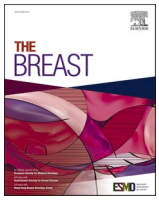
The Breast 82 (2025) 104506

Contents lists available at ScienceDirect

The Breast

journal homepage: www.journals.elsevier.com/the-breast



Integrating genetic polymorphisms and clinical data to develop predictive models for skin toxicity in breast cancer radiation therapy

Ester Aguado-Flor [a,b,1], Victoria M. Reyes [c,1], Víctor Navarro [d], Meritxell Mollà [c,2], Miguel E. Aguado-Barrera [e,f], Manuel Altabas [c], David Azria [g], Adinda Baten [h], Celine Bourgier [g], Renée Bultijnck [i,j], Jenny Chang-Claude [k,l], Maria Carmen De Santis [m], Alison M. Dunning [n], Laura Duran-Lozano [a], Rebecca M. Elliott [o], Marie-Pierre Farcy Jacquet [p], Carlotta Giandini [m], Alexandra Giraldo [c], Sheryl Green [q], Maarten Lambrecht [h], Carlos Lopez-Pleguezuelos [e,f], Chris Monten [i,j], Tiziana Rancati [r], Tim Rattay [s], Barry S. Rosenstein [q], Dirk De Ruysscher [t], Orland Diez [a,u], Petra Seibold [k], Elena Sperk [v], R Paul Symonds [s], Hilary Stobart [w], Ana Vega [e,f,x], Liv Veldeman [i,j], Guillermo Villacampa [d], Adam J. Webb [y], Caroline Weltens [h], Paolo Zunino [z], Christopher J. Talbot [y], Catharine M. West [o], Jordi Giralt [c,aa], Sara Gutiérrez-Enríquez [a,*], on behalf of the REQUITE consortium[3]

[a] *Hereditary Cancer Genetics Group, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain*
[b] *University of Barcelona, Barcelona, Spain*
[c] *Department of Radiation Oncology, Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain*
[d] *Biostatistics Unit, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain*
[e] *Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Santiago de Compostela, A Coruña, Spain*
[f] *Fundación Pública Galega de Medicina Xenómica (FPGMX), Santiago de Compostela, A Coruña, Spain*
[g] *Institut de Recherche en Cancérologie de Montpellier, University Federation of Radiation Oncology of Mediterranean Occitanie, Université de Montpellier, INSERM U1194 IRCM, Montpellier, France*
[h] *Radiation Oncology, UZ Leuven, Leuven, Belgium*
[i] *Department of Human Structure and Repair, Ghent University, Ghent, Belgium*
[j] *Department of Radiation Oncology, Ghent University Hospital, Ghent, Belgium*
[k] *Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany*
[l] *Cancer Epidemiology Group, University Cancer Center Hamburg, University Medical Center Hamburg-Eppendorf, Hamburg, Germany*
[m] *Radiation Oncology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*
[n] *Centre for Cancer Genetic Epidemiology, Dept of Oncology, Strangeways Research Laboratory, University of Cambridge, Cambridge, CB1 8RN, UK*
[o] *Translational Radiobiology Group, Division of Cancer Sciences, University of Manchester, Manchester Academic Health Science Centre, The Christie NHS Foundation Trust, Manchester, United Kingdom*
[p] *University Federation of Radiation Oncology of Occitanie Méditerranée, CHU Nîmes, ICG, Rue Henri Pujol, 34000, Nîmes, France*
[q] *Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, NY, USA*
[r] *Unit of Data Science, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy*
[s] *Leicester Cancer Research Centre, University of Leicester, Leicester, United Kingdom*
[t] *MAASTRO Clinic, GROW School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, Netherlands*
[u] *Area of Clinical and Molecular Genetics, Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain*
[v] *Department of Radiation Oncology, Mannheim Cancer Center, Universitätsmedizin Mannheim, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany*
[w] *Patient Advocate, Independent Cancer Patients' Voice, United Kingdom*
[x] *Biomedical Network on Rare Diseases (CIBERER), 15706, Santiago de Compostela, Spain*
[y] *Department of Genetics and Genome Biology, University of Leicester, Leicester, United Kingdom*
[z] *MOX, Math Department, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milan, 20133, Italy*
[aa] *Radiation Oncology Group, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain*

* Corresponding author. Sara Gutiérrez-Enríquez. Hereditary Cancer Genetics Group, Vall d'Hebron Institute of Oncology (VHIO), Office 4.08, Cellex Center, c/ Natzaret, 115-117, 08035, Barcelona, Spain
E-mail address: sgutierrez@vhio.net (S. Gutiérrez-Enríquez).
[1] Ester Aguado-Flor and Victoria M Reyes contributed equally to this article.
[2] Present address: Radiation Oncology Department, Hospital Clinic of Barcelona, Spain.
[3] Consortium members are listed at the end of the manuscript.

https://doi.org/10.1016/j.breast.2025.104506

Available online 22 May 2025

ARTICLE INFO

Keywords:
Breast cancer
Radiation therapy-induced side-effects
Non-coding RNA polymorphisms
Literature candidate polymorphisms' validation
Predictive models

ABSTRACT

Background: We aim to develop and validate predictive models for acute and late skin toxicity in breast cancer (BC) patients undergoing radiation therapy (RT). Models incorporate a genetic profile—comprising candidate single nucleotide polymorphisms (SNPs) in non-coding RNAs and previously reported toxicity-associated variants—combined with clinical variables.

Methods: The study involved 1979 BC patients monitored for two to eight years post-RT in a multi-centre study. We assessed acute (oedema/erythema) and late (atrophy/fibrosis) toxicity using logistic regression and Cox proportional hazards models. The cohort was divided into training and validation datasets.

Results: Six SNPs demonstrated to be predictors of acute (rs13116075, rs12565978, rs72550778 and rs7284767) and late toxicity (rs16837908 and rs61764370) either in the training or validation cohort. However, none of these SNPs were consistently associated with toxicity across both stages of analysis. The rs13116075, rs12565978 and rs16837908 were previously reported to be associated with RT toxicity. In the validation phase, SNP-based models showed limited predictive ability, with AUC values of 0.49 and c-index of 0.54 for acute and late toxicity, respectively. Models incorporating either clinical variables alone or in combination with SNPs achieved similar AUC and c-index values of ~0.60 for acute and late toxicity, respectively. However, the combined model exhibited the highest predictive accuracy for acute and late toxicity, both in the training and the validation cohorts.

Conclusions: Our findings highlight the importance of combining clinical data with genetic markers to enhance the accuracy of models predicting RT toxicity in BC.

1. Background

As long-term cancer survival rates rise, survivorship issues and quality of life are becoming an increasingly important research focus in oncology. Radiation therapy (RT) is an integral component of breast cancer (BC) treatment, with an increasing number of BC patients receiving potentially curative or palliative RT. RT effectively reduces the risk of local recurrence and contributes to a decrease in overall mortality [1]. However, breast RT can be associated with several side-effects due to normal tissue responses to ionising radiation, which can be acute and/or late [2]. Acute toxicities occur within 90 days of treatment, tend to be transient and include breast erythema, oedema, and desquamation [3]. Late toxicities may appear months or years after RT and are concerning due to their persistence, potential severity and impact on quality of life [4]. Examples of late toxicities mainly include telangiectasia, skin induration (fibrosis), skin hyperpigmentation, arm lymphoedema, and atrophy [5].

The therapeutic window of RT, which is the range of radiation doses that can effectively treat a tumour while minimizing damage to surrounding healthy tissues, is narrow and calculated from the average response, although patient-to-patient variability is high [6]. Patient stratification according to their risk of radiation toxicity would allow clinicians to adjust the treatment for each patient.

The risk of developing RT related toxicity is driven by patient, tumour, and treatment-related factors together with individual genetic predisposition-derived sensitivity. In the context of breast RT, numerous studies have explored the correlation between clinical and treatment risk factors and the occurrence of acute and late skin toxicity [7–12]. However, only a limited number of these studies have presented a comprehensive clinical prediction model [9]. In addition, the incorporation of genetic susceptibility markers into these predictive models has been even rarer, with such integration having been partially explored and modelled so far only in the case of prostate cancer [13,14].

Radiosensitivity is believed to be a complex, inherited, and polygenic trait [15]. Previous research has provided evidence that the risk of radiation-induced toxicity is influenced by common low-penetrance single nucleotide polymorphisms (SNPs) [16–20]. Increasing evidence point to the non-coding RNAs as important biological regulators of numerous cellular processes and sequence variants in such regulatory elements have the potential to affect phenotype through altered gene expression [18]. To our knowledge, there is limited research regarding the impact of non-coding RNAs, such as microRNAs (miRNAs) and long-non-coding RNAs (lncRNAs), on the susceptibility to side effects from RT [18,21].

This study aims to examine the association between a set of SNPs located in non-coding RNAs in addition to gene variants previously linked to RT toxicity, and the development of acute and late RT-induced toxicity in 1979 BC patients followed for a minimum of two years and up to eight years after RT. Furthermore, the study seeks to develop and validate predictive models by integrating these SNPs and relevant clinical variables.

2. Methods

2.1. Patients

Germline blood DNA samples, collected before RT, along with clinical data recorded before and after RT, were obtained from two distinct cohorts of BC patients. The first cohort consisted of 115 BC patients treated at Vall d'Hebron University Hospital and prospectively recruited between 2009 and 2014 (Vall d'Hebron prospective cohort). The second cohort comprised 2057 BC patients treated at various European and USA hospitals recruited consecutively as part of the REQUITE European project between 2014 and 2016 (REQUITE cohort) [22].

For the two cohorts (Vall d'Hebron and REQUITE), the inclusion criteria encompassed patients suitable for postoperative RT for BC after breast-conserving surgery, including both invasive and *in situ* cases, as well as patients receiving primary systemic therapy. Male patients as well as patients who underwent concomitant chemo-radiation, mastectomy, partial breast irradiation, or had bilateral BC were excluded. All patients in both cohorts gave written informed consent and their respective protocols were approved by Vall d'Hebron and local ethics committees in REQUITE participating countries. REQUITE was registered at www.controlled-trials.com (ISRCTN98496463).

Patients recruited through Vall d'Hebron prospective cohort were followed for up to eight years. Patients from the REQUITE cohort were initially followed for a minimum of two years after RT. Their follow-up continued for up to eight years through the REQUITEplus and RADprecise projects.

Genotyping was successful for 2125 patients. For this study, only patients of European descent were included in the analysis, resulting in the exclusion of 123 patients with diverse ethnicities. Patients that only

had baseline data and no follow-up were also excluded from the analysis (n = 23). Finally, a total of 1979 BC patients were included in the study (Fig. 1).

2.2. Data collection and toxicity endpoints definition

In both cohorts, the occurrence of acute and late effects of RT were monitored using Common Terminology Criteria for Adverse Events (CTCAE)v4.0 system. Toxicity data, as well as clinical data, were documented before RT (baseline), at the end of RT and at each annual follow-up after RT.

We defined as cases patients with acute RT-induced toxicity who had oedema or erythema grade ≥2 at the end of RT, provided that their baseline toxicity was 0 or 1. Patients with baseline grade ≥2 at baseline in oedema or erythema toxicity were excluded (n = 10). Patients without acute toxicity (controls) were those with oedema or erythema grade 0 or 1 both at baseline and at the end of RT. Seven patients did not have the toxicity recorded in the post-RT visit and thus were excluded from the acute toxicity study. In total there were 1962 patients for the acute toxicity study (Fig. 1).

For late toxicity, patients who had atrophy or fibrosis grade ≥2 at two years until maximum follow-up after RT and baseline toxicity grades 0 or 1 were defined as cases. Patients with baseline grade ≥2 in atrophy, nipple retraction, telangiectasia, oedema, fibrosis or arm lymphedema toxicity were excluded (n = 126). Controls were those with grade 0 or 1 at baseline and at two years until maximum follow-up after RT in atrophy, nipple retraction, telangiectasia, fibrosis or arm lymphedema. Patients who experienced a downgrade from grade 2 or 3 to grade 0 at consecutive years of the follow-up were excluded, as these were considered as unclear toxicities (n = 54). Two hundred thirty-nine patients did not have follow-up information after 12 months and were therefore excluded from the study. In total there were 1560 patients for the late toxicity study (Fig. 1).

2.3. Polymorphism selection and genotyping

Two sets of SNPs were examined in this study (Supp. Fig. 1). The first set comprised 53 SNPs located in miRNAs, the gene 3′ untranslated regions (3′UTRs), and lncRNAs which we consider novel candidates as they had not been previously evaluated. To improve model performance, the second set included 16 SNPs that had been associated with normal tissue reactions to RT in previously published GWAS or independent candidate gene studies. Blood DNA was genotyped using MassArray Agena Bioscience or Illumina Oncoarray. In total, 44 SNPs were successfully genotyped and included in the study: 31 non-coding RNA SNPs and 13 SNPs previously reported in the literature (Supp. Fig. 1 and Supp. Tables 1–4). See Supplementary Material for more details on the SNPs' selection criteria and genotyping techniques.

2.4. Statistical analysis

Clinical and treatment variables with more than 20 % missing values were excluded from further analysis. To address potential bias from analysing only complete cases, we utilized multivariate imputation via chained equations (MICE) to replace randomly missing values [23]. Differences between the imputed and non-imputed datasets were assessed using standardized mean differences (SMDs). As shown in Supplementary Table 5, all SMDs were below 0.2. According to Cohen's guidelines, an SMD below 0.2 indicates a small effect size, suggesting minimal bias introduced by the imputation process [24].

We analysed the association of SNPs, clinical data, and their combination with acute and late toxicity separately. For each toxicity analysis, the dataset was split into training (acute toxicity: n = 920 [47 %]; late toxicity: n = 739 [47 %]) and validation sets (acute toxicity: n = 1042 [53 %]; late toxicity: n = 821 [53 %]) (Fig. 1). Patients were divided by hospital or treatment centre to ensure that the training and validation datasets included patients from distinct, non-overlapping locations. This approach maintained independence between centres; however, no adjustments were made based on the analysed endpoints (Fig. 1 and Supp. Table 6).

In the training set and for acute toxicity, a logistic model was employed to estimate odds ratio (OR). In contrast, for late toxicity, a Cox proportional hazards model was selected to obtain hazard ratios (HR). Additionally, the Kaplan-Meier method was employed to calculate the cumulative incidence of late toxicity. The least absolute shrinkage and selection operator (Lasso) regression was used to select factors for the multivariable regression analysis [25]. Lasso with the minimum lambda was applied to clinical variables, while for SNPs, lambda was adjusted to maximize the area under the curve (AUC) or c-index while utilizing the minimum number of variables. Additionally, relevant clinical predictors identified through multivariate analyses previously reported in the literature, were retained in the analysis regardless of the Lasso selection output [9,26].

A risk prediction score was estimated using coefficients derived from the logistic regression and Cox proportional hazards multivariate models. The Youden index was used to determine an optimal cut-off of the scores obtained from the different toxicity prediction models (acute and late), which allowed us to categorize patients into high and low risk toxicity strata [27]. This cut-off obtained from the multivariate analysis with the training data was then applied to make predictions on both the training and validation sets.

We implemented bootstrapping as an internal validation technique to address over-fitting and correct for over-optimism [28]. Specifically, we iterated the model development process using 1000 bootstrap samples for internal validation.

For evaluating performance, AUC-ROC was used for acute toxicity and c-index [29], for late toxicity. To assess whether the differences between AUC-ROC and c-index values were statistically significant, we applied the DeLong test [30] for acute toxicity, and the one-shot nonparametric approach described by Kang et al. (2015) for late toxicity, which is specifically designed for comparing two correlated c-indices in the presence of right-censored survival data [31]. Additionally, we computed various performance metrics, including accuracy, sensitivity and specificity [32], based on the predictions made by the model using the determined cut-off. Calibration plots for acute toxicity were evaluated through grouped real proportions versus mean predicted probability, while for late toxicity calibration the plots were estimated through the Cox-Snell residuals [33] on the cumulative probability scale. Overall accuracy for acute toxicity was evaluated using the Brier score.

All analyses were conducted using the R statistical software version 4.2.2.

3. Results

3.1. Toxicity distribution among training and validation cohorts

Fig. 2 presents the distribution of acute and late toxicities in the training and validation cohorts. In the training cohort, 29 % of patients experienced grade ≥2 acute toxicities, compared to 18 % in the validation cohort. For late toxicities, 20 % of patients in the training cohort exhibited grade ≥2, whereas the validation cohort showed a higher prevalence at 25 %.

3.2. Acute toxicity study

3.2.1. Patient characteristics

A total of 1962 BC patients were included in the acute toxicity study, with available SNP, clinical, treatment, and toxicity data after RT. Table 1 provides an overview of the patients' clinical and treatment characteristics.

3.2.2. Clinical variables analysis

In the training cohort, comprising 920 patients, the Lasso method was performed in a total of 45 clinical variables to select those with the highest impact on acute toxicity for inclusion in the multivariate analysis. The following 13 variables were identified by Lasso and included in the multivariate analysis: whole breast RT dose, breast volume in the planning computed tomography (CT), pathologic UICC stage, bra cup size, diabetes, RT toxicity family history, RT hypofractionation, tumour histological grade, analgesics, rheumatoid arthritis, 3D RT, intensity-modulated RT (IMRT), and surgery type. Seven clinically significant variables, previously known for their predictive value in toxicity, were also included in the analysis: body mass index (BMI), tamoxifen use, smoking status, axillary surgery type, chemotherapy, age at RT start, and RT boost [9,26].

After performing multivariate logistic regression analysing the effect on acute toxicity of the 20 variables listed above, breast volume in the planning CT and whole breast RT dose were consistently found to be significantly associated with acute toxicity in both in the training and validation cohorts [p-value (p) < 0.05] (Supp. Fig. 2).

We estimated a risk score for each patient of the training set using the coefficients derived from the previous logistic regression model involving the 20 clinical variables (Supp. Table 7). This score was then dichotomized into two levels (high risk and low risk) using the Youden index. This variable was found to exhibit a significant association with the occurrence of acute toxicity in the training cohort [OR, 5.84; 95 % confidence interval (CI), 4.19–8.24; $p < 0.001$] as well as in the validation cohort (OR, 1.42; CI, 1.03–1.94; $p = 0.030$) (Fig. 3A).

3.2.3. SNPs analysis

After Lasso selection, 18 out of the 44 studied SNPs were selected to be included in the multivariate analysis. Twelve of them were in non-coding RNAs. Multivariate analysis of acute toxicity in the training cohort identified three SNPs with significant associations: rs13116075 (OR, 1.45; CI, 1.09–1.93; $p = 0.01$), rs72550778 (OR, 1.73; CI, 1.06–2.81; $p = 0.03$), and rs7284767 (OR, 1.29; CI, 1.03–1.61; $p = 0.03$) (Supp. Fig. 3 and Supp. Table 8). These three SNPs were in non-coding RNAs, including rs13116075, which was previously shown to be associated with overall late toxicity in an independent dataset [16]. However, when testing the model in the validation cohort one different SNP, the rs12565978 (located near *PLXNA2*), previously reported to be linked to late toxicity in Barnett GC et al. Radiotherapy and Oncology (2014) [16], was significantly associated with acute toxicity (OR, 1.52; CI, 1.10–2.07; $p = 0.009$) (Supp. Fig. 3).

The risk score obtained from the logistic regression model using the 18 SNPs (Supp. Table 8), was found to be significantly associated with acute toxicity (OR, 2.31; CI, 1.73–3.10; $p < 0.001$) in the training cohort. However, when tested in the validation cohort, this association was not significant (OR, 0.98; CI, 0.71–1.34; $p = 0.877$) (Fig. 3A).

3.2.4. Integration of clinical variables and SNPs

Multivariate analysis involving clinical and SNP variables identified the same significant variables as those observed in separate models (Suppl. Fig. 4). The dichotomized risk score derived from the logistic regression model analysing the impact of the combination of the 20 clinical variables with the 18 SNPs (Supp. Table 9), was found to be significantly associated with the acute toxicity (OR, 7.24; CI, 5.25–10.07; $p < 0.001$) in the training cohort and in the validation cohort (OR, 1.66; CI, 1.21–2.29; $p = 0.0017$) (Fig. 3A).

3.2.5. Model performance

Fig. 4A, a ROC curve in the training cohort, demonstrated that the model utilizing clinical variables with an AUC of 0.747 (AUC after bootstrap optimism corrected (AUC_{om}) = 0.708) outperformed the model using SNPs (AUC = 0.633, AUC_{om} = 0.589), while combining

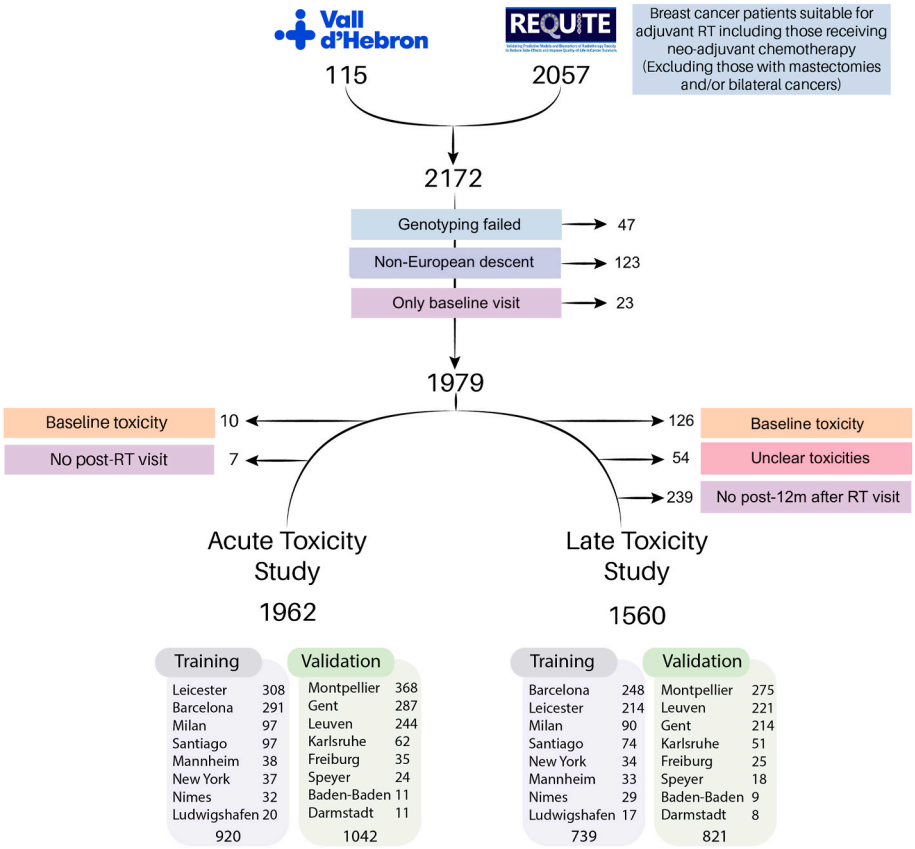


Fig. 1. Study flow diagram depicting numbers of analysed patients included in training and validation cohorts. RT: radiation therapy. 12m: 12 months.

both had the highest AUC of 0.776 (AUC_{om} = 0.721). In Fig. 4B, we observed that all models performed worse in the validation cohort (SNPs model AUC = 0.487, the clinical model AUC = 0.599, and combined model AUC = 0.597). Supplementary Table 10 shows the statistical differences between the AUCs. In the training cohort, the SNPs model achieved 61 % accuracy, the clinical model achieved 66 %, and the combined model performed the best with 72 %. However, these values dropped to 55 % for the clinical and the genetic model, and to 62 % for the combined model in the validation cohort (Fig. 4C). Sensitivity and specificity values also decreased in the validation cohort (Fig. 4D and E). In the validation cohort, calibration slopes of −0.142, 0.286, and 0.250, along with Brier scores of 0.178, 0.173, and 0.170, were observed for the SNPs, clinical, and combined models, respectively (Suppl. Fig. 5).

3.3. Late toxicity study

3.3.1. Patient characteristics

In the late toxicity study 1560 BC patients were included, all of whom had complete SNP, clinical, treatment, and toxicity data available for analysis ranging from two to eight years after RT treatment. Table 2 provides an overview of the patients' clinical and treatment characteristics.

3.3.2. Clinical variables analysis

In the training cohort (739 patients), a total of 46 clinical variables were analysed using the Lasso method to select relevant variables for multivariate analysis. The following 16 variables were identified by Lasso and included in the multivariate analysis: whole breast RT dose, breast volume in planning CT scan, smoking status, alcohol intake, age at RT start, pathologic UICC stage, depression, hypertension, breast cancer phenotype (as defined in Table 2), tumour histological grade, tumour quadrant, RT hypofractionation, rheumatoid arthritis, IMRT, chemotherapy and BMI. Additionally, three variables were retained in the multivariate analysis based on previously data reporting their predictive value: tamoxifen use, axillary surgery type, RT boost [9,34].

In the multivariate Cox proportional hazards model, only BMI was consistently found to be significantly associated with late toxicity in both training and validation cohorts ($p < 0.05$) (Supp. Fig. 6).

The categorised risk score derived from the Cox proportional hazards model for the 19 clinical variables was significantly associated with late toxicity [Hazard Ratio (HR), 4.12; CI, 2.93–5.79; $p < 0.001$] (Supp. Table 11) (Fig. 3B). The significance of this association remained robust after testing the model in the validation cohort consisting of 821 patients (HR, 1.83; CI, 1.37–2.44; $p < 0.001$) (Fig. 3B).

3.3.3. SNPs analysis

Following Lasso selection, 18 out of the 44 studied SNPs were selected to be included in the multivariate analysis. Thirteen of them were in non-coding RNAs. After multivariate analysis, rs61764370, which is in the 3′UTR of *KRAS*, was significantly associated with late toxicity in the training cohort (HR, 0.58; CI, 0.38–0.90; $p = 0.01$) (Supp. Fig. 7). However, when testing the model in the validation cohort one different SNP, rs16837908, was significantly associated with late toxicity (HR, 0.37; CI, 0.14–0.97; $p = 0.04$) (Supp. Fig. 7). This SNP was previously reported in Barnett GC et al. Radiotherapy and Oncology (2014) [16] to be associated with late RT toxicity, and it is located near *INO80D*.

The risk score obtained using the 18 SNPs (Supp. Table 12) was significantly associated with late toxicity, (HR, 3.26; CI, 1.99–5.34; $p < 0.001$), but not in the validation cohort (HR, 1.18; CI, 0.87–1.62; $p = 0.2849$) (Fig. 3B).

3.3.4. Integration of clinical variables and SNPs

Multivariate analysis involving clinical and SNP variables identified BMI and smoking as significant variables in both training and validation cohorts (Suppl. Fig. 8). The risk prediction score obtained using the combination of the 19 clinical variables and 18 SNPs (Supp. Table 13), was found to be significantly associated with late toxicity (HR, 3.26; CI, 1.99–5.34; $p < 0.001$) in the training cohort as well as in the validation cohort (HR, 1.63; CI, 1.24–2.14; $p < 0.001$) (Fig. 3B).

3.3.5. Cumulative incidence analysis

The Kaplan-Meier cumulative incidence plots (Fig. 5) display the predicted probabilities of developing toxicity for both high-risk and low-risk groups. In the training cohort, we observed that all model predictions were able to significantly discriminate between the two groups (Fig. 5A, B, C, respectively, at the top). However, in the validation cohort, while the clinical and combined models remained significant, the SNP model did not.

3.3.6. Model performance

In the training cohort, the SNP model performed worse than the clinical model (c-index: 0.62 vs. 0.75). The combined model incorporating clinical and SNP variables achieved the highest c-index of 0.77 (Fig. 6A). However, after internal validation with bootstrapping, all models showed decreased performance, further dropping in the validation cohort, with clinical and combined models at 0.59 and the SNP model at 0.54. The statistical differences between the c-index values are shown in Supplementary Table 10. Accuracy values, illustrated in Fig. 6B, showed that in the training cohort, the SNP model achieved 44 % accuracy, the clinical model 71 %, and the combined model performed the best with 74 %. However, in the validation cohort, these values dropped to 40 % for the SNPs, 54 % for the clinical, and 57 % for the combined model (Fig. 6B). In Fig. 6C and D, the sensitivity and specificity values are presented, showing consistent patterns of decreasing performance in the validation cohort compared to the training cohort. However, it is noteworthy that the sensitivity of the clinical model exhibited an improvement from 62 % in the training cohort to 65 % in the validation cohort. Calibration plots in Suppl. Fig. 9 represent the agreement between observed and predicted proportions.

4. Discussion

This study aimed to develop predictive models for acute and late skin toxicity following RT in BC patients by incorporating SNPs and relevant clinical factors. The models were developed through training and validation stages and included previously unexplored SNPs in non-coding genes, alongside variants previously associated with RT-induced toxicity in BC patients. The findings offer insights into potential predictors of RT side effects.

In the multivariate analysis, four SNPs were identified as independent predictors of acute toxicity, either in the training or validation cohort. However, none of these SNPs were consistently associated with toxicity across both stages of analysis (Supp. Table 8 and Suppl. Fig. 3). Of note, **rs13116075** (an intronic variant of the *LOC105377448* lncRNA) and **rs12565978** (an intergenic variant near *PLXNA2*) had been previously reported to be associated with late toxicity in Barnett GC et al. Radiotherapy and Oncology (2014) [16] (Supp. Table 4). Among these, **rs13116075** showed one of the strongest multivariable associations with overall late RT toxicity in BC patients [16]. The remaining two SNPs were in non-coding genes: **rs72550778** within *miR-34b* (Supp. Table 2) and **rs7284767** within the *TUG1* lncRNA (Supp. Table 3). Interestingly, Kishan et al. (2022) identified another SNP (rs4938723) in the promoter region of *miR-34b* that was associated with long-term genitourinary toxicity following RT in prostate cancer patients [35]. For late toxicity, two SNPs exhibited significant associations in the multivariate analysis within either the training or validation cohort (Supp. Table 12 and Suppl. Fig. 7). These included **rs16837908**, an intergenic variant near *INO80D* (Supp. Table 4) previously linked to late RT toxicity in BC patients [16], and **rs61764370**, located in the 3′UTR of *KRAS* (Supp. Table 1).

Both sets of 18 SNPs selected by Lasso for acute and late toxicity did

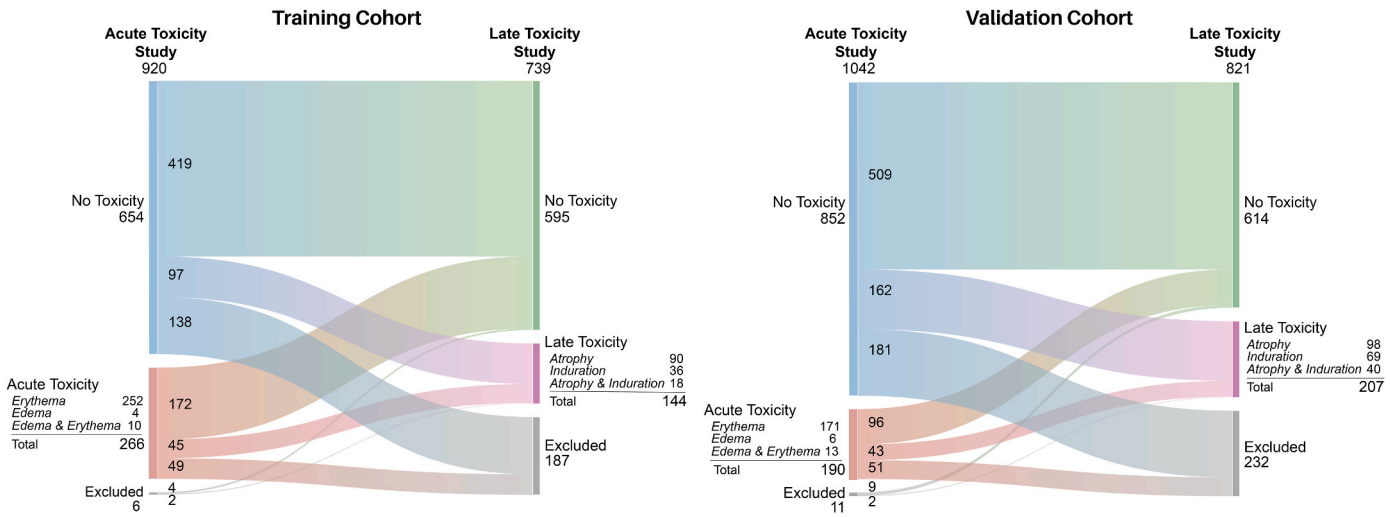


Fig. 2. Sankey diagram illustrating the distribution of acute and late toxicity among the training and validation cohorts. Patients without acute toxicity are depicted in blue, while those experiencing acute toxicity are represented in red. Patients without late toxicity are shown in green, and those with late toxicity are displayed in violet. Excluded patients are represented in grey.

not exhibit a significant association in the validation cohort, with AUC = 0.487 and accuracy = 55 % for acute toxicity and c-index = 0.54 and accuracy = 40 % for late toxicity. This suggests that the profile containing the 31 non-coding related candidate SNPs, and the 13 SNPs previously reported in the literature (Supp. Tables 1–4) confer a very low risk or no risk to develop acute or late toxicity after RT in BC patients. In a recently published radiogenomic analysis on the same REQUITE cohort as in our study, the association between toxicity following RT and 10 out of the 13 SNPs previously documented in the literature was also not confirmed [17]. Overall, these findings indicate that the existing literature on this topic may be influenced by false-positive associations, which can be attributed to factors such as small sample sizes and multiple testing. This highlights the importance of validation studies and the need for caution when interpreting the significance of individual SNPs.

Multivariate analysis of clinical and treatment variables in the validation cohort (either analysed separately or combined with genetic variants) identified the following independent predictors of toxicity after RT: breast volume and whole-breast RT dose for acute toxicity, and BMI and smoking for late toxicity (Supp. Figs. 2, 4, 6 and 8). These findings align with previous studies, emphasizing the significant role of patient and breast related factors on the development of acute and late toxicities [7–9,36,37].

Previous research has primarily assessed the correlation between clinical/treatment factors and acute skin toxicity in BC patients. However, a key limitation is the lack of validation in separate or external cohorts, and these studies have not specifically addressed the development of predictive models [7,37–39]. Two notable studies sought to fill this gap [9,40]. Rattay et al. validated in the whole REQUITE cohort known clinical risk factors for acute erythema, achieving an AUC of 0.65. In our study, a model for acute oedema and erythema using clinical variables had a significant association (OR, 1.42; CI 1.03–1.94; p = 0.030) but a moderate AUC of 0.599 during validation. Differences in our approach, where we considered both toxicities collectively, as opposed to individual assessment, as well as in the study design may explain varying AUC values. Aldraimli et al. used machine learning to predict breast acute desquamation in the REQUITE cohort, achieving an AUC of 0.77 in the validation cohort, that was created using a 50 % and 50 % cross-validation and therefore differing to our split-validation method selecting non-overlapping hospitals.

The two mentioned studies suggested that adding genetic markers to clinical factors might boost predictive performance for acute toxicity. However, in our study, including selected genetic variables did not improve the model performance, with the AUC staying at 0.597 in the validation cohort. Still, the combined model was the most accurate (62 %) when compared to clinical (55 %) and SNP (55 %) models individually (Fig. 4C).

For late RT-induced toxicities, predictive models incorporating clinical data have been created for different cancer sites, with AUC values ranging from 0.60 to 0.75 [41–43]. Notably, there are no such models tailored specifically for breast radiation toxicity. Our study aimed to fill this gap, initially achieving a promising predictive power of c-index = 0.75 using clinical variables. However, in the validation cohort, the power dropped to c-index = 0.59. Combining clinical variables and SNPs in a single model showed higher accuracy compared to separate models but also dropped to a c-index of 0.59 during validation (Fig. 6A).

Our findings, from both the acute and late studies, indicate that clinical and patient-related factors have a higher predictive value than that offered by the specific set of SNPs included in our study (Figs. 4 and 6A-B), underscoring the importance of incorporating clinical variables alongside genetic association studies [13,44].

So far, the relationship between acute and late effects has not yet been clarified [45–47]. Interestingly, our data reveals that only around 25 % of patients who experience acute toxicity subsequently develop late toxicity (Fig. 2). We included in the Lasso selection acute toxicity for predicting late toxicity and it was not selected, suggesting that acute toxicity is not a mandatory precursor for the occurrence of late toxicity for BC patients at least for the studied endpoints.

As expected, our models showed better performance in the training dataset but dropped in the validation cohort (Figs. 4 and 6). This suggests potential overfitting, where the models don't generalize well to new data. Likewise, it is worth noting that there were differences in radiation techniques between the training and validation cohorts (Tables 1 and 2). In the training cohort there was a higher prevalence of hypofractionated treatment, less IMRT and more 3D-CRT employment. Additionally, boost treatments were less common in the training cohort, especially for brachytherapy and electron-photon boost. These differences suggest that the construction of predictive models for RT side effects may benefit from subgroup analysis based on specific RT modalities.

Another limitation of this study is the inclusion of only patients of European descent, which may affect the generalizability of our findings to more diverse populations. This restriction was necessary to minimize potential confounding due to population stratification; however, it may limit the applicability of our model across different ethnic groups.

Table 1
Patient characteristics in the acute toxicity cohort.

Characteristic	Overall population N = 1962	Training population N = 920	Validation population N = 1042	Missings
Patient characteristics				
Age at RT start (years) (mean ± SD)	58.46 ± 11.09	58.25 ± 11.35	58.66 ± 10.85	4 (0.20 %)
BMI (kg/m^2) (mean ± SD)	26.47 ± 5.60	27.21 ± 6.18	25.85 ± 4.98	73 (3.72 %)
Smoker				21 (1.07 %)
Never	1087 (55.40 %)	527 (57.28 %)	560 (53.74 %)	
Ex	590 (30.07 %)	286 (31.09 %)	304 (29.17 %)	
Current	264 (13.46 %)	106 (11.52 %)	158 (15.16 %)	
Bra cup size				163 (8.31 %)
AA-A	144 (7.34 %)	61 (6.63 %)	83 (7.97 %)	
B	623 (31.75 %)	273 (29.67 %)	350 (33.59 %)	
C	538 (27.42 %)	265 (28.80 %)	273 (26.20 %)	
≥D	494 (25.18 %)	218 (23.70 %)	276 (26.49 %)	
Diabetes	113 (5.76 %)	57 (6.20 %)	56 (5.37 %)	1 (0.05 %)
Breast Cancer Phenotype				16 (0.82 %)
DCIS	190 (9.68 %)	61 (6.63 %)	129 (12.38 %)	
HER2+	55 (2.80 %)	35 (3.80 %)	20 (1.92 %)	
Luminal	424 (21.61 %)	291 (31.63 %)	133 (12.76 %)	
Luminal A	686 (34.96 %)	242 (26.3 %)	444 (42.61 %)	
Luminal B HER2+	133 (6.78 %)	66 (7.17 %)	67 (6.43 %)	
Luminal B HER2-	327 (16.67 %)	159 (17.28 %)	168 (16.12 %)	
Triple Negative	131 (6.68 %)	54 (5.87 %)	77 (7.39 %)	
Side of primary tumour				4 (0.20 %)
Left	1008 (51.38 %)	467 (50.76 %)	541 (51.92 %)	
Right	950 (48.42 %)	449 (48.8 %)	501 (48.08 %)	
Cancer treatment				
Surgery type				11 (0.56 %)
Segmentectomy/Quadrantectomy	1097 (55.91 %)	618 (67.17 %)	479 (45.97 %)	
Wide local excision	854 (43.53 %)	296 (32.17 %)	558 (53.55 %)	
Axillary surgery type				265 (13.51 %)
No axillary surgery	145 (7.39 %)	56 (6.09 %)	89 (8.54 %)	
Sentinel node biopsy	1223 (62.33 %)	500 (54.35 %)	723 (69.39 %)	
Planned axillary dissection	135 (6.88 %)	76 (8.26 %)	59 (5.66 %)	
Sentinel node biopsy + axillary dissection	194 (9.89 %)	89 (9.67 %)	105 (10.08 %)	
Chemotherapy	630 (32.11 %)	350 (38.04 %)	280 (26.87 %)	2 (0.10 %)
Neoadjuvant chemotherapy	189 (9.63 %)	115 (12.50 %)	74 (7.10 %)	4 (0.20 %)
Adjuvant chemotherapy	462 (23.55 %)	254 (27.61 %)	208 (19.96 %)	3 (0.15 %)
Tamoxifen	817 (41.64 %)	351 (38.15 %)	466 (44.72 %)	347 (17.69 %)
Anti-HER2 therapy	158 (8.05 %)	86 (9.35 %)	72 (6.91 %)	356 (18.14 %)
Whole breast RT dose (Gy) (mean ± SD)	45.28 ± 5.48	44.94 ± 4.97	45.57 ± 5.88	1 (0.05 %)
Breast volume in the CT (cm^3) (mean ± SD)	806.77 ± 500.02	876.26 ± 536.12	745.45 ± 457.40	21 (1.07 %)
RT hypofractionation	735 (37.46 %)	429 (46.63 %)	306 (29.37 %)	4 (0.20 %)
IMRT	952 (48.52 %)	340 (36.96 %)	612 (58.73 %)	3 (0.15 %)
3D-CRT	1574 (80.22 %)	869 (94.46 %)	705 (67.66 %)	21 (1.07 %)
RT boost	1358 (69.22 %)	510 (55.43 %)	848 (81.38 %)	0 (0.00 %)
RT boost type				0 (0.00 %)
Electrons	324 (16.51 %)	203 (22.07 %)	121 (11.61 %)	
Photons	701 (35.73 %)	306 (33.26 %)	395 (37.91 %)	
Brachytherapy	111 (5.66 %)	1 (0.11 %)	110 (10.56 %)	
Electrons + Photons	222 (11.31 %)	0 (0.00 %)	222 (21.31 %)	

SD: Standard deviation; RT: Radiotherapy; BMI: Body mass index; CT: computed tomography; IMRT: Intensity-modulated radiotherapy; 3D RT: 3D conformal radiation therapy.

Future studies should include more diverse cohorts to ensure broader validity and to mitigate potential biases in genetic associations.

Our study's strengths include extended, prospective follow-up beyond the conventional two years, revealing reliable long-term toxicity trends. Additionally, our development of predictive models coupled with the use of training and validation sets involving different hospitals ensures reliability, albeit at the potential cost of reduced statistical power due to smaller sample sizes. However, a limitation is our use of candidate genetic variants for SNP profiling, potentially missing the full genetic complexity of toxicity responses.

5. Conclusions

When combining clinical variables and SNPs, the model exhibited the highest predictive accuracy for acute and late toxicity, highlighting the potential benefits of integrating both types of variables for optimising model performance. However, the combined models' limited improvement in AUC or c-index during validation underlines the ongoing difficulties in developing robust, clinically meaningful SNP-based predictive models. Further research is warranted to refine/validate the predictive models, considering larger and diverse patient cohorts. Additionally, incorporating other relevant factors, such as dosimetry, additional genetic markers or SNP interactions, and using machine learning approaches may enhance the accuracy and generalizability of toxicity prediction models. Furthermore, integrating multi-omics approaches, including emerging predictive factors such as radiomic parameters, the radiation-induced lymphocyte apoptosis (RILA) test, transcriptomic data, and immune response markers, may contribute to improving risk stratification and could inform future research that complements our findings.

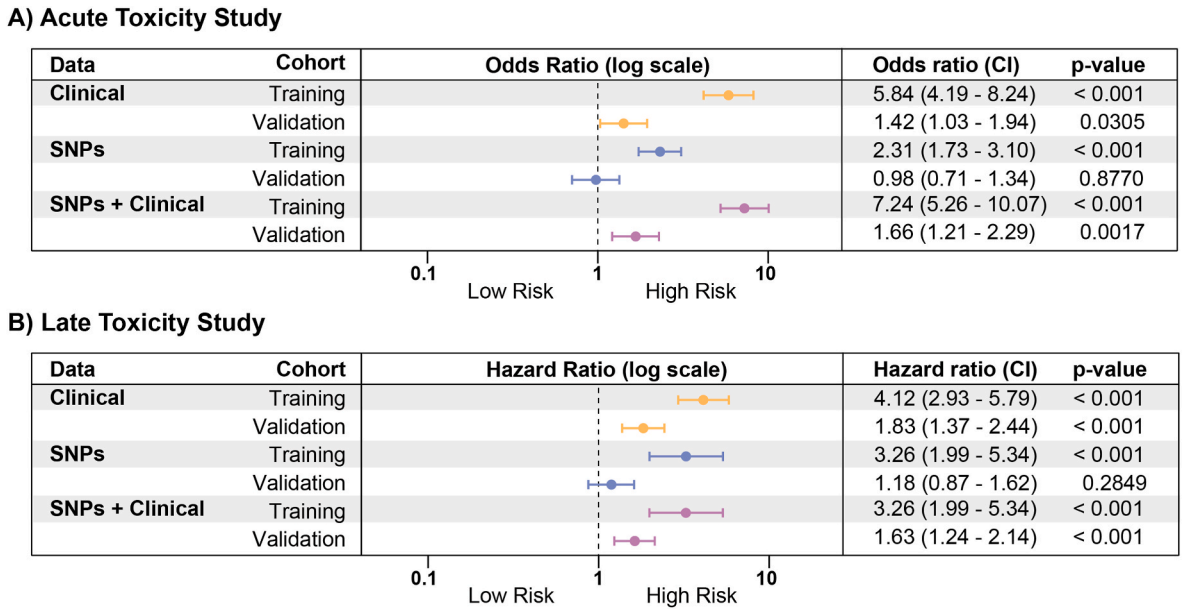


Fig. 3. Forest plots showing the association of clinical (orange), SNPs (blue) and SNPs + Clinical factors (violet) with radiation therapy-induced according to dichotomized high and low risk toxicity groups, in the training and validation cohorts from the acute toxicity study (A) and late toxicity study (B). CI: Confidence interval.

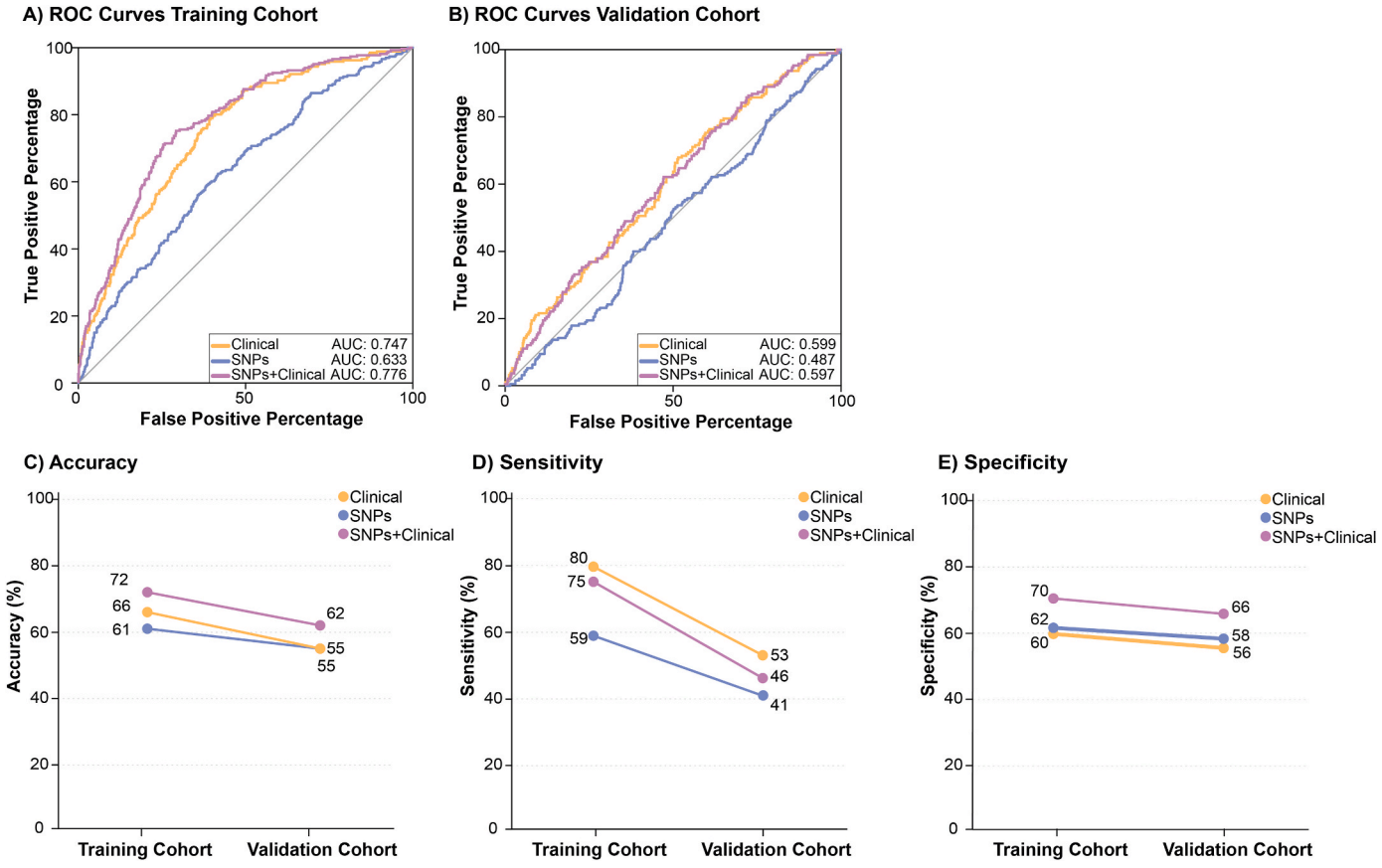


Fig. 4. Comparative performance metrics of the three predictive models (Clinical in orange, SNPs in blue, and SNPs + Clinical in violet) in training and validation cohort for acute toxicity. **A)** Receiver operating characteristic (ROC) curves in the training cohort. **B)** ROC curves in the validation cohort. **C)** Accuracy comparison in the training and validation cohorts. **D)** Sensitivity evaluation in the training and validation cohorts. **E)** Specificity assessment in the training and validation cohorts.

Table 2
Patient characteristics in the late toxicity cohort.

Characteristic	Overall population N = 1560	Training population N = 739	Validation population N = 821	Missings
Patient characteristics				
Time of follow-up (months) (median [min-max])	48 [24–96]	48 [24–84]	36 [24–96]	0 (0.00 %)
Age at RT start (years) (mean ± SD)	58.02 ± 10.79	57.64 ± 10.91	58.35 ± 10.68	4 (0.26 %)
BMI (kg/m^2) (mean ± SD)	26.31 ± 5.33	26.97 ± 5.87	25.76 ± 4.77	64 (4.10 %)
Smoker				16 (1.03 %)
Never	885 (56.73 %)	434 (58.73 %)	451 (54.93 %)	
Ex	454 (29.1 %)	223 (30.18 %)	231 (28.14 %)	
Current	205 (13.14 %)	82 (11.1 %)	123 (14.98 %)	
Bra cup size				151 (9.68 %)
AA-A	115 (7.37 %)	51 (6.9 %)	64 (7.8 %)	
B	484 (31.03 %)	217 (29.36 %)	267 (32.52 %)	
C	436 (27.95 %)	216 (29.23 %)	220 (26.8 %)	
≥D	374 (23.97 %)	162 (21.92 %)	212 (25.82 %)	
Diabetes	90 (5.77 %)	43 (5.82 %)	47 (5.72 %)	1 (0.06 %)
Breast Cancer Phenotype				15 (0.96 %)
DCIS	150 (9.62 %)	50 (6.77 %)	100 (12.18 %)	
HER2+	47 (3.01 %)	32 (4.33 %)	15 (1.83 %)	
Luminal	308 (19.74 %)	206 (27.88 %)	102 (12.42 %)	
Luminal A	560 (35.9 %)	210 (28.42 %)	350 (42.63 %)	
Luminal B HER2+	106 (6.79 %)	51 (6.9 %)	55 (6.7 %)	
Luminal B HER2-	270 (17.31 %)	139 (18.81 %)	131 (15.96 %)	
Triple Negative	104 (6.67 %)	41 (5.55 %)	63 (7.67 %)	
Side of primary tumour				3 (0.19 %)
Left	810 (51.92 %)	377 (51.01 %)	433 (52.74 %)	
Right	747 (47.88 %)	360 (48.71 %)	387 (47.14 %)	
Cancer treatment				
Surgery type				9 (0.58 %)
Segmentectomy/Quadrantectomy	897 (57.5 %)	529 (71.58 %)	368 (44.82 %)	
Wide local excision	654 (41.92 %)	206 (27.88 %)	448 (54.57 %)	
Axillary surgery type				199 (12.76 %)
No axillary surgery	112 (7.18 %)	47 (6.36 %)	65 (7.92 %)	
Sentinel node biopsy	982 (62.95 %)	410 (55.48 %)	572 (69.67 %)	
Planned axillary dissection	104 (6.67 %)	61 (8.25 %)	43 (5.24 %)	
Sentinel node biopsy + axillary dissection	163 (10.45 %)	77 (10.42 %)	86 (10.48 %)	
Chemotherapy	520 (33.33 %)	289 (39.11 %)	231 (28.14 %)	2 (0.13 %)
Neoadjuvant chemotherapy	158 (10.13 %)	95 (12.86 %)	63 (7.67 %)	4 (0.26 %)
Adjuvant chemotherapy	381 (24.42 %)	211 (28.55 %)	170 (20.71 %)	3 (0.19 %)
Tamoxifen	666 (42.69 %)	280 (37.89 %)	386 (47.02 %)	267 (17.12 %)
Anti-HER2 therapy	130 (8.33 %)	73 (9.88 %)	57 (6.94 %)	275 (17.63 %)
Whole breast RT dose (Gy) (mean ± SD)	45.53 ± 5.38	45.35 ± 4.95	45.69 ± 5.74	0 (0.00 %)
Breast volume in the CT (cm^3) (mean ± SD)	791.03 ± 468.01	855.57 ± 506.86	733.3 ± 422.33	14 (0.90 %)
RT hypofractionation	533 (34.17 %)	309 (41.81 %)	224 (27.28 %)	3 (0.19 %)
IMRT	739 (47.37 %)	246 (33.29 %)	493 (60.05 %)	1 (0.06 %)
3D-CRT	1272 (81.54 %)	700 (94.72 %)	572 (69.67 %)	14 (0.9 %)
RT boost	1124 (72.05 %)	436 (59 %)	688 (83.8 %)	0 (0.00 %)
RT boost type				0 (0.00 %)
Electrons	286 (18.33 %)	174 (23.55 %)	112 (13.64 %)	
Photons	563 (36.09 %)	260 (35.18 %)	303 (36.91 %)	
Brachytherapy	101 (6.47 %)	1 (0.14 %)	100 (12.18 %)	
Electrons + Photons	174 (11.15 %)	1 (0.14 %)	173 (21.07 %)	

SD: Standard deviation; RT: Radiotherapy; BMI: Body mass index; CT: computed tomography; IMRT: Intensity-modulated radiotherapy; 3D RT: 3D conformal radiation therapy.

Preprint

The present manuscript has been previously published as a preprint: Aguado-Flor E et al., Exploring the Potential of Combining Non-Coding Rna Polymorphisms and Clinical Data in Predicting Radiotherapy-Induced Toxicity Among Breast Cancer Patients. SSRN, Preprint posted online 13 June 2024, https://doi.org/10.2139/ssrn.4858319.

CRediT authorship contribution statement

Ester Aguado-Flor: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Victoria M. Reyes:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation. **Víctor Navarro:** Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Data curation. **Meritxell Mollà:** Writing – review & editing, Methodology, Data curation. **Miguel E. Aguado-Barrera:** Writing – review & editing, Data curation. **Manuel Altabas:** Writing – review & editing, Resources, Investigation. **David Azria:** Writing – review & editing, Supervision, Resources, Investigation, Conceptualization. **Adinda Baten:** Writing – review & editing, Resources, Investigation. **Celine Bourgier:** Writing – review & editing, Resources, Investigation. **Renée Bultijnck:** Writing – review & editing, Resources, Investigation. **Jenny Chang-Claude:** Writing – review & editing, Supervision, Resources, Investigation, Funding acquisition, Conceptualization. **Maria Carmen De Santis:** Writing – review & editing, Resources, Investigation. **Alison M. Dunning:** Writing – review

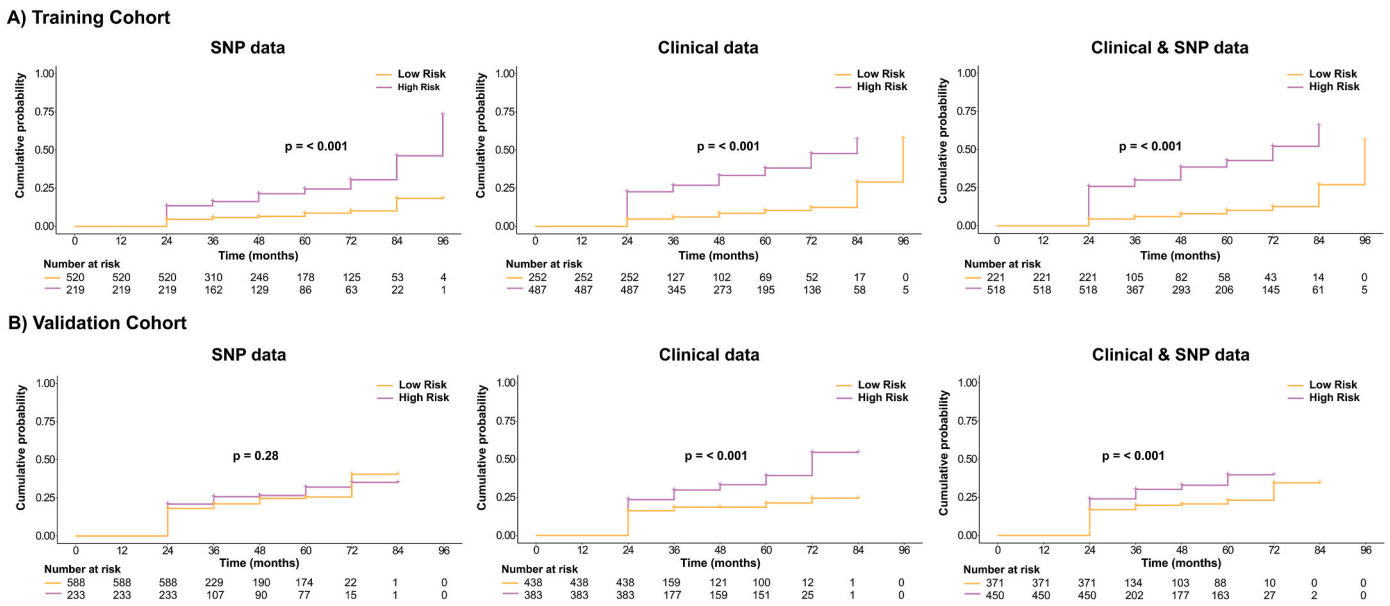


Fig. 5. Cumulative incidence curves reporting the late toxicity probability in the high risk and low risk strata. Analysis of the SNP model: training cohort of 520 high versus 219 low risk (**A, left**); validation cohort of 588 high versus 233 low risk (**B, left**). Analysis of the clinical data model: training cohort of 252 high versus 487 low risk (**A, middle**); validation cohort of 438 high versus 383 low risk **(B, middle**). Analysis of the combined model (clinical + SNP data): training cohort of 221 high versus 518 low risk (**A, right**); validation cohort of 371 high versus 450 low risk (**B, right**). P values by Cox model are reported.

& editing, Supervision, Resources, Methodology, Investigation. **Laura Duran-Lozano:** Writing – review & editing, Resources, Investigation. **Rebecca M. Elliott:** Writing – review & editing, Supervision, Project administration, Investigation. **Marie-Pierre Farcy Jacquet:** Writing – review & editing, Resources, Investigation. **Carlotta Giandini:** Writing – review & editing, Resources, Investigation. **Alexandra Giraldo:** Writing – review & editing, Resources, Investigation. **Sheryl Green:** Writing – review & editing, Resources, Investigation. **Maarten Lambrecht:** Writing – review & editing, Resources, Investigation. **Carlos Lopez-Pleguezuelos:** Writing – review & editing, Investigation. **Chris Monten:** Writing – review & editing, Resources, Investigation. **Tiziana Rancati:** Writing – review & editing, Supervision, Resources, Investigation. **Tim Rattay:** Writing – review & editing, Resources, Investigation. **Barry S. Rosenstein:** Writing – review & editing, Supervision, Resources, Investigation. **Dirk De Ruysscher:** Writing – review & editing, Supervision, Resources, Investigation. **Orland Diez:** Writing – review & editing, Investigation. **Petra Seibold:** Writing – review & editing, Supervision, Project administration, Investigation, Conceptualization. **Elena Sperk:** Writing – review & editing, Resources, Investigation. **R Paul Symonds:** Writing – review & editing, Resources, Investigation. **Hilary Stobart:** Writing – review & editing, Conceptualization. **Ana Vega:** Writing – review & editing, Supervision, Resources, Investigation. **Liv Veldeman:** Writing – review & editing, Resources, Investigation. **Guillermo Villacampa:** Writing – review & editing, Resources, Investigation, Formal analysis. **Adam J. Webb:** Writing – review & editing, Resources, Methodology, Investigation, Data curation. **Caroline Weltens:** Writing – review & editing, Resources, Investigation. **Paolo Zunino:** Writing – review & editing, Investigation. **Christopher J. Talbot:** Writing – review & editing, Supervision, Resources, Investigation, Conceptualization. **Catharine M. West:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Jordi Giralt:** Writing – review & editing, Supervision, Funding acquisition. **Sara Gutiérrez-Enríquez:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

TRIPOD checklist

The TRIPOD checklist for prediction model development and validation [48] has been completed and included in the manuscript as Supplementary Table 14.

Ethics approval and consent to participate

The study was conducted according to the Declaration of Helsinki principles. The study was approved by local ethics committees in participating countries. All patients provided informed consent prior to study participation.

Data availability

Data is available on request.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT 3.5 to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Funding information

REQUITE received funding from the European Union's Seventh Framework Programme for research, technological development, and demonstration under grant agreement no. 601826.

REQUITEplus was supported by funding from the German Federal Office for Radiation Protection (BfS; grant no. 3619S42261).

The **RADprecise** project was supported by ERA-NET ERA PerMed JCT2018-244 and by ANR, BMBF (#01KU1912), DS-CAT, FRRB, ISCIII funding.

Petra Seibold was supported by ERA PerMed JCT2018 funding (ERAPERMED2018-244, BMBF #01KU1912) and BfS funding (#3619S42261).

Tiziana Rancati was partially funded by Fondazione Italo Monzino.

Tim Rattay was funded by a National Institute of Health Research (NIHR) Clinical Lectureship (CL 2017-11-002). He was previously funded by an NIHR Doctoral Research Fellowship (DRF 2014-07-079).

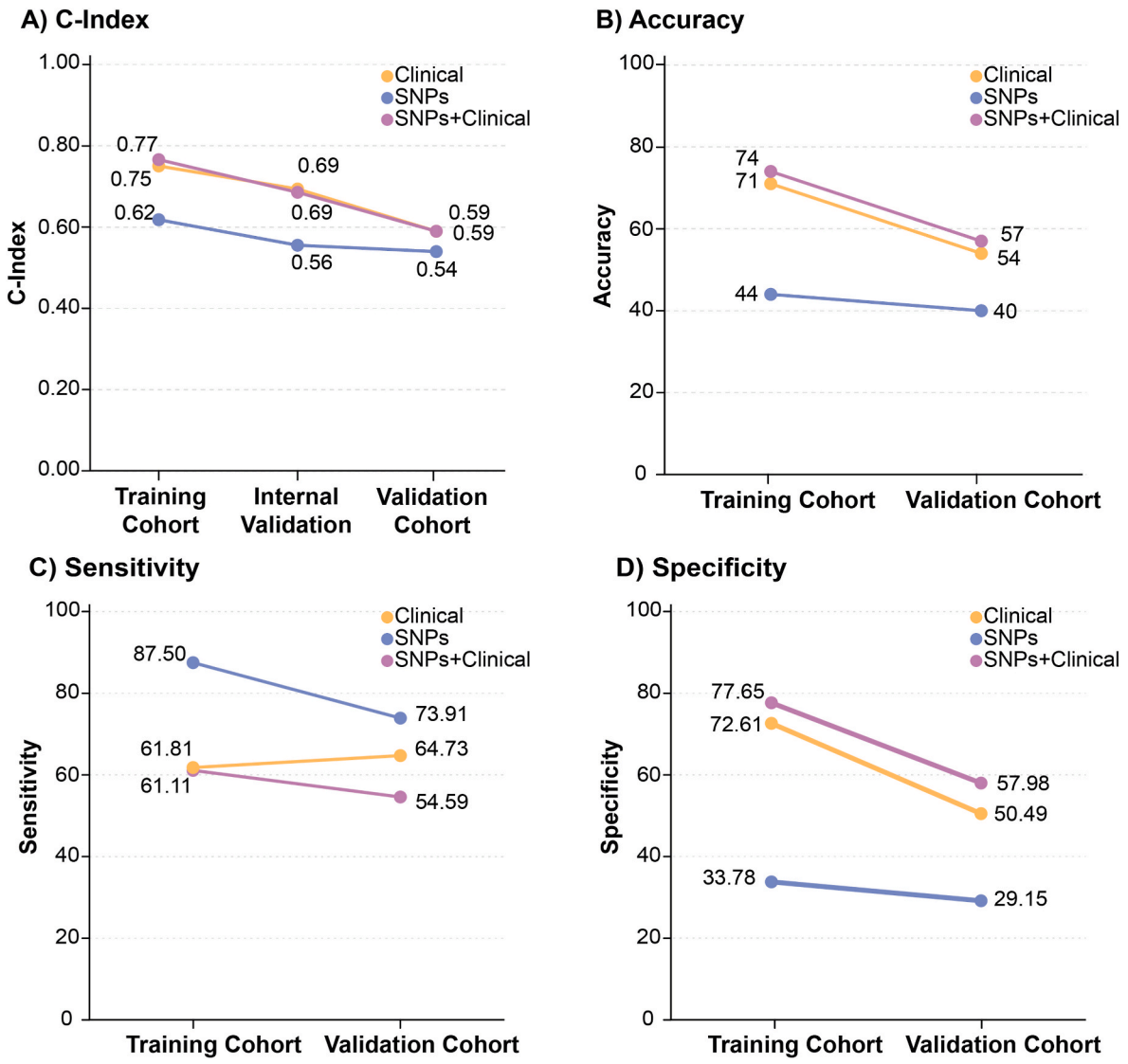


Fig. 6. Comparative performance metrics of the three predictive Models (Clinical in orange, SNPs in blue, and SNPs + Clinical in violet) in training and validation cohorts for late toxicity. A) c-index comparison in the training, internal validation and validation cohorts. **B)** Accuracy comparison in the training and validation cohorts. **C)** Sensitivity evaluation in the training and validation cohorts. **D)** Specificity assessment in the training and validation cohorts.

Ana Vega is supported by Spanish Instituto de Salud Carlos III (ISCIII) funding, an initiative of the Spanish Ministry of Economy and Innovation partially supported by European Regional Development FEDER Funds, Spain (PI22/00589, PI19/01424, PI16/00046, PI13/02030, PI10/00164; INT20/00071, INT17/00133, INT16/00154, INT15/00070), through the Autonomous Government of Galicia (Consolidation and structuring program: IN607B), by ERAPerMed JTC2018 funding (ERAPERMED2018-244) and by the AECC (PRYES211091VEGA).

Catharine West and **Rebecca Elliott** are supported by the NIHR Manchester Biomedical Research Centre and Catharine West is supported by Cancer Research UK (C1094/A18504, C147/A25254).

Sara Gutiérrez-Enríquez is supported by ERAPerMed JTC2018 funding (ERAPERMED2018-244 and SLT011/18/00005) and the Generalitat de Catalunya (2021SGR01112).

Patient recruitment in **Leicester** was supported by the Leicester Experimental Cancer Medicine Centre.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgements

We thank all patients who participated in the REQUITE study and all the members of the REQUITE project consortium in:

Belgium: Ghent University Hospital; KU Leuven.

France: ICM Montpellier, CHU Nîmes (Department of Radiation Oncology, CHU Nîmes, Nîmes, France).

Germany: Klinikum Darmstadt GmbH (Prof. Dr. Christian Weiβ and team); Zentrum für Strahlentherapie Freiburg (Dr. Petra Stegmaier and team); ViDia Christliche Kliniken Karlsruhe (Prof. Johannes Claβen and team); Universitätsklinikum Mannheim: Anette Kipke, Stefanie Kolb, and Christiane Zimmermann; Strahlentherapie Speyer (Dr. Jörg Schäfer). The researchers at DKFZ also thank Anusha Müller, Irmgard Helmbold, Juan Camilo Rosas Romero, Thomas Heger.

Italy: Fondazione IRCCS Istituto Nazionale dei Tumori, Milano; Candiolo Cancer Institute – FPO, IRCCS.

Spain: In Santiago: Complexo Hospitalario Universitario de Santiago de Compostela. In Barcelona: Vall d'Hebron Hospital Universitari, Vall d'Hebron Barcelona Hospital Campus; VHIO also acknowledge the Cellex Foundation for providing research facilities, the CERCA Programme/Generalitat de Catalunya for institutional support, and the

Agencia Estatal de Investigación for their financial support as a Center of Excellence Severo Ochoa (CEX2020-001024-S/AEI/10.13039/501100011033). VHIO authors would like to express their gratitude to Joanna Domènech-Vivó for her contribution in annotating genetic variants.

UK: University Hospitals of Leicester NHS Trust: Theresa Beaver, Sara Barrows, Monika Kaushik, Frances Kenny, Jaroslaw Krupa, Kelly V Lambert, Simon M Pilgrim, Sheila Shokuhi, Kalliope Valassidou, Kiran Kancherla, Kufre Sampson, Ahmed Osman and Kaitlin Walker. This publication presents independent research funded by the NIHR. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

USA: Mount Sinai Hospital, New York.

REQUITE consortium

Rita Aerts[a], Donna Appleton[b], Soumia Arredouani[a], Barbara Avuzzi[c], Luis Aznar-Garcia[d], Gillian C. Barnett[e], Ion Bioangiu[d], Thomas Blaschke[f], Francoise Bons[g], Muriel Brengues[h], Erik Briers[i], Patricia Calvo-Crespo[j,k], Ana Carballo[j,k], Ananya Choudhury[l], Alessandro Cicchetti[m], Johannes Claßen[n], Charlotte E. Coles[e], Wilfried De Neve[o,p], Gilles Defraene[a], Elena Delmastro[q], Pieter Deseyne[o,p], Hannah Dobbelaere[r], Isabel Dominguez-Rios[k], Roxana Draghici[h], Fréderic Duprez[o,p], Michael Ehmann[s], Elhaseen Elhamin[o,p], Claire P. Esler[t], Laura Fachal[u], Corinne Faivre-Finn[v], Irene Fajardo-Paneque[k], Juan Fernández-Tajes[w,j], Valérie Fonteyne[o,p], Karen Foweraker[t], Marzia Franceschini[x], Olivia Fuentes-Rios[w,j], Pietro Gabriele[q], David García-Relancio[y], Elisabetta Garibaldi[q], Benjamin Gauter-Fleckenstein[s], Tommaso Giandini[z], Ulrich Giesche[n], Frank Giordano[s], Ava Golchin[aa], Antonio Gomez-Caamano[j,k], Victoria Harrop[ab], Karin Haustermans[a], Irmgard Helmbold[ac], Carsten Herskind[s], Daniel S. Higginson[ad], Kerstie Johnson[ae], Kiran Kancherla[d], Monika Kaushik[b], Manjusha Keni[af], Frances Kenny[b], Christopher Kent[d], Sarah L. Kerns[ag], Hazem Khout[b], Jaroslaw Krupa[b], Kelly V. Lambert[b], Maarten Lambrecht[a], Samuel Lavers[r], William Li[ah], Yolande Lievens[o,p], Ramon Lobato-Busto[ai], Laura Lozza[c], Marcus Mareel[o,p], Sara Morlino[c], Anusha Müller[ac], Burkhard Neu[aj], Ahmed Osman[d], Piet Ost[o,p], Leen Paelinc[o,p], Abigail Pascoe[t], Debbie Payne[ak], Stephanie Peeters[a], Paula Peleteiro[j,k], Gabriele Pietro[q], Simon Pilgrim[b], Bibiana Piqué-Leiva[y], Giselle Post[p], Monica Ramos[y], Belina Rodriguez-Lage[w,j], Kufre Sampson[d], Claudia Sangalli[c], MarieLuise Sautter-Bihl[aj], Jörg Schäfer[al], Thomas Schnabel[am], Alejandro Seoane-Ramallo[an], Sheila Shokuhi[b], Paloma Sosa-Fajardo[j,k], Thiagarajan Sridhar[d], Petra Stegmaier[ao], Richard G. Stock[aa], Richard Stock[aa], Martijn Swimberghe[o,p], Begoña Taboada-Valladares[j,k], Sridhar Thiagarajan[d], Laura Torrado Moya[j,k], Kalliope Valassiadou[b], Riccardo Valdagni[c,ap,m], Marc van Eijkeren[o,p], Annick Van Greveling[p], Erik van Limberghen[a], Tom Vercauteren[o,p], Ben Vanneste[aq], Subramaniam Vasanthan[d], Marlon Veldwijk[s], Katrien Vandecasteele[o,p], Tim Ward[ar], Christian Weiß[as], Christian Weißenberger[ao], Simon Wright[d]

[a]Department of Radiation Oncology, University Hospitals Leuven/KU Leuven, Leuven, Belgium

[b]Department of Breast Surgery, Glenfield Hospital, University Hospitals of Leicester NHS Trust, Leicester, UK

[c]Department of Radiation Oncology 1, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[d]Department of Oncology, Leicester Royal Infirmary, University Hospitals of Leicester NHS Trust, Leicester, UK

[e]Department of Oncology, University of Cambridge, Cambridge, UK

[f]Praxis für Strahlentherapie, Baden-Baden, Germany

[g]Fédération Universitaire d'Oncologie Radiothérapie d'Occitanie Méditérranée, Département d'Oncologie Radiothérapie, CHU Carémeau, Nîmes, France

[h]Fédération Universitaire d'Oncologie Radiothérapie d'Occitanie Méditérranée, Département d'Oncologie Radiothérapie, ICM Montpellier, INSERM U1194 IRCM, University of Montpellier, Montpellier, France

[i]Patient advocate, Hasselt, Belgium

[j]Instituto de Investigación Sanitaria de Santiago de Compostela, Santiago de Compostela, Spain

[k]Department of Radiation Oncology, Complexo Hospitalario Universitario de Santiago, Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Spain

[l]Translational Radiobiology Group, Division of Cancer Sciences, University of Manchester, Manchester Academic Health Science Centre, The Christie NHS Foundation Trust, Manchester, UK

[m]Prostate Cancer Program, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[n]Klinik für Strahlentherapie, St. Vincentius-Kliniken gAG, Karlsruhe, Germany

[o]Department of Radiation Oncology, Ghent University Hospital, Ghent, Belgium

[p]Department of Human Structure and Repair, Ghent University, Ghent, Belgium

[q]Department of Radiation Oncology, Fondazione del Piemonte per l'Oncologia Candiolo Cancer Institute, Candiolo, Italy

[r]Department of Genetics and Genome Biology, University of Leicester, Leicester, UK

[s]Department of Radiation Oncology, Universitätsklinikum Mannheim, Medical Faculty Mannheim, Heidelberg University, Heidelberg, Germany

[t]City Hospital, Nottingham University Hospitals NHS Trust, Nottingham, UK

[u]Centre for Cancer Genetic Epidemiology, Department of Oncology, Strangeways Research Laboratory, University of Cambridge, Cambridge, UK

[v]Division of Cancer Sciences, University of Manchester, Manchester, UK

[w]Fundación Pública Galega de Medicina Xenómica, Grupo de Medicina Xenómica (USC), Santiago de Compostela, Spain

[x]Department of Radiotherapy 2, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[y]Radiation Oncology Department, Vall d'Hebron University Hospital, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain

[z]Department of Medical Physics, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[aa]Department of Radiation Oncology, The Icahn School of Medicine at Mount Sinai, New York, US

[ab]Queen Elizabeth Hospital, University Hospitals Birmingham NHS Trust, Birmingham, UK

[ac]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

[ad]Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, US

[ae]Cancer Research Centre, University of Leicester, Leicester, UK

[af]Department of Oncology, Derby Teaching Hospitals NHS Foundation Trust, Derby, UK

[ag]Departments of Radiation Oncology and Surgery, University of Rochester Medical Center, Rochester, US

[ah]Queens Hospital Center, New York, US

[ai]Department of Medical Physics, Complexo Hospitalario Universitario de Santiago, Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Spain

[aj]Klinik für Radioonkologie und Strahlentherapie, Städtisches Klinikum, Karlsruhe, Germany

[ak]Centre for Integrated Genomic Medical Research (CIGMR), Manchester, UK

[al]Strahlentherapie Speyer, Speyer, Germany

[am]Klinik für Strahlentherapie und Radiologische Onkologie, Klinikum der Stadt Ludwigshafen gGmbH, Ludwigshafen, Germany

[an]Medical Physics Department, Vall d'Hebron University Hospital, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain

[ao]Zentrum für Strahlentherapie, Karlsruhe, Germany

[ap]Department of Hematology and Hemato-Oncology, Università

degli Studi di Milano, Milan, Italy
[aq]MAASTRO Clinic, GROW School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, Netherlands
[ar]Patient advocate, Pelvic Radiation Disease Association, UK
[as]Klinikum Darmstadt, Institut für Radionkologie und Strahlentherapie, Darmstadt, Germany.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.breast.2025.104506.

References

[1] Darby S, McGale P, Correa C, Taylor C, Arriagada R, Clarke M, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10 801 women in 17 randomised trials. Lancet 2011;378:1707–16. https://doi.org/10.1016/S0140-6736(11)61629-2.
[2] Popanda O, Marquardt JU, Chang-Claude J, Schmezer P. Genetic variation in normal tissue toxicity induced by ionizing radiation. Mutat Res 2009;667:58–69. https://doi.org/10.1016/j.mrfmmm.2008.10.014.
[3] Xu L, Osei B, Osei E. A review of radiation genomics: integrating patient radiation response with genomics for personalised and targeted radiation therapy. J Radiother Pract 2019;18:198–209. https://doi.org/10.1017/S1460396918000547.
[4] Kerns SL, West CML, Andreassen CN, Barnett GC, Bentzen SM, Burnet NG, et al. Radiogenomics: the search for genetic predictors of radiotherapy response. Future Oncol 2014;10:2391–406. https://doi.org/10.2217/fon.14.173.
[5] Bentzen SM. Preventing or reducing late side effects of radiation therapy: radiobiology meets molecular pathology. Nat Rev Cancer 2006;6:702–13. https://doi.org/10.1038/nrc1950.
[6] Scaife JE, Barnett GC, Noble DJ, Jena R, Thomas SJ, West CMLL, et al. Exploiting biological and physical determinants of radiotherapy toxicity to individualize treatment. Br J Radiol 2015;88:2015017. https://doi.org/10.1259/bjr.20150172.
[7] Ciammella P, Podgornii A, Galeandro M, Micera R, Ramundo D, Palmieri T, et al. Toxicity and cosmetic outcome of hypofractionated whole-breast radiotherapy: predictive clinical and dosimetric factors. Radiat Oncol 2014;9:97. https://doi.org/10.1186/1748-717X-9-97.
[8] Barnett GC, Wilkinson JS, Moody AM, Wilson CB, Twyman N, Wishart GC, et al. The cambridge breast intensity-modulated radiotherapy trial: patient- and treatment-related factors that influence late toxicity. Clin Oncol 2011;23:662–73. https://doi.org/10.1016/j.clon.2011.04.011.
[9] Rattay T, Seibold P, Aguado-Barrera ME, Altabas M, Azria D, Barnett GC, et al. External validation of a predictive model for acute skin radiation toxicity in the REQUITE breast cohort. Front Oncol 2020;10:575909. https://doi.org/10.3389/fonc.2020.575909.
[10] Twardella D, Popanda O, Helmbold I, Ebbeler R, Benner A, von Fournier D, et al. Personal characteristics, therapy modalities and individual DNA repair capacity as predictive factors of acute skin toxicity in an unselected cohort of breast cancer patients receiving radiotherapy. Radiother Oncol 2003;69:145–53. https://doi.org/10.1016/S0167-8140(03)00166-X.
[11] De Langhe S, Mulliez T, Veldeman L, Remouchamps V, van Greveling A, Gilsoul M, et al. Factors modifying the risk for developing acute skin toxicity after whole-breast intensity modulated radiotherapy. BMC Cancer 2014;14:711. https://doi.org/10.1186/1471-2407-14-711.
[12] Tortorelli G, Di Murro L, Barbarino R, Cicchetti S, di Cristino D, Falco MD, et al. Standard or hypofractionated radiotherapy in the postoperative treatment of breast cancer: a retrospective analysis of acute skin toxicity and dose inhomogeneities. BMC Cancer 2013;13:230. https://doi.org/10.1186/1471-2407-13-230.
[13] Kerns SL, Kundu S, Oh JH, Singhal SK, Janelsins M, Travis LB, et al. The prediction of radiotherapy toxicity using single nucleotide polymorphism-based models: a step toward prevention. Semin Radiat Oncol 2015;25:281–91. https://doi.org/10.1016/j.semradonc.2015.05.006.
[14] Franco NR, Massi MC, Ieva F, Manzoni A, Paganoni AM, Zunino P, et al. Development of a method for generating SNP interaction-aware polygenic risk scores for radiotherapy toxicity. Radiother Oncol 2021;159:241–8. https://doi.org/10.1016/j.radonc.2021.03.024.
[15] Andreassen CN. The future has begun in radiogenomics. Radiother Oncol 2014;111:165–7. https://doi.org/10.1016/j.radonc.2014.04.006.
[16] Barnett GC, Thompson D, Fachal L, Kerns S, Talbot C, Elliott RM, et al. A genome wide association study (GWAS) providing evidence of an association between common genetic variants and late radiotherapy toxicity. Radiother Oncol 2014;111:178–85. https://doi.org/10.1016/j.radonc.2014.02.012.
[17] Jandu HK, Veal CD, Fachal L, Luccarini C, Aguado-Barrera ME, Altabas M, et al. Genome-wide association study of treatment-related toxicity two years following radiotherapy for breast cancer. Radiother Oncol 2023;187:109806. https://doi.org/10.1016/J.RADONC.2023.109806.
[18] Naderi E, Crijns APG, Steenbakkers RJHM, van den Hoek JGM, Boezen HM, Alizadeh BZ, et al. A two-stage genome-wide association study of radiation-induced acute toxicity in head and neck cancer. J Transl Med 2021;19:481. https://doi.org/10.1186/s12967-021-03145-1.
[19] Naderi E, Aguado-Barrera ME, Schack LMH, Dorling L, Rattay T, Fachal L, et al. Large-scale meta-GWAS reveals common genetic factors linked to radiation-induced acute toxicities across cancers. JNCI Cancer Spectr 2023;M. https://doi.org/10.1093/JNCICS/PKAD088. ACCEPTED.
[20] Kerns SL, Fachal L, Dorling L, Barnett GC, Baran A, Peterson DR, et al. Radiogenomics consortium genome-wide association study meta-analysis of late toxicity after prostate cancer radiotherapy. J Natl Cancer Inst 2020;112:179–90. https://doi.org/10.1093/jnci/djz075.
[21] Herskind C, Talbot CJ, Kerns SL, Veldwijk MR, Rosenstein BS, West CML. Radiogenomics: a systems biology approach to understanding genetic risk factors for radiotherapy toxicity? Cancer Lett 2016;382:95–109. https://doi.org/10.1016/j.canlet.2016.02.035.
[22] Seibold P, Webb A, Aguado-Barrera ME, Azria D, Bourgier C, Brengues M, et al. REQUITE: a prospective multicentre cohort study of patients undergoing radiotherapy for breast, lung or prostate cancer. Radiother Oncol 2019;138:59–67. https://doi.org/10.1016/j.radonc.2019.04.034.
[23] van Buuren S, Groothuis-Oudshoorn K, , From with contributions, Vink G, Schouten R, Robitzsch A. Package "mice": multivariate imputation by chained equations. CRAN Repository; 2021. p. 1–188.
[24] Cohen J. Statistical power analysis for the behavioural sciences. Hillside. NJ: Lawrence Earlbaum Associates; 1988.
[25] Tibshirani R. Regression shrinkage and selection via the Lasso. J Roy Stat Soc B 1996;58:267–88. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
[26] Batenburg MCT, Bartels M, Maarse W, Witkamp A, Verkooijen HM, van den Bongard HJGD. Factors associated with late local radiation toxicity after post-operative breast irradiation. Breast J 2022;2022:6745954. https://doi.org/10.1155/2022/6745954.
[27] Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32–5. https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-35.
[28] Efron B, Tibshirani RJ. An introduction to the bootstrap. https://doi.org/10.1201/9780429246593; 1994.
[29] Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA, J Am Med Assoc 1982;247:2543–6. https://doi.org/10.1001/jama.1982.03320430047030.
[30] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44. https://doi.org/10.2307/2531595.
[31] Kang L, Chen W, Petrick NA, Gallas BD. Comparing two correlated C indices with right-censored survival outcome: a one-shot nonparametric approach. Stat Med 2015;34. https://doi.org/10.1002/sim.6370.
[32] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology (Camb, Mass) 2010;21:128–38. https://doi.org/10.1097/EDE.0b013e3181c30fb2.
[33] Cox DR, Snell EJ. A general definition of residuals. J Roy Stat Soc B 1968;30:248–65. https://doi.org/10.1111/j.2517-6161.1968.tb00724.x.
[34] Barnett GC, Kerns SL, Noble DJ, Dunning AM, West CML, Burnet NG. Incorporating genetic biomarkers into predictive models of normal tissue toxicity. Clin Oncol 2015;27:579–87. https://doi.org/10.1016/j.clon.2015.06.013.
[35] Kishan AU, Marco N, Schulz-Jaavall M-B, Steinberg ML, Tran PT, Juarez JE, et al. Germline variants disrupting microRNAs predict long-term genitourinary toxicity after prostate cancer radiation. Radiother Oncol 2022;167:226–32. https://doi.org/10.1016/j.radonc.2021.12.040.
[36] Brunt AM, Wheatley D, Yarnold J, Somaiah N, Kelly S, Harnett A, et al. Acute skin toxicity associated with a 1-week schedule of whole breast radiotherapy compared with a standard 3-week regimen delivered in the UK FAST-Forward Trial. Radiother Oncol 2016;120:114–8. https://doi.org/10.1016/j.radonc.2016.02.027.
[37] Allali S, Carton M, Everhard S, Rivera S, Ghannam Y, Peignaux K, et al. CANTO skin: evaluation of skin toxicity risk factors in patients treated for breast cancer. Int J Cancer 2023;153:1–12. https://doi.org/10.1002/ijc.34664.
[38] Deantonio L, Gambaro G, Beldì D, Masini L, Tunesi S, Magnani C, et al. Hypofractionated radiotherapy after conservative surgery for breast cancer: analysis of acute and late toxicity. Radiat Oncol 2010;5:112. https://doi.org/10.1186/1748-717X-5-112.
[39] Back M. Impact of radiation therapy on acute toxicity in breast conservation therapy for early breast cancer. Clin Oncol 2004;16:12–6. https://doi.org/10.1016/j.clon.2003.08.005.
[40] Aldraimli M, Osman S, Grishchuck D, Ingram S, Lyon R, Mistry A, et al. Development and optimization of a machine-learning prediction model for acute desquamation after breast radiation therapy in the multicenter REQUITE cohort. Adv Radiat Oncol 2022;7:100890. https://doi.org/10.1016/j.adro.2021.100890.
[41] D'Avino V, Palma G, Liuzzi R, Conson M, Doria F, Salvatore M, et al. Prediction of gastrointestinal toxicity after external beam radiotherapy for localized prostate cancer. Radiat Oncol 2015;10:80. https://doi.org/10.1186/s13014-015-0389-5.
[42] Tomatis S, Rancati T, Fiorino C, Vavassori V, Fellin G, Cagna E, et al. Late rectal bleeding after 3D-CRT for prostate cancer: development of a neural-network-based predictive model. Phys Med Biol 2012;57:1399–412. https://doi.org/10.1088/0031-9155/57/5/1399.
[43] De Langhe S, De Meerleer G, De Ruyck K, Ost P, Fonteyne V, De Neve W, et al. Integrated models for the prediction of late genitourinary complaints after high-dose intensity modulated radiotherapy for prostate cancer: making informed decisions. Radiother Oncol 2014;112:95–9. https://doi.org/10.1016/j.radonc.2014.04.005.
[44] Andreassen CN, Barnett GC, Langendijk JA, Alsner J, De Ruysscher D, Krause M, et al. Conducting radiogenomic research – do not forget careful consideration of

the clinical data. Radiother Oncol 2012;105:337–40. https://doi.org/10.1016/j.radonc.2012.11.004.
[45] Burnet NG, Wurm R, Nyman J, Peacock JH. Normal tissue radiosensitivity - how important is it? Clin Oncol 1996;8:25–34. https://doi.org/10.1016/S0936-6555(05)80035-4.
[46] Tucker SL, Turesson I, Thames HD. Evidence for individual differences in the radiosensitivity of human skin. Eur J Cancer 1992;28A:1783–91. https://doi.org/10.1016/0959-8049(92)90004-L.
[47] Bentzen SM, Overgaard M, Overgaard J. Clinical correlations between late normal tissue endpoints after radiotherapy: implications for predictive assays of radiosensitivity. Eur J Cancer 1993;29A:1373–6. https://doi.org/10.1016/0959-8049(93)90004-Y.
[48] Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med 2015;162. https://doi.org/10.7326/M14-0697.