

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/196139>

Please be advised that this information was generated on 2019-12-04 and may be subject to change.



# SNP association study in PMS2-associated Lynch syndrome

Sanne W. ten Broeke<sup>1</sup> · Fadwa A. Elsayed<sup>2</sup> · Lisa Pagan<sup>1</sup> · Maran J. W. Olderode-Berends<sup>3</sup> · Encarna Gomez Garcia<sup>4</sup> · Hans J. P. Gille<sup>5</sup> · Liselot P. van Hest<sup>5</sup> · Tom G. W. Letteboer<sup>6</sup> · Lizet E. van der Kolk<sup>7</sup> · Arjen R. Mensenkamp<sup>8</sup> · Theo A. van Os<sup>9</sup> · Liesbeth Spruijt<sup>8</sup> · Bert J. W. Redeker<sup>9</sup> · Manon Suerink<sup>1</sup> · Yvonne J. Vos<sup>3</sup> · Anja Wagner<sup>10</sup> · Juul T. Wijnen<sup>1</sup> · E. W. Steyerberg<sup>11</sup> · Carli M. J. Tops<sup>1</sup> · Tom van Wezel<sup>2</sup> · Maartje Nielsen<sup>1</sup>

Published online: 17 November 2017

© The Author(s) 2017. This article is an open access publication

## Abstract

Lynch syndrome (LS) patients are at high risk of developing colorectal cancer (CRC). Phenotypic variability might in part be explained by common susceptibility loci identified in Genome Wide Association Studies (GWAS). Previous studies focused mostly on *MLH1*, *MSH2* and *MSH6* carriers, with conflicting results. We aimed to determine the role of GWAS SNPs in *PMS2* mutation carriers. A cohort study was performed in 507 *PMS2* carriers (124 CRC cases), genotyped for 24 GWAS SNPs, including SNPs at 11q23.1 and 8q23.3. Hazard ratios (HRs) were calculated using a weighted Cox regression analysis to correct for ascertainment bias. Discrimination was assessed with a concordance statistic in a bootstrap cross-validation procedure. Individual SNPs only had non-significant associations with CRC occurrence with HRs lower than 2, although male carriers of allele A at rs1321311 (6p21.31) may have increased risk of CRC (HR = 2.1, 95% CI 1.2–3.0). A polygenic risk score (PRS) based on 24 HRs had an HR of 2.6 (95% CI 1.5–4.6) for the highest compared to the lowest quartile, but had no discriminative ability (c statistic 0.52). Previously suggested SNPs do not modify CRC risk in *PMS2* carriers. Future large studies are needed for improved risk stratification among Lynch syndrome patients.

**Keywords** Lynch syndrome · PMS2 · SNP · Cancer risk · Colorectal cancer · Modifiers

## Introduction

Lynch syndrome (LS) accounts for 2–4% of all CRCs and is characterized by a high risk for developing malignancies, most notably colorectal cancer (CRC) and endometrial

cancer (EC). The underlying cause is a germline mutation in one of the mismatch repair (MMR) genes: *MLH1*, *MSH2* (*EPCAM*), *MSH6* or *PMS2*. Mutations in all MMR genes are associated with a significantly increased cancer risk compared to the general population, although *MSH6* and *PMS2* carriers show lower penetrance compared to *MLH1* and *MSH2* carriers [1–3]. Within and between family variability is commonly observed and a range of theories have

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10689-017-0061-3>) contains supplementary material, which is available to authorized users.

✉ Sanne W. ten Broeke  
tenbroeke@lumc.nl

<sup>1</sup> Department of Clinical Genetics, Leiden University Medical Centre, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

<sup>2</sup> Department of Pathology, Leiden University Medical Centre, Leiden, The Netherlands

<sup>3</sup> Department of Clinical Genetics, University Medical Centre Groningen, Groningen, The Netherlands

<sup>4</sup> Department of Clinical Genetics, Maastricht University Medical Centre, Maastricht, The Netherlands

<sup>5</sup> Department of Clinical Genetics, VU Amsterdam, Amsterdam, The Netherlands

<sup>6</sup> Department of Clinical Genetics, University Medical Centre Utrecht, Utrecht, The Netherlands

<sup>7</sup> Department of Clinical Genetics, The Netherlands Cancer Institute, Amsterdam, The Netherlands

<sup>8</sup> Department of Clinical Genetics, Radboud University Medical Centre, Nijmegen, The Netherlands

<sup>9</sup> Department of Clinical Genetics, Academic Medical Centre, Amsterdam, The Netherlands

<sup>10</sup> Department of Clinical Genetics, Erasmus Medical Centre, Rotterdam, The Netherlands

<sup>11</sup> Department of Medical Statistics, Leiden University Medical Centre, Leiden, The Netherlands

been proposed to explain the phenomenon, such as genotype–phenotype correlations, parent-of-origin effects, lifestyle factors and the influence of common susceptibility loci. The latter, mainly single nucleotide polymorphisms (SNPs), were identified in genome wide association studies (GWAS) in large cohorts consisting of sporadic CRC cases [4]. Among these candidate SNPs, previous studies have identified statistically significant effects of multiple SNPs in LS patients, and independent studies replicated the effect of SNPs rs3802842 (11q23.1) and rs16892766 (8q23.3) among *MLH1* carriers [5, 6]. It should be noted, however, that others have failed to replicate these findings [7, 8]. Although the latter studies analyzed cohorts of similar size to our own, few or no *PMS2* carriers were included [7]. Due to a relatively low penetrance and high phenotypic variability, this specific subset of LS patients might be of particular interest [9]. In a previous study among 377 *PMS2* carriers, we found age at CRC diagnosis to vary widely (range 26–86 years) and mean age of index carriers and mutation-positive family members differed by 10 years [3]. In the current study, we aim to determine whether these SNPs modify CRC risk in a large cohort of *PMS2* mutation carriers.

## Materials and methods

### Sample collection

DNA extracted from leucocyte DNA was collected from 8 Dutch family cancer clinics. Index carriers included in this study were sent in between 2007 and 2016 to the Clinical Genetics department, because of a clinical suspicion of LS, e.g. LS-associated cancer at a young age and/or a positive family history. Mutation analysis was initiated based on the presence of histological hallmarks [microsatellite instability (MSI) or loss of *PMS2* expression in the tumor] and/or when the family complied with the Bethesda Criteria [10]. Participating clinics provided DNA samples and clinical data on CRC, age at diagnosis, other cancer development and polypectomy. Controls were defined as carriers that were tested pre-symptomatically, after a pathogenic mutation was identified in the index carrier of the family. All carriers are referred to gastroenterology departments after the diagnosis has been established, which then adhere to international surveillance guidelines, i.e. colonoscopies every 1–2 years from 25 years of age [11]. Data was analyzed anonymously. The study was approved by the medical ethical committee of Leiden University Medical Centre, protocol ID P01-019.

### Genotyping

*PMS2* genotyping in this cohort was carried out as previously described [3]. SNP genotyping was done at the

LUMC laboratory using a KASp genotyping assay (LGC Genomics, Hoddesdon, UK). Primers were designed using Primerpicker (KBioscience, Hoddesdon, UK) and are available upon request. All oligonucleotides were obtained from Eurofins Genomics (Ebersberg, Germany). Genotypes were called using the CFX manager software v3.0 (Bio-Rad, Veenendaal, the Netherlands).

### Statistical analysis

*PMS2* carriers were analyzed as a birth cohort. A Cox-proportional hazards regression model was fitted to estimate hazard ratios (HRs), with age at CRC as endpoint and SNP genotype as independent variable. Patients without CRC were censored at the last age known to be alive. The HR was calculated separately for heterozygous and homozygous carriers of the risk allele, with homozygotes of the non-risk allele as reference category. We also calculated the per allele HR (additive model). Cox-regression analyses were also stratified for gender. These sub-analyses only includes a per allele (additive) model, due to multiple testing risks. Missing age at CRC diagnosis ( $n = 3$ ) was imputed using median age of CRC in the general population (age 70,  $n = 2$ ) or set at 1 year before death ( $n = 1$ ). The proportional hazards assumption was investigated by examining the scaled Schoenfeld residuals with a formal statistical test and by visual inspection.

Previous studies have described the oversampling of cases in clinic-based cohorts. Moreover, affected family members are more likely to be tested for the mutation and this too results in oversampling of cases. To adjust for this non-random sampling, we used a weighted cohort approach as previously described [12]. Standard errors were corrected for familial clustering of risk by using the Huber–White sandwich estimator [13].

We also calculated two polygenic risk scores (PRS) [14] based on (1) the odds ratios (ORs) reported in the meta-analysis by Ma et al. (Supplemental Table 1) and (2) based on our HRs from the current study [4]. Kaplan Meier (KM) and Cox regression analysis were concurrently performed. A concordance statistic was calculated to assess the discriminative value of the 24 SNP model. The optimism in the concordance statistic was estimated by fitting the model in each of 500 bootstrap samples (drawn with replacement), and validation in the original sample. Analyses were initially performed for patients with complete data. We also performed imputation of missing values based on the correlation structure between SNPs and with the outcome (transcan function in R software, version 3.2). Since results were similar, we only present complete case results.

Lastly, a post hoc power analysis was performed to assess the chance of finding significantly increased risks using the collected cohort, which contains all currently known *PMS2*

mutation carriers in the participating centers (Supplemental Fig. 1). We had at least 80% power to find an HR of 1.5 for the majority of SNPs. A more detailed description of the statistical analysis is available in the supplementary methods (Supplemental File 1). Statistical analyses were performed using STATA version 14 (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP) and R software (version 3.2, using the rms library).

## Results

### Cohort

In total, 521 samples from carriers with a germline *PMS2* mutation were genotyped, derived from 152 families. Fourteen patients were excluded because (1) they were younger than 25 years at censoring and were therefore not yet at risk of developing CRC (n = 11) or (2) insufficient clinical data was available (n = 3, including one CRC case). The analyzed cohort consisted of 124 cases (*PMS2* carriers with CRC) and 383 controls (*PMS2* carriers without CRC), with attributed person years of 6527 and 19,549, respectively. Person-years

were calculated until age of CRC for cases (*PMS2* carriers with CRC, n = 125), and age at polypectomy, age of death, or last known age alive (whichever occurred first) for controls (*PMS2* carriers without CRC, n = 1, n = 1 and n = 381 respectively). The mean age was 52.5 for CRC cases and 51.0 for non-cases (Table 1). For a detailed description of the families including genotypes see Supplemental Tables 2a + 2b.

### Hardy–Weinberg equilibrium

Two SNPs, rs1048943 (15q24.1) and rs4925386 (20q13), were not in Hardy Weinberg equilibrium (HWE). Violation of the HWE was present in both cases and controls and as this might be the result of a genotyping error, these SNPs were removed from the analysis. Ultimately, 24 SNPs were included in the final analysis.

### Risk of colorectal cancer

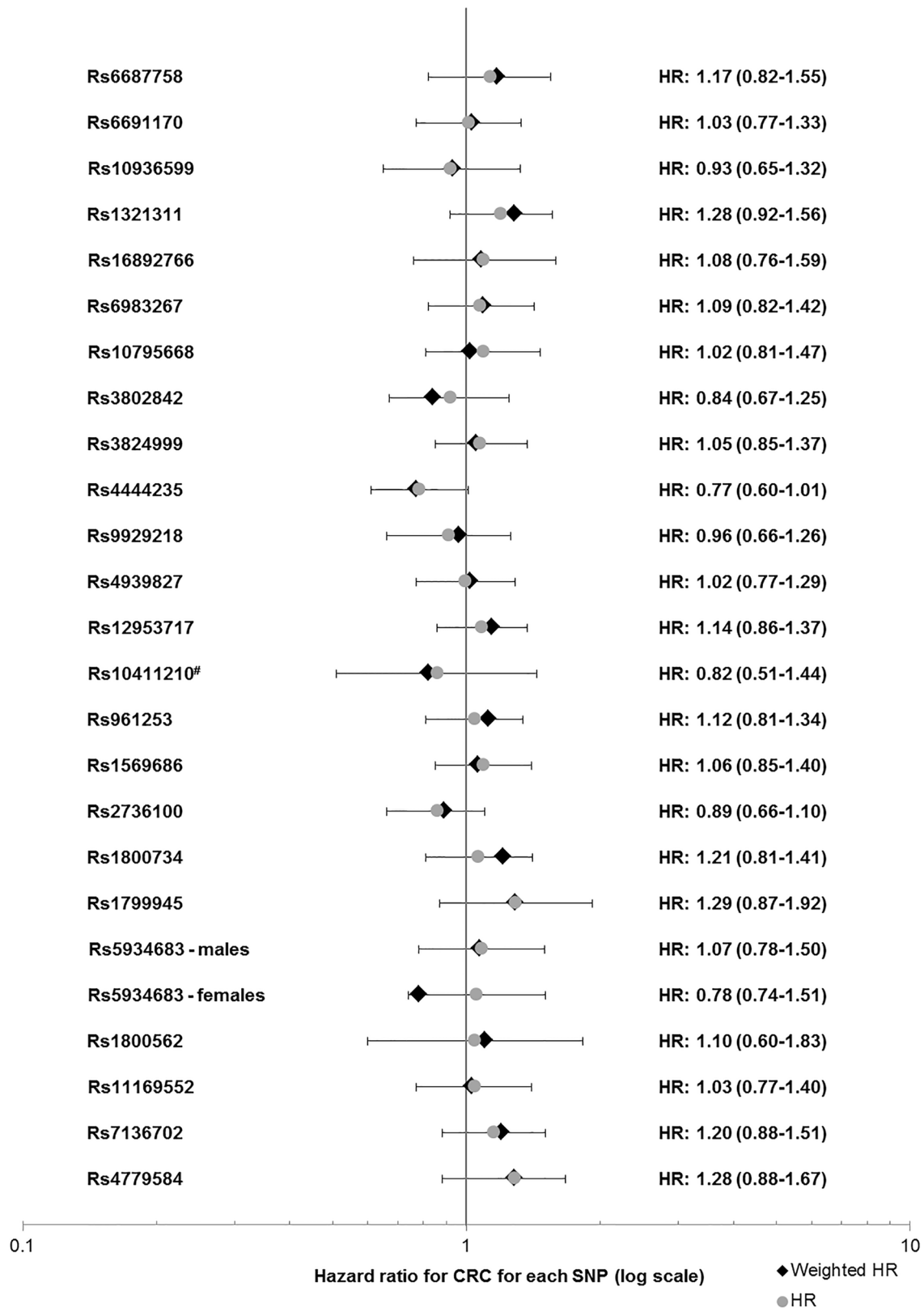
None of the SNPs individually showed a clear risk modifying effect (Fig. 1, Supplemental Table 1). There was a difference in HR between male and female *PMS2* carriers for

**Table 1** Cohort description

	No CRC—controls (n = 383)	CRC—cases (n = 124)	All (n = 507)
Sex			
Male	133 (34%)	60 (48%)	193 (38%)
Female	250 (65%)	64 (52%)	314 (62%)
Age (CRC or censoring)			
Mean (s.d.)	51.0 (14.2)	52.5 (12.7)	
Range	25–88	27–88	
Index carrier			
Yes	38 (10%)	89 (72%)	127 (25%)
No, family member	345 (90%)	35 (28%)	380 (75%)
Other cancers (no. of carriers)			
Endometrial cancer <sup>#</sup>	30	9	39
Ovarian <sup>#</sup>	4	0	4
Duodenal cancer <sup>#</sup>	4	2	6
Breast <sup>#</sup>	10	3	13
Urothelial <sup>#</sup>	4	2	6
Esophagus	1	0	1
Leukemia	0	3	3
Testis	2	0	2
Prostate	1	1	2
Vagina	0	1	1
Mesothelioma	0	1	1

'Index carrier' means the first person to be tested. Incidence of cancer in the group of index carriers without CRC: 20 endometrial cancers, 4 ovarian cancers, 3 breast cancers, 3 cancers of the small intestine, 1 testis cancer and 1 carcinoid. Ten of these index carriers had not developed any cancer at the time of DNA diagnostics; they were tested because of polyps at an early age or because they had an (affected) deceased family member

<sup>#</sup> Lynch syndrome associated cancer



**Fig. 1** Forest plot of HRs for all SNPs. rs5934683 lies on the X chromosome and was therefore stratified for gender. \*SNPs previously associated with increased risk in *MLH1* mutation carriers. <sup>#</sup>Reference

category: homozygous for risk allele (due to low number of homozygous carriers of the non-risk allele). *HR* Hazard ratio

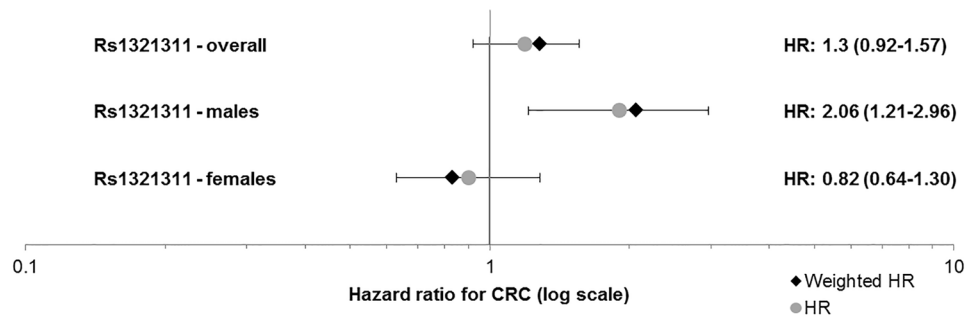
rs1321311 (6p21.31), with an HR for the each additional A allele of 2.1 (95% CI 1.2–3.0,  $p=0.005$ ) and 0.83 (95% CI 0.63–1.28,  $p=0.56$ ) for males and females, respectively (Fig. 2; Forest plot, Supplemental Fig. 2: KM curve).

**Combination of rs3802842 and rs16892766**

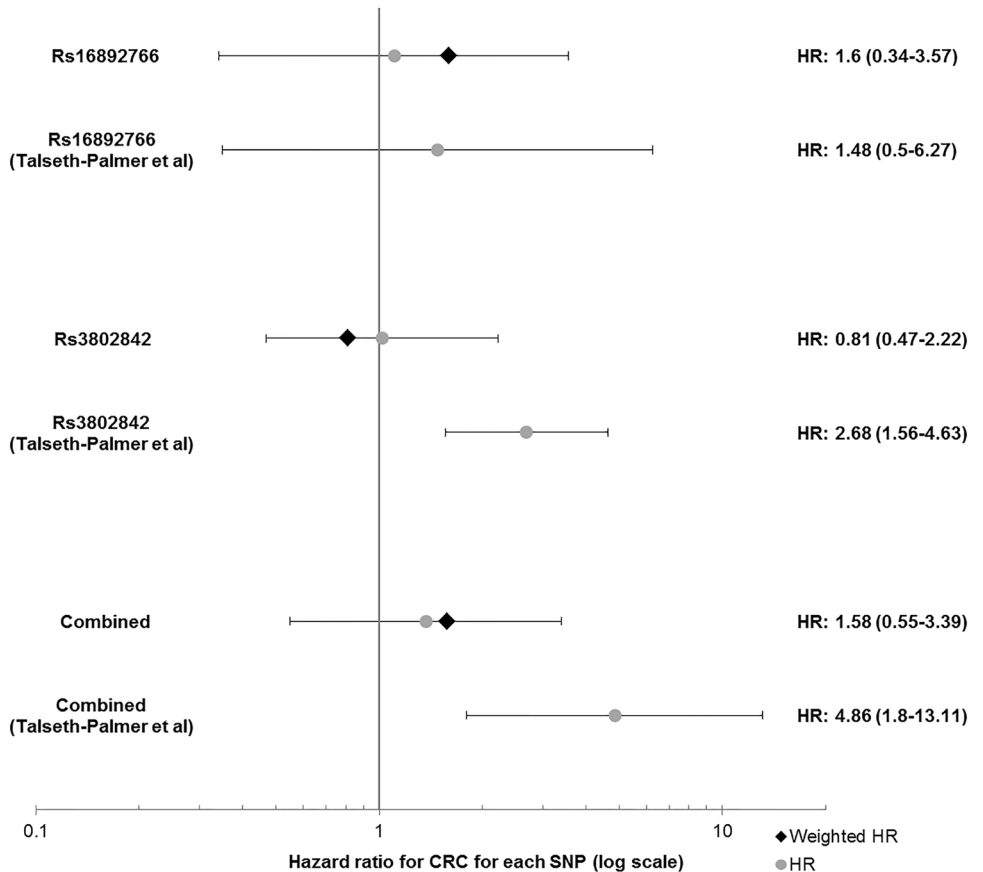
A previous meta-analysis reported a significant pairwise effect on CRC risk of rs3802842 (11q23.1) and rs16892766 (8q23.3) in *MLH1* mutation carriers [6]. The HR in the additive model for this combination in our *PMS2* cohort was 0.95 (95% CI 0.80–1.25,  $p=0.99$ ). For carriers of more than three risk alleles the HR was 1.58 (95% CI 0.55–3.39) compared

to patients with no risk alleles, see Fig. 3 for a comparison of previously publishes HRs and results from this study. The mean age at CRC diagnosis for 0, 1, and more than 1 risk alleles was 52.8; 52.9 and 50.4 respectively. The corresponding median ages were 54 [interquartile range (IQR): 43–62], 51 (IQR:43–63) and 47 (IQR:39–63). There was a statistically non-significant difference between the median age of CRC diagnosis between male and female carriers of two or more risk alleles, namely 53 (IQR:39–64,  $n=7$ ) and 43.5 years to age (IQR:38–63,  $n=10$ ,  $p=0.56$ , Mann–Whitney test).

**Fig. 2** Forest plot of HRs for rs1321311.  $p=0.005$  for males. HR Hazard ratio



**Fig. 3** Forest plot of HRs for rs3802842 and rs16892766. For the combination of the two SNPs, the plotted HR represents a comparison for carriers of three vs. no risk alleles. HR Hazard ratio

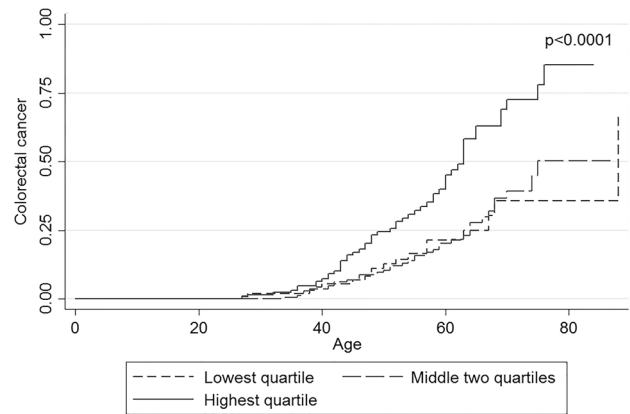


## Polygenic risk score

The polygenic risk score was calculated for 444 *PMS2* carriers with complete genotyping. The medians for PRS1 (meta-analysis derived ORs) were  $-0.12$  [interquartile range (IQR):  $-0.48$  to  $0.30$ ] for controls and  $-0.03$  (IQR:  $-0.39$  to  $0.40$ ) for CRC cases. The HRs for group 2 (second and third quartile) and 3 (fourth quartile) were  $1.33$  (95% CI  $0.76$ – $2.33$ ) and  $1.50$  (95% CI  $0.82$ – $2.72$ ) respectively (Table 2). The medians for PRS2 (based on HRs from our own data) were  $0.30$  (IQR:  $-0.057$  to  $0.55$ ) for controls and  $0.51$  (IQR:  $0.068$ – $0.75$ ) for CRC cases. The corresponding HRs for group 2 and 3 were  $1.05$  (95% CI  $0.59$ – $1.89$ ) and  $2.62$  (95% CI  $1.49$ – $4.60$ ) respectively (Table 2). The KM curves for PRS1 and 2 are shown in and Supplemental Fig. 3 and Fig. 4, respectively. The difference between survival curves was highly significant for PRS2 ( $p < 0.0001$ ). The optimism-corrected c statistic was only  $0.52$ , indicating no discriminatory value.

## Discussion

*PMS2* carriers currently represent a relatively small proportion of LS patients. However, the number of *PMS2*-associated LS cases is expected to rise with the implementation of population-based screening protocols for all CRC below age 70. Identification of *PMS2* carriers has been challenging in the past due to difficulties in mutation analysis, a milder phenotype and many families not fulfilling clinical selection criteria [3, 9, 15–17]. Obtaining a better understanding of the specific *PMS2*-associated phenotype is particularly relevant, as it appears to differ markedly from phenotypes associated with other MMR mutations. Unfortunately, we were



**Fig. 4** Kaplan Meier survival curve for PRS2. This plot compares curves for the lowest, the two middle and the highest quartile of the PRS. PRS2 is based on hazard ratios from the current study. *HR* Hazard ratio, *PRS* polygenic risk score

unable to confirm any risk modifying effects of rs3802842 (11q23.1) and rs16892766 (8q23.1), two SNPs previously shown to be associated with enhanced risk in *MLH1* mutation carriers [5, 6]. Studies in *MLH1* mutation carriers reported that a higher number of risk alleles in a carrier is associated with a younger onset of disease (28 years younger for 3 compared to 0 risk alleles) [6]. In our cohort, mean ages where 52.8 and 50.4 for 0 compared to more than 1 risk allele, respectively. As such, there seems to be no clinical utility of rs3802842 and rs16892766 in risk stratification for *PMS2* carriers.

Many studies on (genetic) modifiers in LS patients focus on *MLH1* and *MSH2*, or *MSH6* carriers, while *PMS2* is seldom analyzed. The only study to include *PMS2* carriers ( $n = 40$ ) found that carriers of the G-alleles of rs10795668

**Table 2** Polygenic risk scores

PRS category		Controls	CRC	Expected events	HR	p for HR
PRS1	Median (IQR)	$-0.12$ ( $-0.48$ to $0.30$ )	$-0.03$ ( $-0.39$ to $0.40$ )			
	First quartile	84	22	28	Ref	0.41
	Second and third quartile	167	54	52	$1.33$ ( $0.76$ – $2.33$ )	0.31
	Fourth quartile	84	33	28	$1.50$ ( $0.82$ – $2.72$ )	0.19
PRS2	Median (IQR)	$0.30$ ( $-0.057$ to $0.55$ )	$0.51$ ( $0.068$ – $0.75$ )			
	First quartile	84	18	26	Ref	$<0.0001$
	Second and third quartile	165	38	53	$1.05$ ( $0.59$ – $1.89$ )	0.86
	Fourth quartile	84	53	30	$2.62$ ( $1.49$ – $4.60$ )	0.001

*PRS* polygenic risk score, *IQR* interquartile range

*PRS1* Weighted on odds ratios from general population, i.e. in sporadic CRC cases

*PRS2* Weighted on hazard ratios from this study

Log rank survival curves PRS1:  $p = 0.32$

Log rank survival curves PRS2:  $p < 0.0001$

(10p14) and rs9929218 (16q22.1) were at lower risk of CRC, a notable finding in that this is the opposite effect compared to sporadic CRC [7]. The authors conceded that their results should be confirmed in larger studies. As these findings have not been confirmed in our much larger cohort, we suggest that these previous findings may indeed have been false positives due to the small number of carriers included.

A relevant question is why our study did not confirm reported findings of previous studies of MMR carriers. One explanation might be that although patients have germline mutations in genes with similar functions, carriers are affected by genetic modifiers in different ways. Indeed, comparable studies in *BRCA1* or *BRCA2* mutation carriers have resulted in the identification of SNPs that clearly modify breast cancer risk. However, *BRCA2* carriers appear unaffected by SNPs that confer an increased breast cancer risk in *BRCA1* carriers, even though both genes play a role in homologous recombination [18]. This could also hold for MMR mutation carriers, as illustrated by the observation that while rs3802842 and rs16892766 may increase risk in *MLH1* carriers, they do not appear to have an effect in *MSH2* or *PMS2* carriers [5, 6]. Researchers should therefore concentrate on building cohorts large enough to analyze Lynch patients in a gene-stratified manner.

Gender stratification in our cohort led to the notable finding that male carriers of allele A at locus rs1321311 (6p21.31) show a per allele HR of 2.07 (95% CI 1.21–2.96), while the HR for females was 0.83 (95% CI 0.63–1.28). This SNP has been linked to the *CDKN1A* gene that encodes the p21 protein. p21 is involved in several (p53-independent) pathways as a tumor suppressor, although it also has oncogenic characteristics [19, 20]. Interestingly, down-regulation of p21 is inversely associated with MSI, the hallmark of Lynch-associated tumors. One study found that a larger proportion of Lynch-associated CRCs expressed p21 compared to sporadic CRCs (80 vs. 31%) [21]. However, a recent study reported on expressive Quantitative Trait Loci (eQTL) in colonic tissue based on data from the GTEx project portal (<http://www.gtexportal.org/home/>) and did not find a statistically significant effect of rs1321311 on *CDKN1A* expression in sigmoid and transverse colon tissue ( $p=0.84$  and  $p=1.00$  respectively) [22, 23]. It is also unclear why this effect only appears to be present in male *PMS2* carriers. Although no gender difference was noted by the meta-analysis that identified the SNP [19], it is possible that gender differences exist, as CRC risk in the general population and in Lynch patients is known to be higher in men compared to women [24, 25]. Indeed, mutations are more often found in males than females when assessed for Lynch syndrome [26]. Another possible explanation for this gender specific effect might lie in the effect of other risk modifiers. It is perceivable that other factors than SNPs have a stronger influence on (colorectal) cancer development in women, such as

hormonal factors [27, 28]. Unfortunately data on hormone levels or other factors previously shown to modify cancer risk in LS such as medication use (e.g. aspirin) or environmental factors were unavailable for analysis and as such we were unable to correct for this [29–33]. Similarly, data on smoking and BMI were only available for a small proportion of carriers ( $n=131$ , 26%). It should be emphasized that all results after gender stratification should be interpreted with caution because of small sample size and multiple testing. This could have led to false associations. Further studies are needed to validate these findings.

We also investigated the effect of the 24 SNPs on CRC risk in the *PMS2* cohort by means of a polygenic risk score (PRS). While there did not appear to be a significant effect of the PRS based on ORs from sporadic CRC cohorts, there was a difference in the cumulative incidence of CRC for *PMS2* carriers with a PRS2 (based on HRs in this study) in the highest quartile. Bootstrap validation however refuted this promising observation. Further studies are hence needed in other large cohorts.

There were some limitations to this study. Our study consisted exclusively of Dutch *PMS2* carriers and thus had a relatively homogeneous genetic makeup, implying that differences between our results and previous studies might be due to population-specific effects.

A second limitation might be that we did not correct for the specific mutation present in each family, mainly because in the majority of families the segregating *PMS2* mutation is rare or even unique. A previous study by our group did not identify such a correlation with CRC risk in *PMS2* carriers (Supplemental Tables 2a + 2b Table: for more details) [34].

Unfortunately, we were not able to validate our findings in an external cohort. To our best knowledge this is one of the largest *PMS2* cohorts currently collected, and bootstrap validation is a strong approach to assess discriminative ability of a prediction model [35]. Stratifying our cohort into a discovery and validation cohort was not a viable option as this would have resulted in a substantial decrease in power. Our study might already have been underpowered to detect weak associations. However, while such associations are interesting from a scientific point of view and may be relevant to tumorigenesis, they are not necessarily useful in clinical practice when the effect is small. For the two SNPs previously found to increase risk in *MLH1* mutation carriers, we had 60–80% power to detect an HR of 1.5, which we would consider clinically relevant. The previously reported HR in *MLH1* carriers for rs3802842 was 2.7, an HR for which we have ample power to detect (Supplemental Fig. 1).

Families with a segregating *PMS2* mutation show a high degree of phenotypic variability. We were not able to confirm the risk modifying effect of rs3802842 (11q23.1) and rs16892766 (8q23.3), which were previously found to increase the risk in *MLH1*-associated LS. This, together



with the established lower penetrance, raises the question of whether *PMS2*-associated LS should be considered a separate Lynch disease entity. Additional explanations for phenotypic variability that warrant greater exploration include gene-environment interactions and risk modification by other genetic variants.

**Acknowledgements** The authors would like to thank Ruud van der Breggen (LUMC) for technical support. Writing assistance: The final manuscript was edited by Medactie, funded by the Dutch Cancer Society.

**Funding** This work was supported by the Dutch Cancer Society (Grant: UL-2012-5515). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Compliance with ethical standards

**Conflict of interest** The authors report no conflict of interests.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Dowty JG, Win AK, Buchanan DD, Lindor NM, Macrae FA, Clendenning M, Antill YC, Thibodeau SN, Casey G, Gallinger S, Marchand LL, Newcomb PA, Haile RW, Young GP, James PA, Giles GG, Gunawardena SR, Leggett BA, Gattas M, Boussioutas A, Ahnen DJ, Baron JA, Parry S, Goldblatt J, Young JP, Hopper JL, Jenkins MA (2013) Cancer risks for MLH1 and MSH2 mutation carriers. *Hum Mutat* 34(3):490–497
- Baglietto L, Lindor NM, Dowty JG, White DM, Wagner A, Gomez Garcia EB, Vriends AH, Cartwright NR, Barnetson RA, Farrington SM, Tenesa A, Hampel H, Buchanan D, Arnold S, Young J, Walsh MD, Jass J, Macrae F, Antill Y, Winship IM, Giles GG, Goldblatt J, Parry S, Suthers G, Leggett B, Butz M, Aronson M, Poynter JN, Baron JA, Le ML, Haile R, Gallinger S, Hopper JL, Potter J, de la Chapelle A, Vasen HF, Dunlop MG, Thibodeau SN, Jenkins MA (2010) Risks of Lynch syndrome cancers for MSH6 mutation carriers. *J Natl Cancer Inst* 102(3):193–201
- ten Broeke SW, Brohet RM, Tops CM, van der Klift HM, Velthuisen ME, Bernstein I, Capella Munar G, Gomez Garcia E, Hoogerbrugge N, Letteboer TG, Menko FH, Lindblom A, Mensenkamp AR, Moller P, van Os TA, Rahner N, Redeker BJ, Sijmons RH, Spruijt L, Suerink M, Vos YJ, Wagner A, Hes FJ, Vasen HF, Nielsen M, Wijnen JT (2015) Lynch syndrome caused by germline *PMS2* mutations: delineating the cancer risk. *J Clin Oncol* 33(4):319–325
- Ma X, Zhang B, Zheng W (2014) Genetic variants associated with colorectal cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Gut* 63(2):326–336
- Wijnen JT, Brohet RM, van ER, Jagmohan-Changur S, Middeldorp A, Tops CM, PM van, MG Ausems, Gomez GE, Hes FJ, Hoogerbrugge N, Menko FH, van Os TA, Sijmons RH, Verhoef S, Wagner A, Nagengast FM, Kleibeuker JH, Devilee P, Morreau H, Goldgar D, Tomlinson IP, Houlston RS, Van WT, Vasen HF (2009) Chromosome 8 q23.3 and 11q23.1 variants modify colorectal cancer risk in Lynch syndrome. *Gastroenterology* 136(1):131–137
- Talseth-Palmer BA, Wijnen JT, Brenne IS, Jagmohan-Changur S, Barker D, Ashton KA, Tops CM, Evans TJ, McPhillips M, Groombridge C, Suchy J, Kurzawski G, Spigelman A, Moller P, Morreau HM, Van WT, Lubinski J, Vasen HF, Scott RJ (2013) Combined analysis of three Lynch syndrome cohorts confirms the modifying effects of 8q23.3 and 11q23.1 in MLH1 mutation carriers. *Int J Cancer* 132(7):1556–1564
- Win AK, Hopper JL, Buchanan DD, Young JP, Tenesa A, Dowty JG, Giles GG, Goldblatt J, Winship I, Boussioutas A, Young GP, Parry S, Baron JA, Duggan D, Gallinger S, Newcomb PA, Haile RW, Le ML, Lindor NM, Jenkins MA (2013) Are the common genetic variants associated with colorectal cancer risk for DNA mismatch repair gene mutation carriers? *Eur J Cancer* 49(7):1578–1587
- Houille S, Charbonnier F, Houivet E, Tinat J, Buisine MP, Caron O, Benichou J, Baert-Desurmont S, Frebourg T (2011) Evaluation of Lynch syndrome modifier genes in 748 MMR mutation carriers. *Eur J Hum Genet* 19(8):887–892
- Senter L, Clendenning M, Sotamaa K, Hampel H, Green J, Potter JD, Lindblom A, Lagerstedt K, Thibodeau SN, Lindor NM, Young J, Winship I, Dowty JG, White DM, Hopper JL, Baglietto L, Jenkins MA, de la Chapelle A (2008) The clinical phenotype of Lynch syndrome due to germ-line *PMS2* mutations. *Gastroenterology* 135(2):419–428
- Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, Hamilton SR, Hiatt RA, Jass J, Lindblom A, Lynch HT, Peltomaki P, Ramsey SD, Rodriguez-Bigas MA, Vasen HF, Hawk ET, Barrett JC, Freedman AN, Srivastava S (2004) Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* 96(4):261–268
- Vasen HF, Tomlinson I, Castells A (2015) Clinical management of hereditary colorectal cancer syndromes. *Nat Rev Gastroenterol Hepatol* 12(2):88–97
- Antoniou AC, Goldgar DE, Andrieu N, Chang-Claude J, Brohet R, Rookus MA, Easton DF (2005) A weighted cohort approach for analysing factors modifying disease risks in carriers of high-risk susceptibility genes. *Genet Epidemiol* 29(1):1–11
- Williams RL (2000) A note on robust variance estimation for cluster-correlated data. *Biometrics* 56(2):645–646
- Dudbridge F (2013) Power and predictive accuracy of polygenic risk scores. *PLoS Genet* 9(3):e1003348
- Clendenning M, Hampel H, LaJeunesse J, Lindblom A, Lockman J, Nilbert M, Senter L, Sotamaa K, de la Chapelle A (2006) Long-range PCR facilitates the identification of *PMS2*-specific mutations. *Hum Mutat* 27(5):490–495
- van der Klift HM, Tops CM, Bik EC, Boogaard MW, Borgstein AM, Hansson KB, Ausems MG, Gomez GE, Green A, Hes FJ, Izatt L, van Hest LP, Alonso AM, Vriends AH, Wagner A, van Zelst-Stams WA, Vasen HF, Morreau H, Devilee P, Wijnen JT (2010) Quantification of sequence exchange events between *PMS2* and *PMS2CL* provides a basis for improved mutation scanning of Lynch syndrome patients. *Hum Mutat* 31(5):578–587
- van der Klift HM, Mensenkamp AR, Drost M, Bik EC, Vos YJ, Gille HJ, Redeker BE, Tiersma Y, Zonneveld JB, Garcia EG, Letteboer TG, Olderode-Berends MJ, van Hest LP, van Os TA, Verhoef S, Wagner A, van Asperen CJ, Ten Broeke SW, Hes FJ, de Wind N, Nielsen M, Devilee P, Ligtenberg MJ, Wijnen JT, Tops CM (2016) Comprehensive mutation analysis of *PMS2* in a large Cohort of Proband suspected of lynch syndrome or constitutional mismatch repair deficiency (CMMRD) syndrome. *Hum Mutat* 37(11): 1162–1179

18. Ingham SL, Warwick J, Byers H, Lalloo F, Newman WG, Evans DG (2013) Is multiple SNP testing in BRCA2 and BRCA1 female carriers ready for use in clinical practice? Results from a large Genetic Centre in the UK. *Clin Genet* 84(1):37–42
19. Dunlop MG, Dobbins SE, Farrington SM, Jones AM, Palles C, Whiffin N, Tenesa A, Spain S, Broderick P, Ooi LY, Domingo E, Smillie C, Henrion M, Frampton M, Martin L, Grimes G, Gorman M, Semple C, Ma YP, Barclay E, Prendergast J, Cazier JB, Olver B, Penegar S, Lubbe S, Chander I, Carvajal-Carmona LG, Ballereau S, Lloyd A, Vijayakrishnan J, Zgaga L, Rudan I, Theodoratou E, Starr JM, Deary I, Kirac I, Kovacevic D, Aaltonen LA, Renkonen-Sinisalo L, Mecklin JP, Matsuda K, Nakamura Y, Okada Y, Gallinger S, Duggan DJ, Conti D, Newcomb P, Hopper J, Jenkins MA, Schumacher F, Casey G, Easton D, Shah M, Pharoah P, Lindblom A, Liu T, Smith CG, West H, Cheadle JP, Midgley R, Kerr DJ, Campbell H, Tomlinson IP, Houlston RS (2012) Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. *Nat Genet* 44(7):770–776
20. Abbas T, Dutta A (2009) p21 in cancer: intricate networks and multiple activities. *Nat Rev Cancer* 9(6):400–414
21. Sinicrope FA, Roddey G, Lemoine M, Ruan S, Stephens LC, Frazier ML, Shen Y, Zhang W (1998) *Loss of p21WAF1/Cip1 protein expression accompanies progression of sporadic colorectal neoplasms but not hereditary nonpolyposis colorectal cancers*. *Clin Cancer Res* 4(5):1251–1261
22. Loo LWM, Lemire M, Le Marchand L (2017) In silico pathway analysis and tissue specific cis-eQTL for colorectal cancer GWAS risk variants. *BMC Genom* 18(1):381
23. Battle A, Brown CD, Engelhardt BE, Montgomery SB (2017) Genetic effects on gene expression across human tissues. *Nature* 550(7675):204–213
24. Barrow E, Hill J, Evans DG (2013) Cancer risk in Lynch Syndrome. *FamCancer* 12(2):229–240
25. Haggard FA, Boushey RP (2009) Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clin Colon Rectal Surg* 22(4):191–197
26. Kastrinos F, Uno H, Ukaegbu C, Alvero C, McFarland A, Yurgelun MB, Kulke MH, Schrag D, Meyerhardt JA, Fuchs CS, Mayer RJ, Ng K, Steyerberg EW, Syngal S (2017) Development and validation of the PREMM5 Model for comprehensive risk assessment of lynch syndrome. *J Clin Oncol* 35(19):2165–2172
27. Dashti SG, Chau R, Ouakrim DA, Buchanan DD, Clendenning M, Young JP, Winship IM, Arnold J, Ahnen DJ, Haile RW, Casey G, Gallinger S, Thibodeau SN, Lindor NM, Le Marchand L, Newcomb PA, Potter JD, Baron JA, Hopper JL, Jenkins MA, Win AK (2015) Female hormonal factors the risk of endometrial cancer in Lynch Syndrome. *Jama* 314(1):61–71
28. Jori B, Kamps R, Xanthouleas S, Delvoux B, Blok MJ, Van de Vijver KK, de Koning B, Oei FT, Tops CM, Speel EJ, Kruitwagen RF, Gomez-Garcia EB, Romano A (2015) Germ-line variants identified by next generation sequencing in a panel of estrogen and cancer associated genes correlate with poor clinical outcome in Lynch syndrome patients. *Oncotarget* 6(38):41108–41122
29. Movahedi M, Bishop DT, Macrae F, Mecklin JP, Moeslein G, Olschwang S, Eccles D, Evans DG, Maher ER, Bertario L, Bisgaard ML, Dunlop MG, Ho JW, Hodgson SV, Lindblom A, Lubinski J, Morrison PJ, Murday V, Ramesar RS, Side L, Scott RJ, Thomas HJ, Vasen HF, Burn J, Mathers JC (2015) Obesity, aspirin, and risk of colorectal cancer in carriers of hereditary colorectal cancer: a prospective investigation in the CAPP2 Study. *J Clin Oncol* 33(31): 3591–3597
30. Win AK, Dowty JG, English DR, Campbell PT, Young JP, Winship I, Macrae FA, Lipton L, Parry S, Young GP, Buchanan DD, Martinez ME, Jacobs ET, Ahnen DJ, Haile RW, Casey G, Baron JA, Lindor NM, Thibodeau SN, Newcomb PA, Potter JD, Le Marchand L, Gallinger S, Hopper JL, Jenkins MA (2011) Body mass index in early adulthood and colorectal cancer risk for carriers and non-carriers of germline mutations in DNA mismatch repair genes. *Br J Cancer* 105(1):162–169
31. Van Duijnhoven FJ, Botma A, Winkels R, Nagengast FM, Vasen HF, Kampman E (2013) Do lifestyle factors influence colorectal cancer risk in Lynch syndrome? *Fam Cancer* 12(2):285–293
32. Pande M, Lynch PM, Hopper JL, Jenkins MA, Gallinger S, Haile RW, LeMarchand L, Lindor NM, Campbell PT, Newcomb PA, Potter JD, Baron JA, Frazier ML, Amos CI (2010) Smoking and colorectal cancer in Lynch syndrome: results from the Colon Cancer Family Registry and the University of Texas M.D. Anderson Cancer Center. *Clin Cancer Res* 16(4):1331–1339
33. Botma A, Nagengast FM, Braem MG, Hendriks JC, Kleibeuker JH, Vasen HF, Kampman E (2010) Body mass index increases risk of colorectal adenomas in men with Lynch syndrome: the GEOLynch cohort study. *J Clin Oncol* 28(28):4346–4353
34. Suerink M, van der Klift HM, ten Broeke SW, Dekkers OM, Bernstein I, Capella Munar G, Gomez Garcia E, Hoogerbrugge N, Letteboer TGW, Menko FH, Lindblom A, Mensenkamp A, Moller P, van Os TA, Rahner N, Redeker BJW, Olderoode M, Spruijt L, Vos YJ, Wagner A, Morreau H, Hes FJ, Vasen HFA, Tops CM, Wijnen JT, Nielsen M (2016) The effect of genotypes parent of origin on cancer risk age of cancer development in PMS2 mutation carriers *Genet Med* 18(4):405–409
35. Steyerberg EW (2009) *Clinical Prediction Models*. Statistics for Biology and Health. Springer, New York