The path to individualised breast cancer screening

Lakeman, I.M.M.

CHAPTER 7

7

Discussion and future perspectives

Today, in Dutch clinical genetic services, breast cancer risk prediction is mainly based on family history and carrier status of pathogenic variants in one of the five well known breast cancer genes (i.e. *BRCA1, BRCA2, PALB2, CHEK2*, *ATM*). Family history is an important risk factor for breast cancer. On average, healthy women with at least one first degree relative affected with breast cancer have a relative risk of developing breast cancer of ~2-fold. Last decade, we have gained more knowledge about the aetiology of this familial relative risk, which could improve breast cancer risk prediction in terms of precision and accuracy. Combining all known genetic, familial and lifestyle risk factors will give a more individual based lifetime risk score. In this thesis we have explored the clinical utility of the use of the currently known common low risk variants associated with breast cancer, which explain ~18% of the familial relative risk, for individual breast cancer risk prediction. Especially for families that visit the clinical genetic services in the Netherlands.

7.1 Dutch breast cancer families

In **chapters 2 and 3**, we explored the clinical applicability of the Polygenic Risk Score (PRS) for risk prediction in a cohort of breast cancer families not explained by *BRCA1* or *BRCA2* pathogenic variants, that had visited the clinical genetic services in the Netherlands. It was known that the PRS could improve the discriminative power between breast cancer cases and controls[1-5], but little was known about this discriminative power within families and the additive impact on family-based risk prediction in these families.

In **chapter 2**, high risk breast cancer families were analysed that were selected for genetic research purposes and counselled between 1990 and 2012. An advantage of this cohort was the availability of a DNA sample of both affected and unaffected family members. While most studies use population controls as a reference group[2-4, 6], we used healthy relatives of breast cancer cases as a reference to make it more compatible with clinical practice in clinical genetic services. Only three previous studies have also genotyped breast cancer cases and their unaffected relatives, but with a lower number of variants included in their PRS[7-9]. The PRS in this study was based on 161 breast cancer associated variants which were known at that time[10]. Within our cohort of high-risk families, affected family members had on average a higher PRS compared to their healthy relatives, suggesting already an association between this PRS and breast cancer. Association analyses proved the effect of the PRS, showing a significant association (HR per SD=1.16) within high-risk families between the PRS and breast cancer. As presented in **chapter 3** and described in the literature as well[2, 7], we observed just a very weak positive correlation between the PRS and the family history score, calculated by BOADICEA version 3 using the complete pedigree. This result underscores the additive value of measuring the PRS for every individual in the family, as opposed to using an estimated PRS based on the family history.

With BOADICEA version 3, in which the PRS was not yet implemented[11], lifetime risks (i.e. breast cancer risk between age 20 and age 80) were calculated with and without the PRS in addition to family history-based risk prediction. By adding the PRS, about 20% of both affected and unaffected women were reclassified to another risk category and would have received a different screening advise based on the Dutch breast cancer screening guideline[12].

In **chapter 3** we selected breast cancer cases with a positive family history for breast cancer that visited one of the clinical genetic services in the Netherlands, without a pathogenic variant in *BRCA1* or *BRCA2*. These cases were more representative of the average breast cancer families counselled in the clinic than those analysed in **chapter 2**. The best predictive PRS for breast cancer known at this moment was calculated based on 313 common low risk variants (PRS_{313}). Again, as expected, this PRS was on average higher for breast cancer cases versus population controls. Furthermore, women who developed an *in situ* carcinoma had on average a lower PRS_{313} compared to women who developed an invasive tumour but a higher PRS compared to population controls. Between family members, 13% of the variance in the PRS_{313} could be explained by the PRS_{313} of the proband (case with the youngest diagnosis), hence the proband's PRS was only modestly predictive of that of family members. A significant association was determined in this family-based cohort between breast cancer and the PRS_{313}, OR per SD=1.97, with a stronger effect for invasive compared with *in situ* carcinoma (OR per SD=2.00, and 1.69 respectively). For the majority, gene panel sequencing was performed for at least *CHEK2, ATM* and *PALB2*. In total 1.8% of the controls and 8.4% of the cases carried a truncating pathogenic variant in one of these genes, most frequently in *CHEK2*. Using BOADICEA version 5 where the PRS_{313} is implemented[13], family history-based breast cancer lifetime risk scores were calculated including the pedigree and gene-panel result. In addition to this family history-based score, the individual PRS_{313} was included. For up to 34% of the gene-panel negative cases, screening recommendations could have changed by adding the PRS_{313} to family history-based risk prediction. Addition of the PRS_{313} had a large impact on screening recommendations for *ATM* and *CHEK2* pathogenic variant carriers as well, corresponding to the suggested polygenic effect of moderate risk breast cancer genes. No change was detected for carriers of a *PALB2* pathogenic variant, who all remained in the high-risk category, although variations in risk scores may have impact on choices that women make regarding prophylactic surgery.

These family-based studies are important for implementation of the PRS in the clinic. Using information from breast cancer families which recently visited clinical genetic services, provides a good representation of the group of counselees from families that are seen in the context of clinical genetic services. Furthermore, an advantage of selecting "genetically enriched" cases is that we had a sufficient number of pathogenic variant

carriers in *CHEK2*, and *ATM* in our cohort to show the reclassification (i.e. change to a different screening category) for this group of women as well. However, selecting families with an average higher risk for developing breast cancer, resulting in a higher prevalence of breast cancer in this group compared to the population, causes ascertainment bias so that the effect sizes obtained in these studies cannot be translated directly in the clinic. The higher effect size in our study (**chapter 3,** OR=1.97) compared to population based cohorts of the Breast Cancer Association Consortium (BCAC) (OR=1.61)[14] and in the Dutch population (**chapter 4**, HR=1.56) possibly reflects a higher genetic predisposition in our families. This is also supported by the on average higher PRS for healthy relatives of breast cancer cases compared to population controls and the lower association effect size of the PRS and breast cancer within high-risk families (**chapter 2**, HR=1.16). Although we adjusted for family history, it does probably not suffice to correct for ascertainment bias. This illustrates an important problem in family history-based studies: they lead to overestimation of disease penetrance, which underscores the need for careful separation of family history and the PRS and estimating their effects for the general population. Although we are seeing this selected group of families in the clinic as well, separation of the two risk factors, family history and PRS, will be more specific for an individual. Separation of these risk becomes more important since, compared to 10 years ago, fewer affected families are counselled at this moment.

Both studies showed a quite large percentage of women who changed to another risk category (reclassification) and would have received a different corresponding clinical advice after including the PRS in addition to family history-based risk prediction. These reclassifications were based on breast cancer lifetime risk scores which were mainly calculated for cases (affected counselees), assuming they were 1 year old and unaffected, while in clinical practice the risk scores are only calculated for unaffected family members. Therefore, the reclassification percentage may be different for healthy relatives of affected counselees; the majority of those will have a higher family history-based score, because the affected proband will be included as affected family member. Some studies address this inconsistency by calculating the score for an additional imaginary healthy sister but because the PRS is an individual score, this is not possible in our studies.

In **chapter 6**, we have performed a small pilot study in which we calculated, by using BOADICEA version 5, similar breast cancer lifetime risk scores for 38 unaffected first-degree relatives of women affected with breast cancer, who had already visited the clinical genetic service for breast cancer counselling and tested negative for germline pathogenic variants in one of five breast cancer predisposing genes (*BRCA1/2, PALB2, CHEK2, ATM*). By including family history, non-genetic risk factors, and the PRS, 18 women (47%) changed to a different screening category as compared to the current standard risk prediction including family history alone [Tüchler et al. *manuscript in preparation*]. These results

suggest that if we introduced bias by including cases only, the true reclassification rate for unaffected relatives is probably not lower as the 34% described in **chapter 3**, keeping the conclusion that our results underscore the utility of including the PRS.

Although we found a large percentage of reclassification, the question remains if the direction is correct. Ideally you want to have a prospective cohort of women with data about their screening uptake, breast cancer development and detection of the tumour, i.e. screen detected or interval carcinoma. Unfortunately this information was lacking in both of our cohorts. We had information about the age of diagnosis that would help determine if the cases were retrospectively placed in the right risk category. However, without having information on the detection of the tumour, it is difficult to interpret which category would be the right one. For example, if a woman who was reclassified into the high risk group was diagnosed with breast cancer at age 56, a mammogram biannually via population screening could have been sufficient to detect it, but the recommended annual mammogram for this risk group following the Dutch guideline[12] might have detected the tumour earlier. Based on the knowledge that the BOADICEA model is well calibrated and validated in different prospective studies[15-17] as well as in our study described in **chapter 4** for the Dutch population, we assume that the reclassification leads to the detection of more breast cancers overall and less side-effects of screening such as false positives and overdiagnosis. However, to optimise these benefits of individual risk-based screening, we need to be confident enough to downgrade screening for a part of the women. But even if we can demonstrate cost-efficiency and accept that the reclassification will on average be better for the total group, it remains difficult to translate it to a specific person as seen by a clinician. As clinician you have to decide for that person at that point in time, which method will best manage the real risk for a person and downgrading may be a challenge.

7.2 Breast cancer in the Dutch population

In **chapter 4**, the performance of BOADICEA version 5 and the association with the PRS_{313} was evaluated for the Dutch population. Furthermore, we illustrated the potential impact of the model in detecting breast cancer in a population screening setting in which women would participate based on their individual risk. Comprehensive risk prediction is possible with BOADICEA version 5, which incorporates the PRS_{313} as well as lifestyle, reproductive and hormonal risk factors, but this was not yet validated in the general Dutch population.

We used a large prospective population-based cohort of women aged 45 years or older with extensive follow-up data of up to 25 years. Women who developed breast cancer during follow up had on average a higher PRS_{313} compared to unaffected women. Furthermore, as seen in **chapter 3** as well, women who developed an invasive breast tumour had on

average a higher PRS_{313} compared to women with an *in situ* breast tumour. The PRS_{313} was significantly associated with breast cancer, with a similar effect size (HR=1.56) as in other prospective series of different geographic origin[14], demonstrating its robustness and potential application to the Dutch population. Similar as described in **chapter 3**, the PRS_{313} was associated with *in situ* breast cancer as well, with a non-significant lower effect size than for invasive breast cancer. Moreover, as described previously for a PRS based on 72 variants[18], the PRS_{313} is specifically associated with breast cancer risk and not with a higher risk for the development for a non-breast carcinoma (HR=1.05, non-significant). As determined in previous studies performed by BCAC[4, 14], we found that the effect size of the PRS declined with increasing age. With the BOADICEA model, cumulative 10-year breast cancer risk scores were calculated using four sets of variables (age; age and PRS; age and risk factors; age, PRS, and risk factors). Above inclusion of age, The PRS_{313} improved the discriminatory ability from 0.531 to 0.636. As expected, based on previous research[13, 19], this could only be marginally improved further (to 0.653) by adding lifestyle, reproductive factors, and anthropometric data. Irrespective of the variables included, BOADICEA underestimated the observed risk of 4.4% especially in the highest risk categories. This underestimation was possibly due to the lack of family history data, mammographic density and information about pathogenic variants in *BRCA1/2*. Overall, the PRS_{313} replicates robustly in the Dutch population and the discriminative power of the BOADICEA model seems appropriate for implementation into breast cancer prevention programs. However, for accurate use of the BOADICEA model in the population, information about family history could be important to add.

We illustrated the potential impact of the BOADICEA model in detecting breast cancer in a population screening setting in which women would participate based on their individual risk. In this scenario, the PRS_{313} alone would have detected more cases than the full BOADICEA model (80% versus 62% respectively), but would also have identified a larger screening group (65% versus 45% of all women). Ideally one would want to find the optimal cost-benefit ratio with the highest detection of breast cancer and the lowest false-positive and overdiagnosis rate. An important question in breast cancer risk prediction is how to include and treat *in situ* carcinomas. Although PRS development studies have so far included only invasive breast cancer[4, 14], we showed in **chapter 3** that the PRS_{313} is associated with *in situ* breast cancer as well, consistent with previous research[20] . However, there was a non-significantly lower effect size for *in situ* carcinomas compared to invasive breast cancer. Preferably, comprehensive risk prediction including the PRS_{313} will lead to a higher detection rate of *in situ* carcinomas that are more prone to become invasive and less detection of *in situ* carcinoma that will never become clinically relevant. For this goal, more knowledge is needed about prognostic markers that distinguish between these types. Previous research showed that besides growth pattern, histological grade of ductal carcinoma in situ (DCIS) has been associated with subsequent development of

invasive disease[21-23]. In our study, all women with grade 3 DCIS were in the group eligible for screening based on the PRS and age model and only 50% of the women with a grade 1/2 DCIS. Although the absolute numbers were low, this supports the notion that the PRS predicts DCIS that is more prone to become invasive breast cancer. However, further research needs to be performed to confirm this.

The high prevalence of *in situ* carcinoma nowadays, ~25% of all breast cancers[24], leads to the question, relevant for both family-members and their counsellors, whether women who develop these breast cancers should be considered as "affected" or "unaffected" in family-based risk prediction. For example, BOADICEA is presented as a model that predicts invasive breast cancer considering only invasive breast tumours in the family[13]. Ideally, it would be possible to include DCIS as well in these risk prediction models. However, epidemiological studies determining the risk for developing breast cancer for a healthy relative of someone with DCIS are lacking. Although probably the majority of DCIS will remain indolent[23], it may be possible that DCIS in some individuals within breast cancer families is more prone to become invasive due to genetic predisposition. Therefore, not including DCIS may lead to an underestimation of breast cancer risk in these families. In my opinion, until we are able to distinguish a clinically relevant DCIS from benign DCIS (i.e., overdiagnosis) or until the associated familial relative risk is known and incorporated, we have to include DCIS as invasive breast cancer in risk prediction models. Accordingly, clinicians need to be aware that by doing so, breast cancer risk in families with DCIS diagnoses, will be probably overestimated.

Another issue that needs to be addressed, is that the PRS is widely validated in the European population, but not for all populations. We have validated the BOADICEA model including the PRS_{313} for the Dutch population. However, we selected for European ancestry while a substantial proportion of the Dutch population is of non-European ancestry. In the Netherlands at least 14% of the population was born themselves outside Western-Europe or one of their parents was born outside Western-Europe (Turkey or a country in Africa, South America or Asia)[25]. This means that we have validated the BOADICEA model for only ~86% of the Dutch population. The lack of ethnic diversity in genetic studies is a known problem. In example, of all included individuals in Genome Wide Association Studies (GWAS) until 2018, 78% are European, 10% are Asian, 2% are African, 1% are Hispanic, and all other ethnicities represent <1%[26]. Because of differences in linkage disequilibrium (LD) across ethnicities, it is uncertain if a causal variant is captured for all populations by the variant identified in GWAS of a single population. Related to this, it is known that some variants may be a risk factor in one population but protective in another population, a phenomenon termed flip-flop[27]. This phenomenon may be due to not targeting the true causal variant.

Table 1. Comparison of PRS performance for predicting overall breast cancer among different ancestries

Reference	PRS	Ancestry	Cases	Controls	OR per SD (95% CI)	AUC (95% CI)
Shieh et al. 2019[28]	180 variants	US Latinas and Latin American women	4,658	7,622	1.58 (1.52-1.64)	0.63 (0.62-0.64)
Ho et al. 2020[29]	287 variants[a]	European	11,225	17,788	1.61 (1.57–1.66)	0.63
		Asian	15,755	16,483	1.52 (1.49–1.56)	0.61
		Asians within North American	1,507	1,212	1.36 (1.25–1.49)	0.58
		Chinese	5,236	5,156	1.58 (1.51–1.65)	0.62 (0.60-0.63)
		Malay	1,084	1,332	1.48 (1.36–1.62)	0.60 (0.58-0.60)
		Indian	580	1,018	1.48 (1.33–1.65)	0.61 (0.59-0.64)
Du et al. 2021[30]	313 variants	African	9,241	10,193	1.27 (1.23-1.31)	0.57 (0.56-0.58)
Liu et al. 2021[31]	209 variants[a]	European	3,960	29,634	1.36 (1.31-1.41)	0.59 (0.58-0.60)
		African	274	3,527	1.15 (1.03-1.30)	0.53 (0.50-0.57)
		Latinx	147	2,049	1.20 (1.01-1.42)	0.53 (0.48-0.58)

[a]out of 313 variants as published by Mavaddat et al.[14]
Abbreviations: AUC, Area Under the Curve; CI, Confidence Interval; PRS, Polygenic Risk Score

Fortunately, there is growing attention for the underrepresentation of ethnically diverse populations in human genetics studies. In recent years, more work has been performed to determine the PRS performance in non-European ethnicities[28-32]. For the Asian population[29] and Latinas[28] the PRS showed similar performance as in the European population, but for the African population[30] there was clearly an attenuated effect size (Table 1). This latter may be due to the large heterogeneity in the African population leading to more variation in LD patterns across the continent[33]. Mapping of the true causal variants may help to obtain a more uniform PRS, useful for different ethnicities. Further research needs to be performed to make optimal use of the PRS for all individuals visiting clinical genetic services. Until more knowledge is gained about the performance of the PRS in women of other ethnicities or ethnicity-specific PRS are available, we have to be cautious when using comprehensive risk prediction including a European ancestry based PRS for these women.

7.3 Contralateral BC

7.3.1. Non-pathogenic variant carriers

In both the family studies described in **chapters 2 and 3** and the population-based studies in **chapters 2 and 4**, the PRS was on average higher for women who developed

a second primary breast tumour compared to women who developed a single breast tumour. These findings suggest an association of the PRS with the development of a second breast tumour which is indeed described in the literature for contralateral breast cancer[6, 34, 35]. However, the effect size of this association was weaker than found for a first breast cancer[35].

7.3.2. *BRCA1/2* pathogenic variant carriers

Previous research showed that the PRS was associated with breast cancer risk in women who carry a pathogenic variant in *BRCA1* or *BRCA2*[36, 37], although with a lower effect size compared to the population[14, 36]. Whether the PRS is associated with contralateral breast cancer risk for *BRCA1/2* pathogenic variant carriers as well, had not been investigated previously. In **chapter 5**, we investigated whether the PRS_{313} is associated with contralateral breast cancer risk among women of European ancestry who carry a pathogenic variant in *BRCA1* or *BRCA2* and explored the implications for contralateral breast cancer risk prediction for these women.

We used retrospective cohort data from carriers of a *BRCA1* or *BRCA2* pathogenic variant participating in the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA)[38] of which we included women of European ancestry with a prevalent first primary invasive breast cancer. We showed significant associations among both *BRCA1* and *BRCA2* pathogenic variant carriers between the PRS and contralateral breast cancer risk. However, as seen for the general population[35], the magnitude of the effect sizes were smaller than previously reported for the first breast cancer[36]. For *BRCA1* pathogenic variant carriers, the largest association was seen with the ER-negative PRS_{313} (HR per SD=1.12), while for *BRCA2* pathogenic variant carriers, both the PRS_{313} and ER-positive PRS_{313} showed similar associations with contralateral breast cancer risk (HR per SD=1.15). These findings are consistent with the higher relative prevalence in this cohort of ER-negative and ER-positive contralateral breast cancers respectively. Although the relative risks of the PRS for contralateral breast cancer were modest, differences in the PRS may still have an important effect on the absolute risk, which is high in BRCA1/2 pathogenic variant carriers. Therefore, the PRS could be used to refine estimates of contralateral breast cancer risks in women who carry a *BRCA1* or *BRCA2* pathogenic variant.

For both *BRCA1* and *BRCA2* pathogenic variant carriers, the strength of the association was greater for ER-positive contralateral breast cancers compared to ER-negative contralateral breast cancers, even if the ER-negative PRS was used for *BRCA1* pathogenic variant carriers. The effect sizes for the PRS are also larger for ER-positive disease in the general population, probably because ER-positive disease is commoner given that >75% of all breast tumours are ER-positive[39]. The same distribution holds for *BRCA2* pathogenic variant carriers as seen in our cohort and described in literature[40]. For *BRCA1* pathogenic variant carriers it is

the opposite, about 75-80% of the tumours are ER-negative[40]. In general, the effect size of the PRS_{313} for developing a first breast cancer[36] and contralateral breast cancer is similar for *BRCA1* and *BRCA2* pathogenic variant carriers. However, our results have shown that the PRS in carriers is mainly associated with ER-positive contralateral tumours and just slightly with ER-negative contralateral tumours. Given this, do we predict contralateral breast cancer risk well enough for *BRCA1* pathogenic variant carriers, or are we predicting only ER-positive contralateral breast cancer? For the first tumour, the ER-negative PRS showed good performance for predicting ER-negative tumours[36], therefore a pragmatic solution for *BRCA1* pathogenic variant carriers is to use the ER-negative PRS for risk prediction of the first tumour. However, this is not yet implemented in breast cancer risk prediction models such as BOADICEA[13]. Another solution would be to predict risks for ER-negative and ER-positive tumours separately. This could also inform clinical management, for example in guiding the choice for chemoprevention in case of an high risk for ER-positive tumour development[41, 42]. With larger datasets, it should be possible to develop better subtype specific PRS for breast cancer and contralateral breast cancer and use this PRS for clinical management choices.

Although the PRS may refine contralateral breast cancer risk estimates for women carrying a pathogenic variant in *BRCA1* or *BRCA2*, the effect size of the PRS seemed to decline with a higher age of first breast cancer diagnosis. For women who were diagnosed with a first tumour after the age of 50, the PRS was of less value for risk prediction for *BRCA2* pathogenic variant carriers and of no value for *BRCA1* pathogenic variant carriers. The decline of the effect size with higher age was seen as well for a first breast cancer for *BRCA1/2* pathogenic variant carriers[36] and for a first breast cancer in the general population[4, 14, 43] including our cohort described in **chapter 4.** However, there was some evidence that the decline in effect size was not linear, given the lower effect size below the age of 40 years described by Mavaddat et al.[14]. This effect was also seen in the population for a contralateral breast tumour[35]. The overall decline with higher age may be caused by a dilution of the effect size due to other risk factors, given that the risk for developing breast cancer in general increases with higher age. For age-dependent breast cancer risk prediction (i.e. 5-year risk or 10-year risk), it is important to take this age-effect into account.

7.4. Future perspectives

This thesis describes the clinical utility of using the PRS for individual breast cancer risk prediction. We have validated the association of the PRS with breast cancer for women in both the Dutch population and breast cancer families and showed a better risk-discrimination by adding the PRS to family-based risk prediction. Although the discrimination accuracy is modest with an AUC<0.70, it is an improvement compared to

family-based risk prediction. Secondly, we have shown that addition of the PRS to family-based risk prediction has an impact on screening recommendations for many non-carriers and carriers of a pathogenic variant in a moderate breast cancer gene. Lastly, there is a prospectively calibrated and externally validated model, BOADICEA, which gained approval as medical device (CE marking) and is implemented in the user-friendly web-interface, the CanRisk tool[44], to calculate breast cancer lifetime risks on the basis of genetic and non-genetic risk factors, including the PRS. The currently ongoing debate whether BOADICEA or other such models (e.g. Tyrer-Cuzick[45]) are good enough for implementation in the clinic and in the population screening setting will no doubt continue for some time; statistical modelling studies have suggested the efficacy of risk-based over age-based screening[46, 47], but these await real-life data from any of the several currently ongoing trials[48, 49] investigating the effect of risk-based screening in (semi-) randomised way.

In my opinion, we are ready for implementation of comprehensive risk prediction in clinical genetic services. However, exactly how to implement this new way of risk prediction has not yet materialised in detail. There remain many issues to be resolved and practicalities to be explored.

First, we have to explore if clinicians are ready to work with comprehensive risk prediction in their consultations. A recent study exploring the acceptability of the CanRisk tool, showed that it was generally acceptable to clinicians, but they were apprehensive about the impact of using this tool on their consultations, which can have impact on the level of implementation[50]. Clinicians are confident with screening advice recommendations based on family history-based risk prediction. As described in this thesis (chapter 2, 3, and 6), for a significant number of women, breast-cancer risks calculated including the PRS in CanRisk will be inconsistent with the risk category and corresponding clinical management advice based on family history only. Before implementation, clinicians must gain confidence in comprehensive risk prediction and corresponding results and we have to explore the effects amongst clinicians regarding their willingness to adjust current advises, especially when screening advices will be downgraded. This latter may also be important for the cost-effectiveness of comprehensive risk prediction. Related to this, we need to explore if comprehensive risk prediction will lead to differences in primary and/or secondary prevention choices by women. Furthermore, comprehensive risk prediction can result in a different screening advice for two family members, for example the two sisters shown in chapter 6. We have to explore the psychosocial effects of personal comprehensive risk prediction if the clinical management advice differs within a family in order to be able to anticipate on these effects. To conclude, before implementation of comprehensive risk prediction we need to know the acceptation of downgrading and different screening advice within families for both clinicians and counselees.

Secondly, due to bias towards European ancestry of Genome Wide Association Studies[26] as described above, the PRS is not yet validated for all ethnicities which may lead to health inequalities[51], resulting in an ethical challenge surrounding implementation of comprehensive risk prediction. Can we offer comprehensive risk prediction to women of European ancestry, if this is not yet possible for all women of non-European ancestries? Ideally, the same care is offered to all women in the population. However, because of other issues to be resolved before implementation, it is possible to start with a small group of women of European ancestry in research-setting to explore the ethical, psychosocial and logistical challenges of implementation. In the meantime, effort has to be made in human genetic research to validate the existing PRS_{313} in other ancestries or to determine ethnicity specific common low risk variants to compute ethnicity specific PRS. In the coming years, the Confluence project will address this by developing a large research resource to uncover breast cancer genetics through genome-wide association studies (GWAS) including cases and controls of different ethnicities[52]. It will be of added value if these results will be implemented in risk prediction models, for example by enabling inclusion of different effect sizes for the PRS.

Currently, if no pathogenic variant is detected in a family, the affected counselee will receive a family letter including the clinical advice for their healthy relatives. A practical issue for implementation of comprehensive risk prediction is that these unaffected relatives need to be referred for counselling for DNA sample collection and risk communication. In addition to the fact that comprehensive risk prediction is still time consuming, this will result in a lot more referrals to clinical genetic services for which we may not have the capacity at this moment. It would be helpful to invest in tools to speed up the process, for example by using pedigree data collection procedures that can be exported into the family tree structures that can be directly uploaded in the CanRisk tool.

Another practical issue is the development of a laboratory test to determine the PRS. This can be performed by direct genotyping each SNP separately, or by using a SNP array with additional imputation of the missing variants. Direct genotyping is technically easier, more efficient, and an advantage is the high reliability of the PRS calculation. Therefore, at this moment, laboratories prefer direct genotyping. However, in my opinion, using a genome-wide SNP-array and imputation has advantages that need to be seriously considered. A SNP array will be more future proof and widely applicable, given the possibility to calculate all kinds of PRS, not just those currently known for breast cancer and European ancestry. For example, it is to be expected that we will have a more extensive PRS for breast cancer in the future, knowing that the current PRS explains about half of the estimated part of the familial relative risk that could be explained by common low risk variants[14] and that recent studies already discovered 38 novel breast cancer susceptibility loci at genome wide significance level[53, 54]. Furthermore, although this is not yet implemented in the

CanRisk tool either, it is possible that we need to use ethnicity-specific PRS rather than adjustment of the weights associated with each variant of the PRS_{313}. In addition, because it may be difficult to define ancestry from non-genetic data (e.g. pedigree or anamnestic information), ancestry can be determined fairly accurately with array data. Finally, a sufficiently dense SNP array can also support the many other PRSs that have been defined today for many other common diseases, such as coronary arterial disease. In summary, direct genotyping might be favoured technically, but it is possible that we have to design multiple genetic test for all different PRS and need to estimate ancestry with non-genetic information.

While there are many challenges still to overcome, we can start in research-setting with implementation of individual comprehensive breast cancer risk prediction including the PRS for women visiting clinical genetic services and their healthy relatives. Hopefully the studies described in this thesis contribute to the first steps towards implementation of comprehensive risk prediction to all women in clinical setting and in the future for the population screening as well.

References

1. Cuzick J, Brentnall AR, Segal C, et al. Impact of a Panel of 88 Single Nucleotide Polymorphisms on the Risk of Breast Cancer in High-Risk Women: Results From Two Randomized Tamoxifen Prevention Trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. Mar 2017;35(7):743-750. doi:10.1200/jco.2016.69.8944
2. Dite GS, MacInnis RJ, Bickerstaffe A, et al. Breast Cancer Risk Prediction Using Clinical Models and 77 Independent Risk-Associated SNPs for Women Aged Under 50 Years: Australian Breast Cancer Family Registry. *Cancer EpidemiolBiomarkers Prev*. 12/16/2015 2015;Not in File. doi:1055-9965.EPI-15-0838 [pii];10.1158/1055-9965.EPI-15-0838 [doi]
3. Shieh Y, Hu D, Ma L, et al. Breast cancer risk prediction using a clinical risk model and polygenic risk score. *Breast cancer research and treatment*. Oct 2016;159(3):513-25. doi:10.1007/s10549-016-3953-2
4. Mavaddat N, Pharoah PD, Michailidou K, et al. Prediction of breast cancer risk based on profiling with common genetic variants. *JNatlCancer Inst*. 5/2015 2015;107(5)Not in File. doi:djv036 [pii];10.1093/jnci/djv036 [doi]
5. van Veen EM, Brentnall AR, Byers H, et al. Use of Single-Nucleotide Polymorphisms and Mammographic Density Plus Classic Risk Factors for Breast Cancer Risk Prediction. *JAMA oncology*. Apr 1 2018;4(4):476-482. doi:10.1001/jamaoncol.2017.4881
6. Sawyer S, Mitchell G, McKinley J, et al. A role for common genomic variants in the assessment of familial breast cancer. *JClinOncol*. 12/10/2012 2012;30(35):4330-4336. Not in File. doi:JCO.2012.41.7469 [pii];10.1200/JCO.2012.41.7469 [doi]
7. Muranen TA, Mavaddat N, Khan S, et al. Polygenic risk score is associated with increased disease risk in 52 Finnish breast cancer families. *Breast cancer research and treatment*. Aug 2016;158(3):463-9. doi:10.1007/s10549-016-3897-6
8. Li H, Feng B, Miron A, et al. Breast cancer risk prediction using a polygenic risk score in the familial setting: a prospective study from the Breast Cancer Family Registry and kConFab. *Genetics in medicine : official journal of the American College of Medical Genetics*. May 12 2016;doi:10.1038/gim.2016.43
9. Evans DG, Brentnall A, Byers H, et al. The impact of a panel of 18 SNPs on breast cancer risk in women attending a UK familial screening clinic: a case-control study. *Journal of medical genetics*. Feb 2017;54(2):111-113. doi:10.1136/jmedgenet-2016-104125
10. Michailidou K, Lindstrom S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. Oct 23 2017;doi:10.1038/nature24284
11. Lee AJ, Cunningham AP, Tischkowitz M, et al. Incorporating truncating variants in PALB2, CHEK2, and ATM into the BOADICEA breast cancer risk model. *Genetics in medicine : official journal of the American College of Medical Genetics*. Dec 2016;18(12):1190-1198. doi:10.1038/gim.2016.31
12. IKNL. Richtlijn Borstkanker - Screening buiten het bevolkingsonderzoek. Accessed 03-12-2021, https://richtlijnendatabase.nl/richtlijn/borstkanker/screening/screening_buiten_het_bob/screening_buiten_het_bevolkingsonderzoek.html

13. Lee A, Mavaddat N, Wilcox AN, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and nongenetic risk factors. *Genetics in medicine : official journal of the American College of Medical Genetics*. Jan 15 2019;21(8):1708-1718. doi:10.1038/s41436-018-0406-9
14. Mavaddat N, Michailidou K, Dennis J, et al. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *American journal of human genetics*. Jan 3 2019;104(1):21-34. doi:10.1016/j.ajhg.2018.11.002
15. Terry MB, Liao Y, Whittemore AS, et al. 10-year performance of four models of breast cancer risk: a validation study. *The Lancet Oncology*. Apr 2019;20(4):504-517. doi:10.1016/s1470-2045(18)30902-1
16. Pal Choudhury P, Brook MN, Hurson AN, et al. Comparative validation of the BOADICEA and Tyrer-Cuzick breast cancer risk models incorporating classical risk factors and polygenic risk in a population-based prospective cohort of women of European ancestry. *Breast Cancer Research*. 2021/02/15 2021;23(1):22. doi:10.1186/s13058-021-01399-7
17. Li SX, Milne RL, Nguyen-Dumont T, et al. Prospective Evaluation over 15 Years of Six Breast Cancer Risk Models. *Cancers*. 2021;13(20):5194.
18. Naslund-Koch C, Nordestgaard BG, Bojesen SE. Common breast cancer risk alleles and risk assessment: A study on 35,441 individuals from the Danish general population. *Annals of oncology : official journal of the European Society for Medical Oncology*. Oct 13 2016;doi:10.1093/annonc/mdw536
19. Maas P, Barrdahl M, Joshi AD, et al. Breast Cancer Risk From Modifiable and Nonmodifiable Risk Factors Among White Women in the United States. *JAMA oncology*. Oct 1 2016;2(10):1295-1302. doi:10.1001/jamaoncol.2016.1025
20. Petridis C, Brook MN, Shah V, et al. Genetic predisposition to ductal carcinoma in situ of the breast. *Breast cancer research : BCR*. Feb 17 2016;18(1):22. doi:10.1186/s13058-016-0675-7
21. Visser LL, Groen EJ, van Leeuwen FE, Lips EH, Schmidt MK, Wesseling J. Predictors of an Invasive Breast Cancer Recurrence after DCIS: A Systematic Review and Meta-analyses. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. May 2019;28(5):835-845. doi:10.1158/1055-9965.Epi-18-0976
22. Groen EJ, Hudecek J, Mulder L, et al. Prognostic value of histopathological DCIS features in a large-scale international interrater reliability study. *Breast cancer research and treatment*. Oct 2020;183(3):759-770. doi:10.1007/s10549-020-05816-x
23. Coleman WB. Breast Ductal Carcinoma in Situ: Precursor to Invasive Breast Cancer. *Am J Pathol*. May 2019;189(5):942-945. doi:10.1016/j.ajpath.2019.03.002
24. van Seijen M, Lips EH, Thompson AM, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. *British journal of cancer*. Aug 2019;121(4):285-292. doi:10.1038/s41416-019-0478-6
25. CBS. Hoeveel mensen met een migratieachtergrond wonen in Nederland? Accessed 07-12-2021, https://www.cbs.nl/nl-nl/dossier/dossier-asiel-migratie-en-integratie/hoeveel-mensen-met-een-migratieachtergrond-wonen-in-nederland-#:~:text=Van%20de%20totale%20Nederlandse%20bevolking,daarmee%20tot%20de%20tweede%20generatie.

26. Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic Studies. *Cell*. Mar 21 2019;177(1):26-31. doi:10.1016/j.cell.2019.02.048
27. Wang S, Qian F, Zheng Y, et al. Genetic variants demonstrating flip-flop phenomenon and breast cancer risk prediction among women of African ancestry. *Breast cancer research and treatment*. Apr 2018;168(3):703-712. doi:10.1007/s10549-017-4638-1
28. Shieh Y, Fejerman L, Lott PC, et al. A polygenic risk score for breast cancer in U.S. Latinas and Latin-American women. *Journal of the National Cancer Institute*. Sep 25 2019;doi:10.1093/jnci/djz174
29. Ho WK, Tan MM, Mavaddat N, et al. European polygenic risk score for prediction of breast cancer shows similar performance in Asian women. *Nat Commun*. Jul 31 2020;11(1):3833. doi:10.1038/s41467-020-17680-w
30. Du Z, Gao G, Adedokun B, et al. Evaluating Polygenic Risk Scores for Breast Cancer in Women of African Ancestry. *Journal of the National Cancer Institute*. Mar 26 2021;doi:10.1093/jnci/djab050
31. Liu C, Zeinomar N, Chung WK, et al. Generalizability of Polygenic Risk Scores for Breast Cancer Among Women With European, African, and Latinx Ancestry. *JAMA Netw Open*. Aug 2 2021;4(8):e2119084. doi:10.1001/jamanetworkopen.2021.19084
32. Evans DG, van Veen EM, Byers H, et al. The importance of ethnicity: Are breast cancer polygenic risk scores ready for women who are not of White European origin? *International journal of cancer*. Jan 1 2022;150(1):73-79. doi:10.1002/ijc.33782
33. Tishkoff SA, Reed FA, Friedlaender FR, et al. The genetic structure and history of Africans and African Americans. *Science (New York, NY)*. May 22 2009;324(5930):1035-44. doi:10.1126/science.1172257
34. Robson ME, Reiner AS, Brooks JD, et al. Association of Common Genetic Variants With Contralateral Breast Cancer Risk in the WECARE Study. *Journal of the National Cancer Institute*. Oct 1 2017;109(10)doi:10.1093/jnci/djx051
35. Kramer I, Hooning MJ, Mavaddat N, et al. Breast Cancer Polygenic Risk Score and Contralateral Breast Cancer Risk. *American journal of human genetics*. Nov 5 2020;107(5):837-848. doi:10.1016/j.ajhg.2020.09.001
36. Barnes DR, Rookus MA, McGuffog L, et al. Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. *Genetics in medicine : official journal of the American College of Medical Genetics*. Oct 2020;22(10):1653-1666. doi:10.1038/s41436-020-0862-x
37. Kuchenbaecker KB, McGuffog L, Barrowdale D, et al. Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. *Journal of the National Cancer Institute*. Jul 01 2017;109(7)doi:10.1093/jnci/djw302
38. Chenevix-Trench G, Milne RL, Antoniou AC, Couch FJ, Easton DF, Goldgar DE. An international initiative to identify genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA). *Breast cancer research : BCR*. 2007;9(2):104. doi:10.1186/bcr1670

39. Tsang JYS, Tse GM. Molecular Classification of Breast Cancer. *Adv Anat Pathol*. Jan 2020;27(1):27-35. doi:10.1097/pap.0000000000000232
40. Spurdle AB, Couch FJ, Parsons MT, et al. Refined histopathological predictors of BRCA1 and BRCA2 mutation status: a large-scale analysis of breast cancer characteristics from the BCAC, CIMBA, and ENIGMA consortia. *Breast cancer research : BCR*. Dec 23 2014;16(6):3419. doi:10.1186/s13058-014-0474-y
41. Nelson HD, Fu R, Griffin JC, Nygren P, Smith ME, Humphrey L. Systematic review: comparative effectiveness of medications to reduce risk for primary breast cancer. *Annals of internal medicine*. Nov 17 2009;151(10):703-15, w-226-35. doi:10.7326/0003-4819-151-10-200911170-00147
42. Li K, Anderson G, Viallon V, et al. Risk prediction for estrogen receptor-specific breast cancers in two large prospective cohorts. *Breast cancer research : BCR*. Dec 3 2018;20(1):147. doi:10.1186/s13058-018-1073-0
43. Gao C, Polley EC, Hart SN, et al. Risk of Breast Cancer Among Carriers of Pathogenic Variants in Breast Cancer Predisposition Genes Varies by Polygenic Risk Score. *Journal of Clinical Oncology*. 0(0):JCO.20.01992. doi:10.1200/jco.20.01992
44. Carver T, Hartley S, Lee A, et al. CanRisk Tool—A Web Interface for the Prediction of Breast and Ovarian Cancer Risk and the Likelihood of Carrying Genetic Pathogenic Variants. *Cancer Epidemiology Biomarkers & Prevention*. 2021;30(3):469-473. doi:10.1158/1055-9965.Epi-20-1319
45. Kim G, Bahl M. Assessing Risk of Breast Cancer: A Review of Risk Prediction Models. *J Breast Imaging*. Mar-Apr 2021;3(2):144-155. doi:10.1093/jbi/wbab001
46. van den Broek JJ, Schechter CB, van Ravesteyn NT, et al. Personalizing Breast Cancer Screening Based on Polygenic Risk and Family History. *Journal of the National Cancer Institute*. Apr 6 2021;113(4):434-442. doi:10.1093/jnci/djaa127
47. Pashayan N, Morris S, Gilbert FJ, Pharoah PDP. Cost-effectiveness and Benefit-to-Harm Ratio of Risk-Stratified Screening for Breast Cancer: A Life-Table Model. *JAMA oncology*. Nov 1 2018;4(11):1504-1510. doi:10.1001/jamaoncol.2018.1901
48. Clift AK, Dodwell D, Lord S, et al. The current status of risk-stratified breast screening. *British journal of cancer*. Oct 26 2021;doi:10.1038/s41416-021-01550-3
49. Pashayan N, Antoniou AC, Ivanus U, et al. Personalized early detection and prevention of breast cancer: ENVISION consensus statement. *Nat Rev Clin Oncol*. Nov 2020;17(11):687-705. doi:10.1038/s41571-020-0388-9
50. Archer S, Babb de Villiers C, Scheibl F, et al. Evaluating clinician acceptability of the prototype CanRisk tool for predicting risk of breast and ovarian cancer: A multi-methods study. *PloS one*. 2020;15(3):e0229999. doi:10.1371/journal.pone.0229999
51. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature genetics*. Apr 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x
52. Confluence project. Accessed 18-01-2022, 2022. https://dceg.cancer.gov/research/cancer-types/breast-cancer/confluence-project

53. Zhang H, Ahearn TU, Lecarpentier J, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nature genetics*. Jun 2020;52(6):572-581. doi:10.1038/s41588-020-0609-2
54. Adedokun B, Du Z, Gao G, et al. Cross-ancestry GWAS meta-analysis identifies six breast cancer loci in African and European ancestry women. *Nat Commun*. Jul 7 2021;12(1):4198. doi:10.1038/s41467-021-24327-x