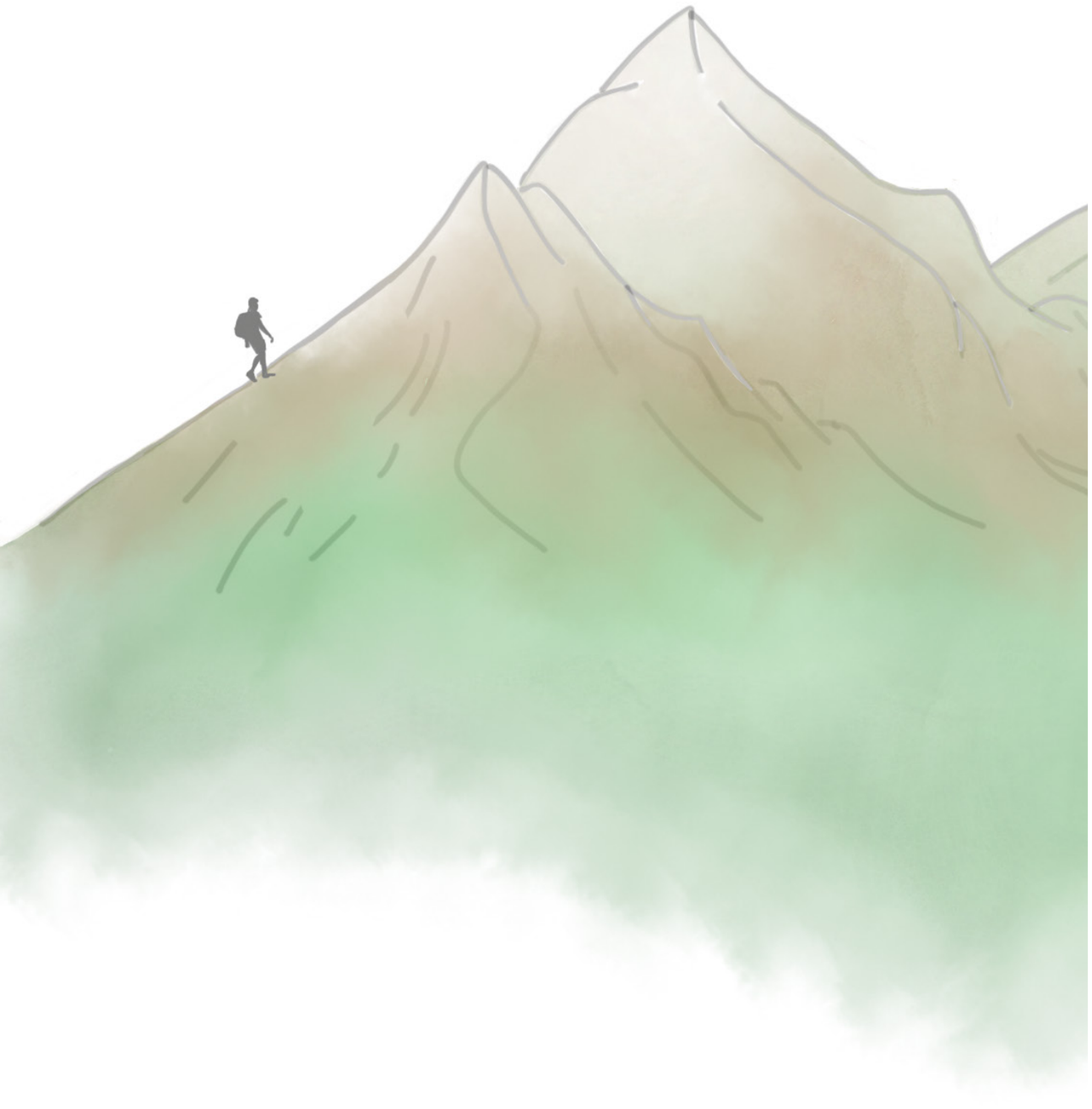
Prediction of contralateral breast cancer: statistical aspects and prediction performance

Giardiello, D.



Chapter 8

General discussion

DISCUSSION

The aim of this thesis was to develop and validate a risk prediction model of contralateral breast cancer (CBC) for women with a first invasive breast cancer and to provide guidelines about performance assessment of risk prediction models with time-to-event outcomes. Large international population-based and hospital-based studies were analyzed to build the CBC prediction models, called PredictCBC models. In addition, the risk of CBC in women diagnosed with ductal carcinoma in situ (DCIS) was investigated. In this chapter, we will discuss the main findings and interpret them in a broader context with particular attention to the potential practical implications in clinical decision making. The methodological challenges of developing and validating a risk prediction model for time-to-event outcomes with and without competing risks are discussed in the CBC risk prediction context, and, more generally, using examples in the context of breast cancer recurrence and survival. Finally, future perspectives of research into prediction of CBC are also given. Future potential research directions in medical statistics are suggested based on methodological gaps that became apparent when analyzing real world breast cancer data.

Main findings and potential clinical implications

Risk factors and risk prediction of CBC in patients with first invasive breast cancer

Accurate CBC risk predictions are essential in clinical decision making regarding contralateral preventive mastectomies (CPMs) or other preventive strategies as personalized treatments and individualized surveillance. A number of patient-, first primary breast cancer-, and treatment characteristics have been suggested to be associated with CBC in several studies in the last 30 years[1-4]. In particular, patients' characteristics such as age at first breast cancer diagnosis, family history of breast cancer, hereditary mutations in the *BRCA1, BRCA2* and *CHEK2* (particularly the *c.1100delC* mutation) genes, as well as specific first primary breast cancer characteristics, i.e., tumor size, lymph node status and breast cancer histology, and (neo)adjuvant systemic therapies[5-11]. Furthermore, less attention has been given to study the association between CBC and lifestyle and reproductive factors. Currently, only a few factors such as body mass index (BMI) and parity (i.e., the number of births preceding first primary breast cancer diagnosis) have been suggested to be associated with CBC[12]. Could these characteristics (in statistical glossary labeled as predictors or covariates) be used to provide accurate CBC risk predictions and therewith support clinical decision making towards the choice of preventive strategies?

To answer this question, we developed and validated CBC risk prediction models (PredictCBC) using large population- and hospital-based studies mostly based in Europe, the United States and Australia with long follow-up (**Chapter 2**). The choice of the predictors in the analyses was based on evidence from the literature, availability of predictors in the studies and experience from clinical practice. PredictCBC models provided a moderate CBC prediction accuracy in terms of discrimination and calibration. These results were in line with other prediction tools currently available for first primary breast cancer and, generally speaking, in the oncology field[13]. Clinical decision making about CPM may be improved using PredictCBC models compared to current clinical practice in the Netherlands, which is mostly based on carriership of germline mutations in the *BRCA1/2* genes. We showed an overlap in the magnitude of CBC risk between *BRCA1/2* mutation carriers and non-*BRCA1/2* carriers. In fact, CPM might not be the preferred choice even in some patients with a *BRCA1/2* germline mutation when other characteristics are highly favorable. On the other hand, some additional preventive strategies might be considered among non-*BRCA1/2* carriers with unfavorable characteristics. We conclude that clinical decision making about preventive strategies should not only be based on a germline mutation in *BRCA1/2* genes or a bilateral breast cancer family history, but the multifactorial context should be considered using, for example, PredictCBC models. However, although PredictCBC models may more objectively estimate CBC risk, the decision making strongly depends on what patients and physicians consider an acceptable risk according to the guidelines and on patients' personal preferences.

Previously, two other tools were proposed to predict CBC: the Manchester formula and CBCrisk[14,15]. The former is a heuristic formula based on a literature review, while the latter is a CBC risk prediction model that was developed and validated in the United States around the same time we developed and validated the PredictCBC models[15,16]. Therefore, we decided to investigate the prediction accuracy of these two tools in comparison with PredictCBC models using the large population- and hospital-based studies we used to develop PredictCBC models (**Chapter 3**). We found that all three CBC tools provided only moderate prediction accuracy. We also found considerable heterogeneity among studies[17].

Other breast cancer risk genes, beyond *BRCA1/2,* have also been shown to be associated with CBC risk[10,18-21]. Can some of these additional genetic markers improve CBC risk prediction and clinical decision making? In **chapter 4,** we updated the PredictCBC models to PredictCBC-2.0 models. We incorporated data on a specific rare mutation in the *CHEK2* gene (*c.1100delC*), a polygenic risk score (PRS) combining 313 common genetic variants associated with breast cancer, and the previously shown relevant factors body mass index and parity[10,12,22]. We showed that the overall CBC prediction accuracy did not substantially increase between PredictCBC and PredictCBC-2.0 models. On the other hand, we demonstrated that the PredictCBC-2.0 model including *CHEK2 c.1100delC* and PRS had higher net benefit compared to the previous PredictCBC models. In other words, clinical decision making might additionally improve and be better tailored

8

incorporating common and specific rare genetic variants associated with CBC, especially among patients with *BRCA1/2* germline mutations and non-*BRCA1/2* carriers.

CBC risk in patients diagnosed with ductal carcinoma in situ

One of the most active research lines among physicians studying ductal carcinoma in situ (DCIS, a potential precursor of cancer) is whether a patient diagnosed with DCIS may develop a subsequent ipsilateral invasive breast cancer in the future[23]. However, it may be equally important to estimate the risk to develop a CBC in patients diagnosed with DCIS. In that light, the comparison of this risk to that of patients diagnosed with a first invasive breast cancer is relevant to help understand the magnitude of risk, the etiology, and treatment strategies. We showed that the CBC risk is slightly higher in patients with DCIS compared to invasive breast cancer patients using a large population-based dataset from 1989 to 2017 of the Dutch Cancer Registry, covering all breast cancer patients diagnosed in the Netherlands (**Chapter 5**). Around five out of 100 patients with DCIS may develop a CBC compared to around four patients with invasive breast cancer within 10 years. We illustrated that this slightly higher CBC risk in DCIS patients might be largely explained by the fact that adjuvant systemic therapies are not currently considered in DCIS patients according to the current Dutch guidelines. This does not imply that physicians should start treating DCIS patients with adjuvant systemic therapy since CBC risk is low and side effects of adjuvant systemic therapies have been demonstrated[24,25]. However, especially in the United States, more DCIS patients ask to undergo a CPM as a consequence of CBC risk overestimation[25]. We concluded that accurate prediction may be also needed for DCIS patients to facilitate decision making about additional treatments or CPM (**Chapter 5**).

Assessing the performance of survival and competing risks prediction models

In **chapter 6 and 7,** we propose frameworks for performance evaluation of predictions and for clinical utility of survival and competing risks models to provide guidance in the context of the STRATOS (STRengthening Analytical Thinking for Observational Studies) international initiative[26]. The objective of STRATOS is to provide accessible and guidance in the design and analysis of observational studies, since the quality of parts of biomedical research urgently needs improvement [26]. The members of the STRATOS initiative are experienced statisticians with different expertise in different topic groups (TG) (https://www.stratos-initiative.org/groups). Two topic groups (TG6 and TG8) involve statisticians with expertise in prediction models and survival analysis, respectively. In collaboration with the members of these two TGs, we provide guidance for different traditional and novel measures that may be used to assess the performance of prediction for survival and competing risk models. Typically, specific time points (also defined time horizons) are chosen as relevant by physicians. For example, physicians may be interested in predicting CBC risk at 5 and 10 years since the first invasive breast cancer diagnosis (**chapter 2, 3 and 4**). For this purpose, several time-dependent discrimination and calibration performance measures were proposed in the literature[27,28]. In **chapter 6 and 7**, we briefly provided an overview of the measures currently available in the literature. Secondly, we suggested to prioritize some of the available measures to evaluate discrimination, calibration and clinical utility of survival and competing risk models according to the literature and software availability. We aimed to guide practitioners and researchers interested in prediction models providing data and the software code. In **chapter 6** we provided both R and SAS code to develop and validate a survival risk prediction model using two free available data sets: the German Breast Cancer Study Group and the Rotterdam breast cancer dataset[29,30]. In **chapter 7**, we used a random sample of 1,000 patients from the Female breast cancer in the elderly; Optimizing Clinical guidelines USing clinico-pathological & molecular Study (FOCUS) and from NCR data to develop and validate a risk prediction model of breast cancer recurrence in the presence of competing risk due to mortality[31]. Both R and SAS codes are freely available in GitHub repositories created and maintained by the author (https://github.com/danielegiardiello/Prediction_performance_survival and https://github.com/survival-lumc/ValidationCompRisks) to facilitate the connection between methodological developments and software availability.

Strengths and limitations of the data used

One of the most important strengths of the studies presented in this thesis is the use of large hospital- and population-based studies with follow-up information of all women diagnosed with invasive breast cancer between 1990 and 2017. Most of the studies were from the Netherlands: the Netherlands Cancer Registry (NCR) has good quality information about patients, first primary breast cancer characteristics, and treatments[32]. We were fortunate to include other studies from the Netherlands such as Amsterdam Breast Cancer Study (ABCS), Breast Cancer Outcome Study of Mutation carriers (BOSOM), Erasmus Medical Center (EMC) study, and Hereditary Breast and Ovarian cancer study (HEBON)[33,34]. Their contribution was essential to incorporate key information about germline mutations in the *BRCA1/2* genes and on the performance of CPM. Last but not least, we incorporated data from the Breast Cancer Association Consortium (BCAC) including studies from other European countries, the United States and Australia. Using BCAC, we included other potentially important information, i.e., the specific rare mutation in the *CHEK2* gene (*c.1100delC*) and the PRS which was developed using BCAC data in a previous study[35,36]. The studies included in this thesis, after combining and harmonizing different sources of data, comprised more than 100,000 women diagnosed with invasive breast cancer or DCIS.

One of the most challenging parts of using different studies was missing data. In **chapter 2, 3 and 4,** the NCR represented more than 60% of all data to develop and

validate PredictCBC and PredictCBC-2.0 models. Family history for breast cancer, germline genetic information (i.e., *BRCA1/2* germline mutation, *CHEK2* c.1100delC and PRS), and CPM are completely unavailable in NCR. However, complete breast cancer characteristics and treatment information available in NCR contributed to developing good performance imputation models based on the correlation matrix of the data[37]. In addition, the remaining predictors were quite complete: more than 70% of patients had at most one missing predictor. Follow-up information regarding some outcomes were incomplete in some studies. For example, in some studies included in **chapter 2, 3 and 4**, CBC and CPM outcome information was incomplete leading to an underestimation of the cumulative incidence. This challenge can only be solved in the future by improvement of data collection at the source (i.e., the original registry from which the studies acquired their data).

Methodological challenges

There are several challenges in methodological research of clinical prediction models when individual data from different studies are available. The most important challenges characterizing all applied works (**chapter 2, 3 and 4**) of this thesis were missing data, heterogeneity between studies, and time-to-event outcomes with a special attention to competing risks.

Missing data are unavoidable in medical research and researchers tend to include only complete information to perform the statistical analyses. However, excluding a large proportion of information will lead to biases[38]. Biases may be substantial or negligible according to the reasons why data are missing. A common classification of missing data is: missing completely at random, missing at random and missing not at random[39]. According to the type of missing data classification, different statistical methods are suggested. These methods, commonly defined as multiple imputation, replace missing values with imputed values. To allow uncertainty about the missing data, imputed values are generated multiple times to create several different plausible imputed data. These results of all imputed data are combined (using different methods) to fully consider uncertainty among imputations. Multiple imputation methods are typically suggested when a missing is at random. When missing is completely at random, complete case analysis is suggested. On the other hand, when the amount of missing values is high, multiple imputation is recommended to avoid data reduction. Additional statistical investigations (e.g. sensitivity analyses) are required when missing is not at random[39]. An overview including the definition of the different types of missing data classification with an example and the suggested statistical method to minimize biased point and variance estimates is shown in **Table 1**.

Table 1: summary of missing data definitions and proposed solutions

Missing data mechanism	Definition	Example	Suggested solution
Missing completely at random	No systematic differences between the missing and observed values	Missing values on a certain predictor occur randomly in the data	Complete case analysis*
Missing at random	Missing values can be explained by differences in observed data	Patients with missing values of *BRCA1/2* tend to be older and with less aggressive tumors.	Multiple imputation
Missing not at random	Differences between missing and observed values may be still present after considering observed data	Patients with missing BMI may tend to have too high or too low self-reported BMI.	Multiple imputation and sensitivity analyses

* multiple imputation may be preferred especially when the amount of missing values is high.

In the context of individual data from multiple studies, missing values may be also systematic. Systematic missing data occurs when a predictor is completely unavailable in one or more studies. For example, genetic information is systematically missing in registry-based data as NCR. In this case, missing data can be considered at random since the missing can be fully explained by the fact that some studies simply did not collect this information. A potential solution for systematic missing values is to use the variable identifying the study as covariate to improve substantially the imputation models. More sophisticated multiple imputation approaches were proposed in the literature using, for example, mixed-effects imputation models to better consider heterogeneity between studies and the hierarchical nature of data[40-42]. Less is known about how to include competing risks in multiple imputation when individual data from multiple studies are available and in presence of systematic missing values[43,44].

Between-study heterogeneity should be also adequately considered both in the analysis, development and in validation of a risk prediction model using multiple studies. The heterogeneity may refer to different baseline risk for different studies, different distribution of the predictors among studies and/or different methods to measure outcomes and predictors. A risk prediction model can be developed using one-stage or a two-stage approach[45]. In one-stage individual patient data, a single model is developed and typically mixed-effects multilevel regression is used to consider within and between studies heterogeneity[42,46-48]. In two-stage individual patient data analysis, in the first step simple regression models are performed by study. Secondly, the estimates are combined using meta-analytic methods[42,47]. Both approaches have advantages and challenges, although some simulations showed a fully specified one stage approach should be preferred, especially in presence of systematic missing data[42,45]. However, few examples and guidelines are available in the literature with time-to-event outcomes[49-53].

8

In case of survival analysis, stratified Cox regression models or flexible parametric survival regression proposed by Royston and Parmar models may be used to account for different baseline risks among studies in one stage individual patient data analysis[53-55]. More sophisticated survival regression models were proposed (e.g. frailty models) in the literature[50]. Currently, to the best of our knowledge, no clear guidelines are available to clarify how to analyze individual patient data using multiple studies in the presence of competing risks outcomes. One of the first questions is whether Fine and Gray or cause-specific hazards models should be used, especially when the aim is to develop and validate a risk prediction model to predict the absolute risk[56,57]. In **chapters 2 and 4,** we developed PredictCBC models with a one stage individual patient data analysis approach using a stratified Fine and Gray method after multiple imputation of missing values. More practical and methodological efforts in the context of competing risks might be useful to better consider heterogeneity in the imputation and analysis models. **Table 2** summarizes the approaches of analyzing individual patient data using multiple studies.

Table 2: a summary of approaches to analyze individual patient data from multiple studies

Approach	Definition	Pros'	Cons'	Potential future developments
One-stage	A single model is developed where heterogeneity should be considered	- Simple; - Consistent with imputation of systematic missing data	- Sophisticated models (e.g. stratification, mixed-models, frailty models) - Computationally demanding	- Clear guidelines about how to develop, validate a competing risks / dynamic prediction models
Two-stage	A single model is developed by study. Estimates are combined using meta-analytical approaches	- Reasonable to fully consider heterogeneity among studies	- In case of systematic missing values, study-specific estimates may be diluted with multiple imputation	

IMPLICATIONS AND FURTHER RESEARCH

Potential clinical implications of PredictCBC models

With the work described in this thesis we have tried to pave a way for more accurate and potentially clinically relevant CBC risk prediction through PredictCBC models. Currently, decision making about CBC preventive strategies is essentially based on *BRCA1/2* germline mutation and/or family history. This choice is still reasonable and practical, although additional clinical and genetic information can refine clinical decision making about CPM. No clear guidelines are currently available about risk management of CBC. For example, in the Netherlands, most of them are based on the risk management for a first primary breast cancer diagnosis[58]. In some circumstances, the BOADICEA risk prediction model, originally developed to predict the risk of first primary breast cancer, is used to have a better idea about the CBC risk[59,60]. However, the risk to develop a CBC is higher among first breast cancer patients compared to the risk to develop a first primary breast cancer among healthy women[61]. Furthermore, the current BOADICEA model does not include crucial additional information from the primary breast cancer, most important being systemic therapies. Last but not least, currently although CBC prediction tools are available like PredictCBC models and CBCrisk, these are not widely used in clinical practice[62]. Generally speaking, PredictCBC models and other CBC risk tools may be used to better identify high risk patients and reassure low risk patients who have worries and fears about the risk to develop a new primary tumor in the opposite breast. An appropriate use of the models may help patients, with a low predicted CBC risk, to avoid CPM or to opt for an alternative preventive strategy such as personalized screening programs.

Suggestions to promote the usage of PredictCBC models

Three fundamental points may encourage physicians to use PredictCBC models in clinical practice: model implementation, model validation and model updating. Model implementation is essential to support public health strategies. One of the most important goals of model implementation is to provide a user-friendly tool to efficiently communicate risks to patients. A well-implemented PredictCBC model may facilitate a more interactive discussion about CBC risk management between patients and physicians. In a parallel PhD trajectory based on the same project, we took a first step towards this implementation. In addition, misconception of the risk may be minimized since breast cancer patients generally tend to overestimate their CBC risk[63,64]. Online website and software applications are largely (and also freely) available nowadays to implement and periodically update risk prediction tools in practice, for example Shiny in R (https://shiny.rstudio.com/) and Evidencio (https://www.evidencio.com/home).

Model validation is important to evaluate a risk prediction in a setting different than the one used to develop the model. PredictCBC models were built using studies from Europe (most of them from the Netherlands), United States, and Australia. We strongly encourage to validate PredictCBC models especially in Asian and African studies to refine CBC risk prediction and to introduce new or updated CBC risk management guidelines in different countries.

PredictCBC models, and generally prediction models, should be periodically monitored over time to provide up-to-date prediction and performances. However, it is still challenging to establish when and how often a periodic surveillance of the risk prediction

performances should be provided, especially when healthcare policies are quite heterogeneous among countries and change over the time[65]. Continuous monitoring and updates of prediction models are expensive and computationally demanding. Centralization, harmonization, and standardization of health care data may represent one of the ongoing/new frontiers in (medical) statistics. Recent and future developments in the field of data engineering, data architecture, and data science may substantially accelerate the usage of more sophisticated electronic health records to answer etiological questions and to develop, validate and monitor risk prediction models.

Further research and future developments in CBC risk prediction

Risk prediction models may need updates and revision. PredictCBC models might be in the future updated as new predictors become available. There are still stimulating opportunities to improve CBC risk prediction performance. For example, polygenic risk scores based on common genetic variants may steadily improve as new biological insights become available. Although the 313-polygenic risk score and *CHEK2 c.1100delC* are currently unlikely to add substantial improvements of CBC risk prediction performance in the general population, better tailored clinical decision making for individual patients was certainly apparent in PredictCBC-2.0 models. Other germline variants in *CHEK2*, and also in the *ATM* and *PALB2* genes are suggested to be associated with higher first breast cancer and CBC risk[9,20,66]. Breast density is a well-established risk factor of first primary breast cancer and it has been suggested to be associated with an increased CBC risk[67,68]. BOADICEA and the CBC specific risk tool CBCrisk include breast density as a predictor; however, no clinical utility evaluation was provided yet. Further research is needed to investigate whether including information about *ATM*, *PALB2* and breast density may improve CBC prediction and decision making. All potential aforementioned predictors are measured at (around) the diagnosis of primary breast cancer. Breast density, lifestyle and reproductive factors may change over time and adequate statistical methods are needed to consider time-dependent predictors to estimate CBC risk over the time.

Conceptually, any kind of risk prediction is challenging, especially far away in the future. It is reasonable to think that risk predictions may improve as new and updated information becomes available close to the prediction time horizon. Breast cancer is a multi-state disease with clinically relevant intermediate outcomes such as, for example, recurrence (local, locoregional, distant) and CBC. Individual patient prognosis can really differ as the intermediate events occurs or information about modifiable risk factors (e.g., BMI or alcohol use) and biomarkers change after the first event (e.g., diagnosis of DCIS or first primary invasive breast cancer). Recently new methods have been developed to simultaneously model intermediate states and incorporate longitudinal data[69-74]. The multi-state and dynamic prediction modeling can be used to reveal the relations between different types of events and to estimate predictions. Prediction can be calculated considering patients' personal characteristics and clinical status prior to the time of prediction (time horizon) at specific landmark time after the first event (e.g., IBC or DCIS) or at the time of an intermediate event (e.g., CBC) occurs. We propose a graphical representation of a multi-state modelling for dynamic prediction, encapsulating the outline of this thesis, as a potential next step in breast cancer prediction. (**Figure 1**). Unfortunately, we were not able to implement any dynamic models because information of intermediate events was incomplete and most of the predictors were not collected over the follow-up time.

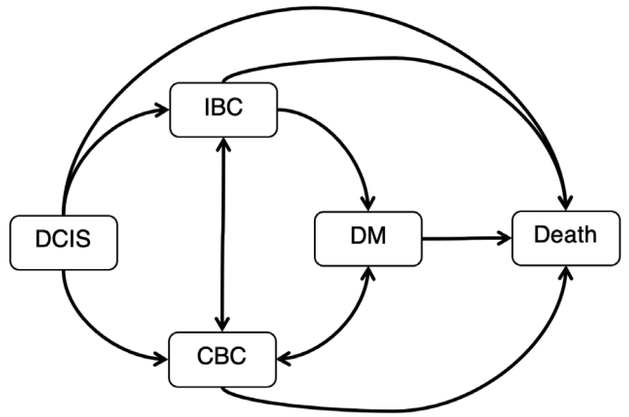


Figure 1: A simplified graphical representation of a multi-state modelling for breast cancer for dynamic prediction.
DCIS: ductal carcinoma in situ; IBC: ipsilateral breast cancer; CBC: contralateral breast cancer; DM: distant metastasis.

Perspectives and suggestions for further methodological research

As mentioned in the paragraphs before, no clear guidelines are currently available about how to analyze individual patient data with multiple studies in presence of competing risks. When missing values are present, how to incorporate competing risks outcomes in the imputation models is still an ongoing research, especially when multiple studies are included and systematic missing values can occur[41,43,44]. An overview about how to develop and validate a risk prediction model in presence of competing risks using individual data including multiple studies is really needed with a real application supported by software code for implementation. These potential guidelines and overviews might be extended in the context of multi-state and dynamic prediction modelling.

Cox proportional hazard models for each event and Fine and Gray regression are the

8

most known models for time-to-event data, although alternatives are possible[54,75]. Over the years, methodological research had focused on extending the potential violation of proportional hazard and random censoring assumptions. Other relevant issues include that many hospital- and population-based studies recruit patients at random times after diagnosis defined as left-truncation or delayed-entry. Few of the current discrimination measures (e.g. c-index and time-dependent Area Under the ROC curve) for time-to-event outcomes consider left-truncation[76]. Discrimination measures should be extended when left-truncation occurs. The potential violation of independent delayed entry assumption in parameter estimation and risk prediction might be additionally investigated. Simulation studies may be challenging in this setting[77-79].

Last but not least, prediction ignores technologies that will be discovered in the future[80]. There is an increasing interest in using and comparing machine learning and modern algorithms with standard statistical methods for risk prediction[81-83]. Thus, further comparison studies are welcome[84,85].

CONCLUSIONS

In conclusion, we paved the road for risk prediction of contralateral breast cancer (CBC) in patients diagnosed with first invasive breast cancer and ductal carcinoma in situ. The potential clinical utility and applications of CBC risk prediction models in clinical practice is largely investigated from a clinical viewpoint. An appropriate implementation and use of PredictCBC models may reassure patients about their fears to develop a new primary breast cancer in the opposite breast and to opt for alternative preventive strategies when their estimated CBC risk is low.

We sketched a framework for performance evaluation of predictions and clinical utility of survival and competing risks models using real word examples and providing the code of the statistical software currently used. There are still many stimulating and challenging opportunities to improve risk prediction and prognosis in the breast cancer field from genetics, biological and clinical perspectives. Much has been achieved in the last 30 years in medical statistics and biostatistics in risk prediction modelling, although identifying predictors that really improve prediction performances in (breast) cancer remains challenging. Further exciting opportunities lie ahead in methodological and applied research with the help of advance technological developments.

REFERENCES

1 Healey, E. A. *et al.* Contralateral breast cancer: clinical characteristics and impact on prognosis. *J Clin Oncol* **11**, 1545-1552, doi:10.1200/JCO.1993.11.8.1545 (1993).

2 Chen, Y., Thompson, W., Semenciw, R. & Mao, Y. Epidemiology of contralateral breast cancer. *Cancer Epidemiol Biomarkers Prev* **8**, 855-861 (1999).

3 Gao, X., Fisher, S. G. & Emami, B. Risk of second primary cancer in the contralateral breast in women treated for early-stage breast cancer: a population-based study. *Int J Radiat Oncol Biol Phys* **56**, 1038-1045, doi:10.1016/s0360-3016(03)00203-7 (2003).

4 Mariani, L. *et al.* Prognostic factors for metachronous contralateral breast cancer: a comparison of the linear Cox regression model and its artificial neural network extension. *Breast Cancer Res Treat* **44**, 167-178, doi:10.1023/a:1005765403093 (1997).

5 van den Broek, A. J. *et al.* Impact of Age at Primary Breast Cancer on Contralateral Breast Cancer Risk in BRCA1/2 Mutation Carriers. *J Clin Oncol* **34**, 409-418, doi:10.1200/JCO.2015.62.3942 (2016).

6 Vichapat, V. *et al.* Risk factors for metachronous contralateral breast cancer suggest two aetiological pathways. *Eur J Cancer* **47**, 1919-1927, doi:10.1016/j.ejca.2011.05.004 (2011).

7 Vichapat, V. *et al.* Prognosis of metachronous contralateral breast cancer: importance of stage, age and interval time between the two diagnoses. *Breast Cancer Res Treat* **130**, 609-618, doi:10.1007/s10549-011-1618-8 (2011).

8 Reiner, A. S. *et al.* Hormone receptor status of a first primary breast cancer predicts contralateral breast cancer risk in the WECARE study population. *Breast Cancer Res* **19**, 83, doi:10.1186/s13058-017-0874-x (2017).

9 Reiner, A. S. *et al.* Breast Cancer Family History and Contralateral Breast Cancer Risk in Young Women: An Update From the Women's Environmental Cancer and Radiation Epidemiology Study. *J Clin Oncol* **36**, 1513-1520, doi:10.1200/JCO.2017.77.3424 (2018).

10 Weischer, M. *et al.* CHEK2*1100delC heterozygosity in women with breast cancer associated with early death, breast cancer-specific death, and increased risk of a second breast cancer. *J Clin Oncol* **30**, 4308-4316, doi:10.1200/JCO.2012.42.7336 (2012).

11 Akdeniz, D. *et al.* Risk factors for metachronous contralateral breast cancer: A systematic review and meta-analysis. *Breast* **44**, 1-14, doi:10.1016/j.breast.2018.11.005 (2019).

12 Akdeniz, D., Klaver, M. M., Smith, C. Z. A., Koppert, L. B. & Hooning, M. J. The impact of lifestyle and reproductive factors on the risk of a second new primary cancer in the contralateral breast: a systematic review and meta-analysis. *Cancer Causes Control* **31**, 403-416, doi:10.1007/s10552-020-01284-2 (2020).

13 Elmore, J. G. & Fletcher, S. W. The risk of cancer risk prediction: "What is my risk of getting breast cancer"? *J Natl Cancer Inst* **98**, 1673-1675, doi:10.1093/jnci/djj501 (2006).

14 Basu, N. N., Ross, G. L., Evans, D. G. & Barr, L. The Manchester guidelines for contralateral risk-reducing mastectomy. *World J Surg Oncol* **13**, 237, doi:10.1186/s12957-015-0638-y (2015).

15 Chowdhury, M., Euhus, D., Onega, T., Biswas, S. & Choudhary, P. K. A model for individualized risk prediction of contralateral breast cancer. *Breast Cancer Res Treat* **161**, 153-160, doi:10.1007/s10549-016-4039-x

(2017).
16 Chowdhury, M. *et al.* Validation of a personalized risk prediction model for contralateral breast cancer. *Breast Cancer Res Treat* **170**, 415-423, doi:10.1007/s10549-018-4763-5 (2018).
17 Giardiello, D. *et al.* Prediction of contralateral breast cancer: external validation of risk calculators in 20 international cohorts. *Breast Cancer Res Treat* **181**, 423-434, doi:10.1007/s10549-020-05611-8 (2020).
18 Thompson, D. & Easton, D. The genetic epidemiology of breast cancer genes. *J Mammary Gland Biol Neoplasia* **9**, 221-236, doi:10.1023/B:JOMG.0000048770.90334.3b (2004).
19 Kramer, I. *et al.* Breast Cancer Polygenic Risk Score and Contralateral Breast Cancer Risk. *Am J Hum Genet*, doi:10.1016/j.ajhg.2020.09.001 (2020).
20 Fanale, D. *et al.* Detection of Germline Mutations in a Cohort of 139 Patients with Bilateral Breast Cancer by Multi-Gene Panel Testing: Impact of Pathogenic Variants in Other Genes beyond BRCA1/2. *Cancers (Basel)* **12**, doi:10.3390/cancers12092415 (2020).
21 Mellemkjaer, L. *et al.* Risk for contralateral breast cancer among carriers of the CHEK2*1100delC mutation in the WECARE Study. *Br J Cancer* **98**, 728-733, doi:10.1038/sj.bjc.6604228 (2008).
22 Kramer, I. *et al.* Breast Cancer Polygenic Risk Score and Contralateral Breast Cancer Risk. *Am J Hum Genet* **107**, 837-848, doi:10.1016/j.ajhg.2020.09.001 (2020).
23 Visser, L. L. *et al.* Predictors of an Invasive Breast Cancer Recurrence after DCIS: A Systematic Review and Meta-analyses. *Cancer Epidemiol Biomarkers Prev* **28**, 835-845, doi:10.1158/1055-9965.EPI-18-0976 (2019).
24 Miller, M. E. *et al.* Contralateral Breast Cancer Risk in Women with Ductal Carcinoma In Situ: Is it High Enough to Justify Bilateral Mastectomy? *Ann Surg Oncol* **24**, 2889-2897, doi:10.1245/s10434-017-5931-2 (2017).
25 Tuttle, T. M. *et al.* Increasing rates of contralateral prophylactic mastectomy among patients with ductal carcinoma in situ. *J Clin Oncol* **27**, 1362-1367, doi:10.1200/JCO.2008.20.1681 (2009).
26 Sauerbrei, W. *et al.* STRengthening analytical thinking for observational studies: the STRATOS initiative. *Stat Med* **33**, 5413-5432, doi:10.1002/sim.6265 (2014).
27 Blanche, P., Dartigues, J. F. & Jacqmin-Gadda, H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* **32**, 5381-5397, doi:10.1002/sim.5958 (2013).
28 Brentnall, A. R. & Cuzick, J. Risk Models for Breast Cancer and Their Validation. *Stat Sci* **35**, 14-30, doi:10.1214/19-STS729 (2020).
29 Schumacher, M. *et al.* Randomized 2 x 2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients. German Breast Cancer Study Group. *J Clin Oncol* **12**, 2086-2093, doi:10.1200/JCO.1994.12.10.2086 (1994).
30 Foekens, J. A. *et al.* The urokinase system of plasminogen activation and prognosis in 2780 breast cancer patients. *Cancer Res* **60**, 636-643 (2000).
31 de Glas, N. A. *et al.* Postoperative complications and survival of elderly breast cancer patients: a FOCUS study analysis. *Breast Cancer Res Treat* **138**, 561-569, doi:10.1007/s10549-013-2462-9 (2013).
32 Font-Gonzalez, A. *et al.* Inferior survival for young patients with contralateral compared to unilateral breast cancer: a nationwide population-based study in the Netherlands. *Breast Cancer Res Treat* **139**, 811-819, doi:10.1007/s10549-013-2588-9 (2013).
33 Schmidt, M. K. *et al.* Breast Cancer Survival of BRCA1/BRCA2 Mutation Carriers in a Hospital-Based Cohort of Young Women. *J Natl Cancer Inst* **109**, doi:10.1093/jnci/djw329 (2017).
34 Pijpe, A. *et al.* Physical activity and the risk of breast cancer in BRCA1/2 mutation carriers. *Breast Cancer Res Treat* **120**, 235-244, doi:10.1007/s10549-009-0476-0 (2010).
35 Michailidou, K. *et al.* Association analysis identifies 65 new breast cancer risk loci. *Nature* **551**, 92-94, doi:10.1038/nature24284 (2017).
36 Mavaddat, N. *et al.* Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am J Hum Genet* **104**, 21-34, doi:10.1016/j.ajhg.2018.11.002 (2019).
37 Madley-Dowd, P., Hughes, R., Tilling, K. & Heron, J. The proportion of missing data should not be used to guide decisions on multiple imputation. *J Clin Epidemiol* **110**, 63-73, doi:10.1016/j.jclinepi.2019.02.016 (2019).
38 Sterne, J. A. *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* **338**, b2393, doi:10.1136/bmj.b2393 (2009).
39 Van Buuren, S. *Flexible imputation of missing data*. Second edn, (Chapman and Hall/CRC, 2018).
40 Jolani, S., Debray, T. P., Koffijberg, H., van Buuren, S. & Moons, K. G. Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using MICE. *Stat Med* **34**, 1841-1863, doi:10.1002/sim.6451 (2015).
41 Resche-Rigon, M. & White, I. R. Multiple imputation by chained equations for systematically and sporadically missing multilevel data. *Stat Methods Med Res*, doi:10.1177/0962280216666564 (2016).
42 Kontopantelis, E. A comparison of one-stage vs two-stage individual patient data meta-analysis methods: A simulation study. *Res Synth Methods* **9**, 417-430, doi:10.1002/jrsm.1303 (2018).
43 White, I. R. & Royston, P. Imputing missing covariate values for the Cox model. *Stat Med* **28**, 1982-1998, doi:10.1002/sim.3618 (2009).
44 Bartlett, J. W. & Taylor, J. M. Missing covariates in competing risks analysis. *Biostatistics* **17**, 751-763, doi:10.1093/biostatistics/kxw019 (2016).
45 Debray, T. P., Moons, K. G., Abo-Zaid, G. M., Koffijberg, H. & Riley, R. D. Individual participant data meta-analysis for a binary outcome: one-stage or two-stage? *PLoS One* **8**, e60650, doi:10.1371/journal.pone.0060650 (2013).
46 Debray, T. P., Koffijberg, H., Vergouwe, Y., Moons, K. G. & Steyerberg, E. W. Aggregating published prediction models with individual participant data: a comparison of different approaches. *Stat Med* **31**, 2697-2712, doi:10.1002/sim.5412 (2012).
47 Burke, D. L., Ensor, J. & Riley, R. D. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med* **36**, 855-875, doi:10.1002/sim.7141 (2017).
48 Ahmed, I., Debray, T. P., Moons, K. G. & Riley, R. D. Developing and validating risk prediction models in an individual participant data meta-analysis. *BMC Med Res Methodol* **14**, 3, doi:10.1186/1471-2288-14-3 (2014).
49 Steyerberg, E. W., Nieboer, D., Debray, T. P. A. & van Houwelingen, H. C. Assessment of heterogeneity in an individual participant data meta-analysis of prediction models: An overview and illustration. *Stat Med* **38**, 4290-4309, doi:10.1002/sim.8296 (2019).
50 Michiels, S., Baujat, B., Mahe, C., Sargent, D. J. & Pignon, J. P. Random effects survival models gave a better understanding of heterogeneity in individual patient data meta-analyses. *J Clin Epidemiol* **58**, 238-245, doi:10.1016/j.jclinepi.2004.08.013 (2005).

51 Bowden, J., Tierney, J. F., Simmonds, M., Copas, A. J. & Higgins, J. P. Individual patient data meta-analysis of time-to-event outcomes: one-stage versus two-stage approaches for estimating the hazard ratio under a random effects model. *Res Synth Methods* **2**, 150-162, doi:10.1002/jrsm.45 (2011).

52 Smith, C. T., Williamson, P. R. & Marson, A. G. Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Stat Med* **24**, 1307-1319, doi:10.1002/sim.2050 (2005).

53 Westeneng, H. J. *et al.* Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *Lancet Neurol*, doi:10.1016/S1474-4422(18)30089-9 (2018).

54 Royston, P. & Parmar, M. K. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* **21**, 2175-2197, doi:10.1002/sim.1203 (2002).

55 Ensor, J. *et al.* Individual participant data meta-analysis for external validation, recalibration, and updating of a flexible parametric prognostic model. *Stat Med*, doi:10.1002/sim.8959 (2021).

56 Wolbers, M. *et al.* Competing risks analyses: objectives and approaches. *Eur Heart J* **35**, 2936-2941, doi:10.1093/eurheartj/ehu131 (2014).

57 Meddis, A., Latouche, A., Zhou, B., Michiels, S. & Fine, J. Meta-analysis of clinical trials with competing time-to-event endpoints. *Biom J* **62**, 712-723, doi:10.1002/bimj.201900103 (2020).

58 *Federatie Medisch Specialisten*, <https://richtlijnendatabase.nl/richtlijn/borstkanker/algemeen.html> (

59 Antoniou, A. C., Pharoah, P. P., Smith, P. & Easton, D. F. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer* **91**, 1580-1590, doi:10.1038/sj.bjc.6602175 (2004).

60 Antoniou, A. C. *et al.* The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* **98**, 1457-1466, doi:10.1038/sj.bjc.6604305 (2008).

61 Brenner, D. J. Contralateral second breast cancers: prediction and prevention. *J Natl Cancer Inst* **102**, 444-445, doi:10.1093/jnci/djq058 (2010).

62 O'Donnell, M. Estimating Contralateral Breast Cancer Risk. *Current Breast Cancer Reports* **10**, 91-97 (2018).

63 Abbott, A. *et al.* Perceptions of contralateral breast cancer: an overestimation of risk. *Ann Surg Oncol* **18**, 3129-3136, doi:10.1245/s10434-011-1914-x (2011).

64 Portschy, P. R. *et al.* Perceptions of Contralateral Breast Cancer Risk: A Prospective, Longitudinal Study. *Ann Surg Oncol* **22**, 3846-3852, doi:10.1245/s10434-015-4442-2 (2015).

65 Jenkins, D. A. *et al.* Continual updating and monitoring of clinical prediction models: time for dynamic prediction systems? *Diagn Progn Res* **5**, 1, doi:10.1186/s41512-020-00090-3 (2021).

66 Breast Cancer Association, C. *et al.* Breast Cancer Risk Genes - Association Analysis in More than 113,000 Women. *N Engl J Med* **384**, 428-439, doi:10.1056/NEJMoa1913948 (2021).

67 Knight, J. A. *et al.* The association of mammographic density with risk of contralateral breast cancer and change in density with treatment in the WECARE study. *Breast Cancer Res* **20**, 23, doi:10.1186/s13058-018-0948-4 (2018).

68 McCormack, V. A. & dos Santos Silva, I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* **15**, 1159-1169, doi:10.1158/1055-9965.EPI-06-0034 (2006).

69 Putter, H., van der Hage, J., de Bock, G. H., Elgalta, R. & van de Velde, C. J. Estimation and prediction in a multi-state model for breast cancer. *Biom J* **48**, 366-380 (2006).

70 van Houwelingen, H. C. & Putter, H. *Dynamic prediction in clinical survival analysis*. (CRC Press, 2011).

71 Rizopoulos, D. *Joint models for longitudinal and time-to-event data: with applications in R*. (CRC Press, 2012).

72 Fontein, D. B. *et al.* Dynamic prediction in breast cancer: proving feasibility in clinical practice using the TEAM trial. *Ann Oncol* **26**, 1254-1262, doi:10.1093/annonc/mdv146 (2015).

73 Geskus, R. B. *Data analysis with competing risk and intermediate states*. (CRC press, 2016).

74 Martin, G. P., Sperrin, M., Snell, K. I. E., Buchan, I. & Riley, R. D. Clinical prediction models to predict the risk of multiple binary outcomes: a comparison of approaches. *Stat Med* **40**, 498-517, doi:10.1002/sim.8787 (2021).

75 Haller, B., Schmidt, G. & Ulm, K. Applying competing risks regression models: an overview. *Lifetime Data Anal* **19**, 33-58, doi:10.1007/s10985-012-9230-8 (2013).

76 Li, S. Estimating time-dependent ROC curves using data under prevalent sampling. *Stat Med* **36**, 1285-1301, doi:10.1002/sim.7184 (2017).

77 Tsai, W.-Y. Testing the assumption of independence of truncation time and failure tim. *Biometrika* **77**, 169-177 (1990).

78 Niels, K. & Moeschberger, M. in *Survival analysis: state of art* (Springer, 1992).

79 Azzato, E. M. *et al.* Prevalent cases in observational studies of cancer survival: do they bias hazard ratio estimates? *Br J Cancer* **100**, 1806-1811, doi:10.1038/sj.bjc.6605062 (2009).

80 Taleb, N. N. *The black swan: the impact of the highly improbable* (Penguin books, 2007).

81 Christodoulou, E. *et al.* A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* **110**, 12-22, doi:10.1016/j.jclinepi.2019.02.004 (2019).

82 Wang, P., Li, Y. & Reddy, C. K. Machine Learning for Survival Analysis: A Survey. *ACM Computing Surveys (CSUR)* **51**, 1-36, doi:10.1145/3214306 (2019).

83 Ming, C. *et al.* Machine learning techniques for personalized breast cancer risk prediction: comparison with the BCRAT and BOADICEA models. *Breast Cancer Res* **21**, 75, doi:10.1186/s13058-019-1158-4 (2019).

84 Giardiello, D., Antoniou, A. C., Mariani, L., Easton, D. F. & Steyerberg, E. W. Letter to the editor: a response to Ming's study on machine learning techniques for personalized breast cancer risk prediction. *Breast Cancer Res* **22**, 17, doi:10.1186/s13058-020-1255-4 (2020).

85 Wilkinson, J. *et al.* Time to reality check the promises of machine learning-powered precision medicine. *Lancet Digit Health* **2**, e677-e680, doi:10.1016/S2589-7500(20)30200-4 (2020).