baseline/follow-up)							
Never/never	95	26	89	57	6	3	<0.001
Never/ever	6	2	6	4	0	0	
Ever/never	15	4	14	9	1	<1	
Ever/ever	247	68	46	30	201	96	
Agreement (%)	94.2		87.1		99.5		
Kappa	0.86		0.72		0.92		
Age at first <sup>a</sup>							
<i>Exact agreement</i>							
Equal	82	36	18	39	64	35	0.599
Baseline < follow-up	70	31	11	24	59	32	
Baseline > follow-up	77	33	17	37	60	33	
Agreement (%)	35.8		39.1		35.0		
Kappa	0.34		0.36		0.33		
Age at first <sup>a</sup>							
<i>Agreement within 1 year</i>							
Equal	135	59	28	61	107	58	0.767
Baseline < follow-up	48	21	7	15	41	22	
Baseline > follow-up	46	20	11	24	35	19	
Agreement (%)	59.0		60.9		58.5		
Kappa	0.58		0.59		0.57		

\*  $P$ -value of Pearson's  $\chi^2$  test difference in proportion agreement between prevalent cases and unaffected carriers

<sup>a</sup> Reliability analysis in women who reported to ever had a mammogram before age 30 in baseline and follow-up questionnaire. The numbers in the categories of age at first mammograms do not always add up to 100% of ever/ever group due to missing values

**Table 6** Estimated Odds Ratios (95% CI) for disagreement on age at first mammogram lifetime and ever/never exposure to fluoroscopies before age 30

Potential determinants of disagreement	Age at 1st mammogram lifetime <sup>a</sup> N = 229		Ever/never exposure to fluoroscopies before age 30 N = 358	
	Univariate OR (95% CI)	Adjusted OR <sup>b</sup> (95% CI)	Univariate OR (95% CI)	Adjusted OR <sup>b</sup> (95% CI)
Case status (unaffected vs. prevalent case)	1.10 (0.57–2.14)	1.44 (0.67–3.11)	0.60 (0.37–0.99)	0.52 (0.26–1.05)
Age at follow-up questionnaire (per year)	1.04 (1.01–1.06)	1.03 (0.98–1.08)	1.02 (1.00–1.04)	1.01 (0.97–1.04)
Time between questionnaires (per year)	0.70 (0.50–0.90)	0.76 (0.53–1.10)	0.92 (0.67–1.26)	0.95 (0.70–1.31)
Educational level (high vs. low)	0.88 (0.45–1.72)	1.16 (0.55–2.44)	0.96 (0.70–1.30)	1.04 (0.75–1.43)
Nulliparous (yes vs. no)	0.71 (0.37–1.37)	1.04 (0.50–2.19)	0.70 (0.39–1.25)	0.77 (0.40–1.48)
BPO (ever vs. never)	1.35 (0.76–2.40)	1.55 (0.75–3.21)	1.08 (0.64–1.82)	0.78 (0.42–1.46)
RRM (ever vs. never)	0.83 (0.49–1.41)	0.72 (0.35–1.50)	0.89 (0.52–1.53)	1.48 (0.71–3.06)
Menopausal status (post vs. pre)	1.73 (1.00–2.98)	0.82 (0.33–2.04)	1.57 (0.93–2.66)	1.23 (0.55–2.75)
Reason 1st mammogram <sup>c</sup> (complaints vs. screening)	0.58 (0.32–1.06)	0.70 (0.35–1.42)	Not applicable	Not applicable
Lifetime no. of mammograms <sup>c</sup> (>7 vs. 1–6)	0.96 (0.56–1.64)	0.64 (0.30–1.38)	Not applicable	Not applicable
Length of recall <sup>d</sup> (per year)	1.01 (0.97–1.05)	0.99 (0.94–1.05)	Not applicable	Not applicable

<sup>a</sup> Prediagnostic mammograms; agreement within 1 year

<sup>b</sup> Obtained from multivariate logistic regression model, adjusted for all applicable variables listed in table

<sup>c</sup> As reported in the follow-up questionnaire

<sup>d</sup> Length of recall: age at follow-up questionnaire completion minus age at 1st mammogram reported in follow-up questionnaire

(consistency within 1 year) and ever/never exposure to fluoroscopies before age 30 (Table 6). In univariate analysis, the chance of disagreement on age at first mammogram was 4% higher per additional year of age at questionnaire completion (OR 1.04, 95% CI 1.01–1.06), higher for postmenopausal women (OR 1.71, 95% CI 1.00–2.98) and lower when the reason for the first mammogram was having complaints (OR 0.58, 95% CI 0.32–1.06). Unexpectedly, increased time between questionnaires was associated with lower chance of disagreement (OR 0.70, 95% CI 0.50–0.90). Disagreement on ever/never having had a fluoroscopy before age 30 was determined by case status and age at questionnaire completion. However, in the multivariate model of both items, there were no significant associations between any of the potential predictors and the chance of disagreement, although age at questionnaire completion remained a marginally significant determinant of disagreement.

## Discussion

To our knowledge, this is the first study on reliability of self-reported diagnostic radiation exposure history in BRCA1/2 carriers. Proportion agreement on reporting ever/never exposure was good (>75%), while the corresponding kappa coefficients were between 0.40 and 0.75, indicating at least moderate reliability beyond chance. Reliability of number of exposures was also good (>75%). Reliability of

reporting age at first mammogram was low (40%) for exact consistency and moderate (60%) for consistency within 1 year. Reliability of age at first mammogram was higher for cases than for unaffected carriers ( $P < 0.001$ ) but this difference disappeared when excluding diagnostic mammograms ( $P = 0.599$ ). Reliability of exact reporting of age at last mammogram was 50% in unaffected carriers. In general, the direction of disagreement on all items was equally distributed, i.e. there was as much underreporting as overreporting of exposure in one questionnaire versus the other. However, for chest X-rays before age 20 and fluoroscopies before age 30, disagreement tended to involve a change from 'ever' in the baseline questionnaire to 'never' in the follow-up questionnaire for both cases and unaffected carriers. Being an unaffected carrier and being younger at questionnaire completion were associated with more consistent reporting of ever/never exposure to fluoroscopy for tuberculosis before age 30. For all other measures, agreement was non-differential by disease status. More consistent reporting of age at first mammogram was mainly determined by younger age at questionnaire completion.

Previous reliability studies [10–16] on self-reported diagnostic radiation history were all on mammography in the setting of evaluation of screening programmes. Reliability varied by measure of self-reported mammography: agreement on ever/never having had a mammogram, lifetime number of mammograms, and date of most recent mammogram was approximately 90% [10–16], 60% [11, 12, 16], and 35% [10, 12, 13, 15, 16], respectively. Our

results on reliability of mammograms are in line with these studies, although we observed a somewhat higher proportion agreement on number of mammograms. Since the interval between questionnaires was longer than in the other reliability studies (i.e. 5.4 years vs. 1 week to 2.6 years, respectively), reliability could have been expected to be lower in our study but this appeared not to be the case. Another difference between our study and the other reliability studies is that most of these studies used in-person and/or telephone interviews [10, 11, 13–16] instead of self-administered mailed questionnaires. Personal or telephone interview may enhance memory but may also lead to over-reporting compared to questionnaire methods [19].

It is generally assumed that cases recall past exposures more accurately than controls. Self-report may also be influenced by whether the respondents are selected from the general population, or from a clinical setting, where they may have become more familiar with medical procedures and may remember the exposures better due to cues in a health care setting. We had expected that reliability would be good and similar for prevalent cases and unaffected carriers since our study population of BRCA1/2 carriers was tested in a clinical setting. Reliability of exact age at first mammogram, which was not investigated in previous studies, was therefore rather disappointing. When including diagnostic mammograms for cases, as expected, reliability of reporting age at first mammogram before age 30 and lifetime was higher in prevalent cases than in unaffected carriers ( $P < 0.001$ ). However, the difference in agreement on age at first mammogram before age 30 disappeared for consistency within 1 year ( $P = 0.375$ ). There was no difference between cases and unaffected carriers in reliability of reporting age at first prediagnostic mammogram ( $P = 0.599$  for exact agreement and  $P = 0.767$  for agreement within 1 year). Unaffected carriers, however, reported more often a younger age at first exposure before age 30 in the follow-up questionnaire than cases but this was not statistically significant ( $P = 0.107$  and  $P = 0.269$  for mammogram and fluoroscopy, respectively). In the multivariate models, age at questionnaire completion was the only predictor of disagreement, although this finding was marginally significant. Menopausal status and reason for first mammogram did not predict consistent reporting. In exploratory analysis we found that for a number of women the large difference in age at first mammogram was due to the fact that on one questionnaire women had reported their first *screening* mammography as being the first mammogram ever made, while in the other they had remembered a single mammogram which was made in the more distant past because of complaints (e.g. when they had felt a lump). This was independent of the direction of disagreement. The reported age at start of screening was a separate question in a different part of the follow-up

questionnaire and only completed when a woman underwent screening at the time of questionnaire completion.

In contrast with other reliability studies [11, 12], we found no evidence of an association between number of mammograms and consistent reporting. Reliability of ever/never exposure to chest X-rays before age 20 and fluoroscopies before age 30 was somewhat lower than for mammograms, while reliability of number of exposures was similar. Interestingly, reliability of ever/never exposure to fluoroscopy for tuberculosis before age 30 was significantly higher for unaffected carriers than for prevalent cases and this was not due to the difference in age at questionnaire completion. Since we examined many differences between cases and unaffected carriers, this may be a chance finding. In general, our results cannot be directly compared with those from validity studies [9, 17–19] that compared self-reported information with medical records. Differences between self-report and medical records may be due to both systematic under- or over-reporting and lack of reliability. Consistent underreporting by unaffected carriers in both questionnaires can only be assessed in a validation study. But validation of e.g. chest X-rays, especially when occurred in the distant past, is extremely difficult if not impossible because of the different locations, destroyed records, and the fact that negative self-reports can not be verified. Still, most validation studies suggest that women tend to over-report the number of diagnostic radiation exposures and underreport the time since their last exposure (“telescoping”) [9, 18, 19]. We found no evidence of telescoping on reliability; the direction of disagreement on age at last mammogram was equally distributed (data not shown).

Thirteen incident cases were excluded from the analysis. In this group, we investigated reliability of ever/never chest X-rays before age 20 (data not shown). Proportion agreement and kappa were slightly higher than for the other groups: 83 and 0.56%, respectively.

When interpreting these results, the strengths and limitations of this study should be considered. This is the first study investigating reliability of self-reported risk factor exposure information in BRCA1/2 mutation carriers. Furthermore, we were able to assess reliability of more than one diagnostic radiation exposure type, to assess reliability of age at first exposure, and to investigate differences in reliability between affected and unaffected women. However, our study also has limitations. The most important limitation concerns the differences between the baseline and the follow-up questionnaire. For instance, general versus indication-based questions and categorical answer categories versus exact numbers for number of exposures. We had expected that more exposures would be reported in the indication-based follow-up questionnaire compared to the baseline questionnaire. However, the amount of

disagreement in both directions was similar for most variables, whereas for chest X-rays and fluoroscopies we observed a tendency of a change from 'ever' in the baseline questionnaire to 'never' in the follow-up questionnaire rather than vice versa. Also, for the analysis on number of fluoroscopies before age 30, we based this number in the baseline questionnaire on lifetime number of fluoroscopies because there was no separate question on number of exposures before age 30. Since in the Netherlands screening for tuberculosis by fluoroscopy was mainly done through mass population screening between 1940 and 1960 in relatively young people [24,] we assumed that most of the reported lifetime number of fluoroscopies had taken place before age 30. Equal distribution of the direction of disagreement suggested that this assumption was justified. On the other hand, the differences between both questionnaires allowed us to improve the questionnaire for future studies. For instance, the proportion of missing values in age at first fluoroscopy for tuberculosis before age 30 was 36% in the baseline and 9% in the follow-up questionnaire. Future studies measuring diagnostic radiation by self-report may want to consider using an indication-based questionnaire where for example, for mammogram exposure age at start screening should be asked first, followed by a question on ever having had a mammogram prior to the first screening mammogram. For number of exposures the exact number may be asked instead of categories, but in addition it should be possible to give a range if one is not sure. Another limitation concerns the use of Cohen's kappa coefficient. If a population is homogeneous with respect to the characteristic being studied, like in our study mammographic screening,  $\kappa$  becomes highly sensitive to small departures from perfect concordance, [25] which explains the lower kappa for several of the variables, even in presence of relatively high proportion agreement.

Our findings indicate that consistency of self-reported diagnostic radiation by BRCA1/2 mutation carriers was mainly non-differential by disease status. These results add to knowledge about the reliability of self-reported diagnostic radiation history and its effect on relative risks in studies on the association between diagnostic radiation and breast cancer; risks will likely be underestimated.

**Acknowledgments** This work was supported by the Dutch Cancer Society (grants NKI1998-1854, NKI2004-3088, NKI 2007-3756).

## Appendix

The Netherlands Collaborative Group on Hereditary Breast Cancer (HEBON): Coordinating center: Netherlands Cancer Institute, Amsterdam: Senno Verhoef, Anouk Pijpe,

Richard Brohet, Frans Hogervorst, Laura van 't Veer, Flora van Leeuwen, Matti Rookus; Erasmus Medical Center, Rotterdam: Margriet Colle 'e, Ans van den Ouweland, Mieke Kriege, Mieke Schutte, Maartje Hooning, Caroline Seynaeve; Leiden University Medical Center, Leiden: Christi van Asperen, Juul Wijnen, Peter Devilee; Radboud University Nijmegen Medical Center, Nijmegen: Nicoline Hoogerbrugge, Marjolijn Ligtenberg; University Medical Center Utrecht, Utrecht: Margreet Ausems, Rob van der Luijt; Amsterdam Medical Center: Cora Aalfs, Theo van Os; VU University Medical Center, Amsterdam: Hanne Meijers-Heijboer, Hans Gille; University Hospital Maastricht, Maastricht: Encarna Gomez-Garcia, Rien Blok; University Medical Center Groningen, Groningen: Jan Oosterwijk, Annemiek van der Hout; Netherlands Foundation for Detection of Hereditary Tumours, Leiden: Hans Vasen, Inge van Leeuwen.

## References

1. Ronckers CM, Erdmann CA, Land CE. Radiation and breast cancer: a review of current evidence. *Breast Cancer Res.* 2005;7(1):21–32.
2. Hall EJ, Brenner DJ. Cancer risks from diagnostic radiology. *Br J Radiol.* 2008;81(965):362–78.
3. Boulton SJ. Cellular functions of the BRCA tumour-suppressor proteins. *Biochem Soc Trans.* 2006;34(Pt 5):633–45.
4. Andrieu N, Easton DF, Chang-Claude J, Rookus MA, Brohet R, Cardis E, et al. Effect of chest X-rays on the risk of breast cancer among BRCA1/2 mutation carriers in the international BRCA1/2 carrier cohort study: a report from the EMBRACE, GENEPSO, GEO-HEBON, and IBCCS Collaborators' Group. *J Clin Oncol.* 2006;24(21):3361–6.
5. Gronwald J, Pijpe A, Byrski T, Huzarski T, Stawicka M, Cybulski C, et al. Early radiation exposures and BRCA1-associated breast cancer in young women from Poland. *Breast Cancer Res Treat.* 2008;112(3):581–4.
6. Goldfrank D, Chuai S, Bernstein JL, Ramon YC, Lee JB, Alonso MC, et al. Effect of mammography on breast cancer risk in women with mutations in BRCA1 or BRCA2. *Cancer Epidemiol Biomarkers Prev.* 2006;15(11):2311–3.
7. Narod SA, Lubinski J, Ghadirian P, Lynch HT, Moller P, Foulkes WD, et al. Screening mammography and risk of breast cancer in BRCA1 and BRCA2 mutation carriers: a case-control study. *Lancet Oncol.* 2006;7(5):402–6.
8. National Breast Cancer Organisation the Netherlands. Guideline Breastcancer. The Netherlands: Utrecht; 2008.
9. Paskett ED, Tatum CM, Mack DW, Hoen H, Case LD, Velez R. Validation of self-reported breast and cervical cancer screening tests among low-income minority women. *Cancer Epidemiol Biomarkers Prev.* 1996;5(9):721–6.
10. Bancej CM, Maxwell CJ, Snider J. Inconsistent self-reported mammography history: findings from the National Population Health Survey longitudinal cohort. *BMC Health Serv Res.* 2004;4(1):32.
11. Rauscher GH, O'Malley MS, Earp JA. How consistently do women report lifetime mammograms at successive interviews? *Am J Prev Med.* 2002;22(1):8–14.

12. Rivera S, Vernon SW, Tiro JA, Coan S, Del Junco D, Chan W, et al. Test-retest reliability of self-reported mammography in women veterans. *Prev Med.* 2006;42(4):320–6.
13. Barratt A, Cockburn J, Smith D, Redman S. Reliability and validity of women's recall of mammographic screening. *Aust N Z J Public Health.* 2000;24(1):79–81.
14. Brownson RC, Jackson-Thompson J, Wilkerson JC, Kiani F. Reliability of information on chronic disease risk factors collected in the Missouri Behavioral Risk Factor Surveillance System. *Epidemiology.* 1994;5(5):545–9.
15. Stein AD, Lederman RI, Shea S. Reproducibility of the women's module of the Behavioral Risk Factor Surveillance System questionnaire. *Ann Epidemiol.* 1996;6(1):47–52.
16. Vacek PM, Mickey RM, Worden JK. Reliability of self-reported breast screening information in a survey of lower income women. *Prev Med.* 1997;26(3):287–91.
17. Pogoda JM, Preston-Martin S. Radiation exposure from diagnostic imaging: agreement between self-report and medical records. *Health Phys.* 2002;83(6):907–17.
18. Berrington de Gonzalez A, Ekobom A, Glass AG, Galanti MR, Grimelius L, Allison MJ, et al. Comparison of documented and recalled histories of exposure to diagnostic X-rays in case-control studies of thyroid cancer. *Am J Epidemiol.* 2003;157(7):652–63.
19. Howard M, Agarwal G, Lytwyn A. Accuracy of self-reports of Pap and mammography screening compared to medical record: a meta-analysis. *Cancer Causes Control.* 2009;20(1):1–13.
20. Norman SA, Localio AR, Zhou L, Bernstein L, Coates RJ, Flagg EW, et al. Validation of self-reported screening mammography histories among women with and without breast cancer. *Am J Epidemiol.* 2003;158(3):264–71.
21. Pijpe A, Manders P, Brohet RM, Collee JM, Verhoef S, Vasen HF, et al. Physical activity and the risk of breast cancer in BRCA1/2 mutation carriers. *Breast Cancer Res Treat.* 2009.
22. Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de Pol A, van Krieken JH, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol.* 2007;29(1):19–24.
23. Fleiss JL. *Statistical methods for rates and proportions.* 2nd ed. New York, NY: John Wiley & Sons; 1981.
24. Sickenga FN. Short history of the tuberculosis suppression in the Netherlands 1900–1960 (Korte geschiedenis van de tuberculosebestrijding in Nederlands 1900–1960). The Hague: KNCV Tuberculosisfoundation; 1980.
25. Maclure M, Willett WC. Misinterpretation and misuse of the kappa statistic. *Am J Epidemiol.* 1987;126(2):161–9.