XRCC1 polymorphism associated with late toxicity after radiotherapy in breast cancer patients

Petra Seibold, PhD*, Sabine Behrens, PhD*, Peter Schmezer, PhD[†], Irmgard Helmbold*, Gillian Barnett, MD, PhD^{††}, Charlotte Coles, MD^{††}, John Yarnold, MD^{||}, Christopher J. Talbot, PhD[¶], Takashi Imai, PhD[#], David Azria, MD, PhD^{**}, C. Anne Koch, MD, PhD^{††††}, Alison M. Dunning, PhD^{|||}, Neil Burnet, PhD^{¶¶}, Judith M. Bliss, MSc^{¶¶¶¶}, R Paul Symonds, MD^{##}, Tim Rattay^{##}, Tomo Suga[#], Sarah L. Kerns, PhD, MPH^{***}, Celine Bourcier^{**}, Katherine A. Vallis, PhD^{††††}, Marie-Luise Sautter-Bihl, MD^{|||}, Johannes Claßen, MD^{¶¶¶}, Juergen Debus, MD^{###}, Thomas Schnabel, MD^{****}, Barry S. Rosenstein, PhD^{####}, Frederik Wenz, MD^{†††††}, Catharine M. West, PhD^{|||||}, Odilia Popanda, PhD^{†,****}, Jenny Chang-Claude, PhD, MHS^{*,*****}

***** *Shared senior authorship*

* Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

† Division of Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Heidelberg, Germany

†† Department of Oncology, Oncology Centre, Cambridge University Hospital NHS Foundation Trust, UK

|| Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, UK

¶ Department of Genetics, University of Leicester, Leicester, UK

Advanced Radiation Biology Research Program, National Institute of Radiological Sciences, Chiba, Japan

** Department of Radiation Oncology and Medical Physics, I.C.M. – Institut regional du Cancer Montpellier, France

††† Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Canada

||| Centre for Cancer Genetic Epidemiology, University of Cambridge, Strangeways Research Laboratory, UK

¶¶ University of Cambridge, Department of Oncology, Oncology Centre, Cambridge University Hospital NHS Foundation Trust, UK

Department of Cancer Studies and Molecular Medicine, University of Leicester, Leicester, UK

*** Rubin Center for Cancer Survivorship, Department of Radiation Oncology, University of Rochester Medical Center, Rochester, NY, USA (present) and Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, USA

†††† CRUK/MRC Oxford Institute for Radiation Oncology, Oxford University, UK

||||| Department of Radiation Oncology, Staedisches Klinikum Karlsruhe, Karlsruhe, Germany

†††† Clinic for radiotherapy and radiation oncology, St. Vincentius-Kliniken gAG, Karlsruhe, Germany

Department of Radiation Oncology, University of Heidelberg, Heidelberg, Germany

**** Clinic for Radiotherapy and Radiation Oncology Klinikum Ludwigshafen, Germany

Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, USA

††††† Department of Radiation Oncology, University Medical Centre Mannheim, University of Heidelberg, Mannheim, Germany

||||||| Institute of Cancer Sciences, The University of Manchester, Manchester Academic Health Science Centre, The Christie, UK

††††† The Institute of Cancer Research, Clinical Trials and Statistics Unit, Sutton, UK

Running title: DNA repair and risk of late radiotherapy toxicity

Corresponding author: Prof. Dr. Jenny Chang-Claude

Division of Cancer Epidemiology (C020)

German Cancer Research Center (DKFZ)

Im Neuenheimer Feld 581

69120 Heidelberg, Germany

Tel.: +49-6221-42-2373

Fax: +49-6221-42-2203

E-mail: j.chang-claude@dkfz.de

Acknowledgements

We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out. We like to

thank Ursula Eilber, Christina Krieg, Muhabbet Celik, and Peter Waas for the most valuable technical assistance; the DKFZ Core Facility (Matthias Schick, Roger Fischer) and Dr. Rainer Claus (Division of Epigenomics and Cancer Risk Factors, DKFZ, Heidelberg) for genotyping (MARIE/ISE); Leila Dorling and Rebecca Elliott for kind assistance with RAPPER genotype data request. BRRG thank Yoshimi Shoji, Atsuko Ishikawa, and Keiko Ogawa for their technical assistance. CIHR thank Vibha Sharma for technical assistance. LeND thank clinicians and radiotherapists in Leicester, Nottingham and Derby and the patients.

Funding

MARIE_{RAD} cohort, Germany: Dietmar Hopp Stiftung (#23017006) and Deutsche Krebshilfe (#108253 and #108419). **ISE-2 cohort**, Germany: Federal Office for Radiation Protection (projects St Sch. 4116 and 4233), and USAMRMC BCRP DAMD17-02-10500. **BRRG cohort**, Japan: NIRS. **CIHR cohort**, Canada: Canadian Institutes of Health Research (MOP-77703). **COHORT cohort**, France: Institutional scientific grant support from Novartis. **LeND cohort**, UK: Breast Cancer Campaign. **RAPPER**, UK: Cancer Research UK (C1094/AA11728, C8197/S10123) and Experimental Cancer Medicine funding. **IMRT trial**, UK: Breast Cancer Campaign. **RACE cohort**, UK: Cancer Research UK, NIHR Biomedical Research Centre. We acknowledge NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden and The Institute of Cancer Research. **Radiogenomics Consortium**: Grants RSGT-05-200-01-CCE from the American Cancer Society, PC074201 from the Department of Defense and 1R01CA134444 from the National Institutes of Health to BSR.

Conflict of interest: none.

Keywords: breast cancer, base excision repair, genetic variants, radiation toxicity

Abbreviations: BER: Base excision repair; GWAS: genome-wide association studies; HWE: Hardy-Weinberg equilibrium; RGC: Radiogenomics Consortium; STAT: Standardized Total Average Toxicity; XRCC1: X-ray repair cross-complementing protein 1

Summary (max 75 words; now 72):

The study aimed to identify genetic variants in oxidative-stress-related genes associated with risk of radiotherapy-related late toxicities. A two stage design with discovery and replication phase was used, including up to 2600 breast cancer patients from eight studies. A genetic variant in *XRCC1* was confirmed to be associated with a significantly decreased risk of late toxicity after radiotherapy. Such variants could be incorporated into clinical tests for prediction of late radiation toxicity.

***XRCC1* polymorphism associated with late toxicity after radiotherapy in breast cancer patients**

Abstract (max 300 words; now 260)

PURPOSE: Risk of normal tissue toxicity limits the radiation doses that can be safely prescribed in radiotherapy. The role of common genetic variants as predictive factors for toxicity is still undefined. This investigation aimed to identify single nucleotide polymorphisms (SNPs) in oxidative-stress-related genes associated with risk of late toxicities in breast cancer patients receiving radiotherapy.

METHODS AND MATERIALS: Using a two stage design, 305 SNPs in 59 candidate genes were investigated in the discovery phase in 753 breast cancer patients from two prospective cohorts from Germany. The ten most promising SNPs in four genes were evaluated in the replication phase in up to 1 883 breast cancer patients from six cohorts identified through the XXX Consortium. Outcomes of interest were late skin toxicity and fibrosis of the breast as well as an overall toxicity score (Standardized Total Average Toxicity, STAT). Multivariable logistic and linear regression models were used to assess associations between SNPs and late toxicity. A meta-analysis approach was used to summarize evidence.

RESULTS: The association of a genetic variant in the base excision repair gene *XRCC1*, rs2682585, with normal tissue late radiation toxicity was replicated in all tested studies. In the combined analysis of discovery and replication cohorts, carrying the rare allele was associated with a significantly lower risk of skin toxicities (multivariate OR 0.77, 95% CI 0.61-0.96, $p=0.02$) and a decrease in STAT scores (-0.08, 95% CI -0.15, -0.02, $p=0.016$).

CONCLUSIONS: Using a stage design with replication, we identified a variant allele in the base excision repair gene *XRCC1*, which could be used in combination with additional variants for developing a test to predict late toxicities after radiotherapy in breast cancer patients.

1 Introduction

2 The majority of breast cancer patients receive radiotherapy as part of their treatment. Some
3 of them will develop moderate to severe normal tissue late side effects due to irradiation (1,
4 2), which can substantially impact on their quality of life (3). Identification of biomarkers of
5 radiosensitivity is important to develop predictive tests that can be used alongside clinical
6 and treatment parameters to individualise the management of high-risk patients and so
7 improve the quality of life of cancer survivors. A number of assays for predicting radiation
8 toxicity have been developed, but none has proved suitable for clinical application. The role
9 of genetic variants as predictive factors for toxicity in breast cancer patients is still undefined.
10 Most studies suffered from limited sample size and lack of replication (4, 5). Therefore, the
11 XXX Consortium was established in 2009 to facilitate collaboration, to pool data, and to
12 share expertise (6). A recent paper highlighted that common variants are likely to have small
13 individual effect sizes, and so large patient cohorts must be studied (7, 8). In support of this
14 suggestion, a XXX Consortium collaborative project involving 2,036 patients was able to
15 replicate an association between genetic variations near *TNF* and radiotherapy toxicity in
16 breast cancer patients (9). Genome-wide association studies (GWAS) have recently
17 identified new candidate variants with no a priori assumptions of which genes might be
18 important, although they also highlight the large sample sizes needed (10, 11).

19 Radiotherapy causes oxidative stress by producing reactive oxygen species, which damages
20 DNA, and causes cell death. Therefore, it was hypothesized that genetic variants in oxidative
21 stress-related genes involved in defense against reactive species and in DNA repair
22 mechanisms may modify the ability of irradiated cells/tissues to recover from oxidative
23 damage. As a consequence, the genetic variants may affect a cancer patient's susceptibility
24 to develop radiotherapy-induced late side effects such as fibrosis and skin toxicities.

25 Previous studies primarily concentrated on single nucleotide polymorphisms (SNPs) in the
26 coding region of candidate genes rather than a comprehensive coverage of genetic variation.
27 GWAS of cancer predisposition however identified SNPs in non-coding regions, which are

now known to be functionally important. A novel aspect of our approach, therefore, was to investigate a panel of more than 300 SNPs (mainly tagging SNPs) to cover comprehensively genetic variability across 59 candidate genes. In addition, we used a staged design for independent replication of observed associations. STROGAR guidelines for reporting radiogenomics studies were followed throughout (12).

Methods

Study design: A two-stage design with discovery (two study cohorts) and replication phase (six study populations) was used (**Figure 1**). Two xxx cohorts were used in the first phase to increase power and to use “internally” consistent first results for selecting SNPS for replication. Of the large panel of SNPs investigated in the discovery phase, SNP associations with p-values <0.05 for late skin toxicities or fibrosis of the breast in at least one of the discovery cohorts and associations in the same direction in both discovery cohorts were selected for replication. For “external” replication, independent study cohorts were sought and selected on the basis of having at least 100 patients and at least one year of follow-up.

Patient populations: The discovery sample comprised two patient cohorts of unilateral primary breast cancer (invasive and *in situ*) treated with radiotherapy in the Rhine-Neckar-Karlsruhe region of southern Germany: (1) the A study in which patients were diagnosed and treated with radiotherapy 2002-2005 and (2) the B study in which patients were diagnosed and treated with radiotherapy 1998-2001. Treatment modalities differed somewhat between the two studies due to additional radiooncology departments being involved in the more recent A study and the different treatment time. The studies were approved by the ethical committee of the University of XX and all patients gave written informed consent. Eligible patients had breast-conserving surgery and did not receive chemotherapy. Standardized toxicity evaluation was conducted by the same study physician for both cohorts. The A study collected detailed information on late toxicity for 414 breast cancer patients who were

recruited in the Rhein-Neckar-Karlsruhe study region of the A study and followed up for a median of six years (13, 14). Genotype data were available for 363 patients who did not receive interstitial or intraoperative boost (9). For the B study, genotype and toxicity data after four years were available for 390 (out of 421) patients without interstitial boost (15, 16). For the replication phase, genotype data, toxicity and treatment data were available for 1,883 patients collated from six study populations of the XXX Consortium (see **Supplementary Table e1** at www.redjournal.org). The six replication cohorts included the C cohort (Japan), the D cohort (Canada), the E study (France), the F study (UK), and the collaborative G study with the H trial and I study (both UK) (**Figure 1**). Selected study characteristics of the eight cohorts are provided in **Table 1** and details on radiotherapy, toxicity information and genotyping in **Supplementary Table e2**. All patients gave written informed consent.

Phenotype(s): The following endpoints were investigated: (1) late skin toxicities (including telangiectasia, pigmentation changes, atrophy); (2) fibrosis of the breast, evaluated at the longest follow-up time points (A, B, D, H, F, I studies) or the maximum grade during follow-up (C,E studies); (3) and additionally overall toxicity measured as a Standardized Total Average Toxicity (STAT) score, combining skin toxicities and fibrosis of the breast, for the replication phase. STAT scores were calculated as described elsewhere (17): First, standardized Z scores were calculated per study for each toxicity endpoint (i) and each individual patient (k) of the respective study:

$$Z_{k,i} = \frac{(s_{k,i} - Mean_i)}{Standard\ Deviation_i}$$

Then, STAT scores were calculated as the average of all non-missing Z scores for an individual patient k :

$$STAT_k = mean\ Z_{k,i}$$

The scoring systems used across the studies included RTOG/EORTC, LENT-SOMA, CTCAE and photographic appearance of the breast.

Genotyping strategy and QC: The discovery SNP panel comprised 305 SNPs in 59 oxidative stress-related genes (mainly tagging SNPs), including base excision repair genes, genotyped using the Illumina GoldenGate™ Assay in the A study (Illumina, San Diego, CA, USA) (18). The list of genes comprised *AKR1C2*, *APEX1*, *APEX2*, *ATM*, *BAX*, *BRCA1*, *CAT*, *CCND1*, *CDKN1A*, *CDKN2A*, *CHEK2*, *CYBA*, *FLT3LG*, *GPX1*, *GSR*, *GSTA1*, *GSTO1*, *H2AFX*, *HMOX1*, *HSPB1*, *LIG3*, *MAPKAPK2*, *MBD4*, *MDC1*, *MDM2*, *MPG*, *MPO*, *MT2A*, *MUTYH*, *NEIL1*, *NEIL2*, *NFE2L2*, *NFKB1*, *NFKBIA*, *NOS3*, *NQO1*, *NTHL1*, *OGG1*, *PARP1*, *PARP2*, *PNKP*, *POLB*, *POLG*, *PRDX1*, *PRDX6*, *SMC1A*, *SMUG1*, *SOD1*, *SOD2*, *TDG*, *TGFB1*, *TNF*, *TP53*, *TXN*, *TXN2*, *TXNRD1*, *TXNRD2*, *UNG*, *XRCC1*. The most promising SNPs (top 15 plus candidate SNPs) were genotyped in the B study using iPLEX application (Sequenom, San Diego, CA, USA).

Ten SNPs in four genes (*NQO1*, *TDG*, *TXN*, *XRCC1*) were selected for replication based on associations with either skin toxicities or fibrosis at p-value <0.05 and/or the same direction in the two discovery cohorts. Genotype data for at least 3 of the 10 replication SNPs were available from each replication study. Genotyping was performed using multiple platforms. All genotype data underwent stringent quality control criteria (call rates, duplicate samples, Hardy-Weinberg equilibrium etc.). The UK G study (H trial, I cohort) provided imputed data using the imputation programme MACH with the HapMap2 reference panel (r^2 between 0.548 and 0.998 for the investigated SNPs). As the analyses were performed separately for each study, there was no formal adjustment for population structure. Details on the genotyping platform and allele frequencies are provided in **Supplementary Tables e2 and e3, respectively** (at www.redjournal.org). There was no deviation from Hardy-Weinberg equilibrium (HWE; all p-values for HWE >0.01).

Data analysis and statistical methods: SAS 9.3 and R 3.1.0 were used for statistical analysis and data visualization. In the discovery phase, we used multivariate binary logistic regression for the dichotomized single endpoints (a) skin toxicities (including telangiectasia) and (b) fibrosis of the breast, and we compared grade ≥ 2 to grade 0 and 1 as reference. Due to the

1 limited sample size, a dominant genetic model (AB/BB vs. AA) and Firth' correction (19) was
2 applied.

3 For the replication, a thorough data harmonization effort was undertaken to ensure
4 comparability of the variables. To achieve more statistical power than by dichotomizing the
5 scores, we used multivariable ordinal logistic regression models (by study) for the single
6 endpoint analyses, comparing higher (more severe) to lower (milder) grades. We additionally
7 investigated overall toxicity in terms of Standardized Total Average Toxicity (STAT) scores
8 using multivariable linear regression per study. In the first regression model, residuals were
9 calculated for each patient using covariates and STAT scores (without the genotype data) to
10 assess toxicity not explained by patient- and treatment-related factors. In a second
11 subsequent regression model, the residuals obtained from model 1 (=difference between
12 observed and estimated toxicity score) were then used to assess associations with genotype.
13 For the G study, which provided imputed data, genotypes were estimated from posterior
14 probabilities. Adjustment for age, radiotherapy dose in terms of Normalized Total Dose (20)
15 and boost yes/no was applied to all studies. Further individual adjustment for variables with
16 $p < 0.1$ in univariate analysis was carried out for single studies, such as follow-up time,
17 chemotherapy, obesity, and smoking status (**Supplementary Table e4** at
18 www.redjournal.org).

19 To summarize and visualize individual study results, random effect models were applied in
20 order to account for different adjustment variables available across studies. The meta-
21 analysis was performed using the R package meta (version 3.6-0).

22 Results

23 Patient characteristics of both the discovery (753 patients) and the replication (1883 patients)
24 phases are provided in **Table 1**. Median follow-up time until late toxicity evaluation ranged
25 from 2 to 6 years, and median age at cancer diagnosis from 53 to 64 years. Two (A,E) of the
26 eight study populations recruited only women above 50 years, while the others also included
27 premenopausal women. In most studies, median whole breast irradiation dose was 50 Gy (in

daily fractions of 1.8 or 2.0 Gy), while two studies had lower median doses of 40 Gy (H trial) and 42 Gy (D study). Between 10% (F cohort) and 90% (B cohort) of the patients received an additional boost. The proportion of grade ≥ 2 skin toxicities ranged between 1% and 30% and for grade ≥ 2 fibrosis between 2% and 25% in the patient cohorts.

Ten SNPs in four genes (*NQO1*_rs10517, *NQO1*_rs2917667, *TDG*_rs10861152, *TDG*_rs2723877, *TXN*_rs2418076, *TXN*_rs2776, *XRCC1*_rs25487, *XRCC1*_rs2682585, *XRCC1*_rs3213356, *XRCC1*_rs3213266) were selected for replication in this analysis. They showed association with either skin toxicity or fibrosis of the breast (p-value < 0.05) in at least one discovery cohort and/or association in the same direction in the two discovery cohorts (**Table 2**). *TNFA*_rs1800629 also fulfilled the selection criteria but these data were published elsewhere (9). In the replication cohorts, some SNPs showed associations in the same direction as in the discovery cohorts, however, none of the SNPs were statistically significantly associated with any of the late toxicity endpoints investigated (**Supplementary Table e5** at www.redjournal.org).

The rs2682585 polymorphism in the base excision repair gene *XRCC1* was consistently associated with a lower risk of normal tissue late toxicity and this inverse association was confirmed in five replication study populations with available data (**Figure 1, Supplementary Table e5**). In the combined data from seven patient cohorts, carriers of the minor allele A of rs2682585 had a significantly lower risk of skin toxicities (multivariate OR 0.77, 95% CI 0.61-0.96, p-value 0.02, **Figure 2A**). Carrying the A allele was also associated with a significantly lower overall risk of late toxicities (**Figure 2C**), corresponding to a 0.08 decrease in STAT score (95% CI -0.15 to -0.02, p-value 0.016), compared to the wild-type genotype GG. No significant association was found for fibrosis of the breast (**Figure 2B**). The association with skin toxicities was in the same order of magnitude in the discovery (OR 0.75, 95% CI 0.55-1.03) and the replication (OR 0.79, 95% CI 0.57-1.08) cohorts and was also very similar for overall late toxicity in the both cohorts. There was no significant study heterogeneity ($p > 0.05$)

for all endpoints examined. The minor allele frequency for rs2682585 was consistent across the Caucasian studies included and ranged between 0.19 and 0.23 (**Supplementary Table e3**). In the Japanese study population (C), the minor allele frequency of 9% was lower; nevertheless the SNP association was consistent with that in the Caucasian study populations. Phenotype by rs2682585 genotype frequencies is provided in **Supplementary Table e6**.

None of the other nine SNPs (*NQO1*_rs10517, *NQO1*_rs2917667, *TDG*_rs10861152, *TDG*_rs2723877, *TXN*_rs2418076, *TXN*_rs2776, *XRCC1*_rs25487, *XRCC1*_rs3213356, *XRCC1*_rs3213266) were replicated and significantly associated with any late toxicity endpoint in the combined analysis.

In a sensitivity analysis, we excluded the D cohort, which had only a few events (12 for skin toxicities, 8 for fibrosis). The results for rs2682585 were robust and remained significant in the combined analysis of the five studies (OR for skin toxicities 0.74, 95% CI 0.59-0.94).

Discussion

In order to overcome the limitations of previous SNP studies due to limited power and lack of replication, we used a two-stage design. We found that a genetic polymorphism of the base excision repair (BER) gene *XRCC1*, rs2682585, was significantly associated with lower risks of skin toxicities and STAT scores in breast cancer patients. *XRCC1* is one of the core BER proteins. It has no known enzymatic function but interacts with other components of the BER pathway such as DNA glycosylases involved in the repair of oxidative damage or DNA polymerase β (POLB) and Ligase III (LIGIII) involved in resealing the repair nick (21). It is thought to be a scaffold protein coordinating the correct sequence of repair steps (22). The rs2682585 SNP is located on chromosome 19 in the *XRCC1* promoter region (1128 base pairs 5' of the transcription start site) where it covers a transcription factor binding site

1 (<http://snpinfo.niehs.nih.gov><http://snpinfo.niehs.nih.gov>).<http://snpinfo.niehs.nih.gov>).

2 Therefore, this SNP may modify *XRCC1* gene expression resulting in altered protein levels.
3 Even minor alterations of a protein component caused by a SNP might affect the
4 stoichiometric association and proper functioning of the complex. Furthermore, rs2682585 is
5 located within the recently identified *PINLYP* (phospholipase A2 inhibitor and LY6/PLAUR
6 domain containing) gene where it causes a His→Arg amino acid change. To our knowledge,
7 there are however currently no experimentally validated functional data available for
8 rs2682585.

9
10 *XRCC1* rs2682585 was not selected for replication with respect to fibrosis and not
11 significantly associated with fibrosis of the breast alone in the combined cohorts. Yet the
12 direction of association was similar to that for skin toxicities, thus the limited number of
13 events for this endpoint might have prevented us from detecting a significant association. It is
14 however also possible that the lack of association reflects different mechanisms underlying
15 the development of skin toxicities and fibrosis.

16
17 As base excision repair is involved in the repair of oxidative damage due to radiotherapy,
18 SNPs in the base excision repair gene *XRCC1* have been investigated previously for
19 associations with radiotoxicity in breast cancer patients, either individually or jointly with
20 SNPs in other repair genes (5, 23-25). However, findings have been inconsistent (26). Some
21 studies suggested a joint effect between the non-synonymous *XRCC1* SNP rs25487
22 (Arg399Gln) and rs1130409 in *APEX1* (Asp148Glu) with regard to the development of
23 radiation-induced acute toxicity and pneumonitis, respectively (27, 28). Neither
24 *XRCC1*_rs25487 alone nor a joint association with *APEX1* was observed for late toxicity in
25 our discovery phase. *XRCC1*_rs25487 was also not significantly associated with late toxicity
26 in the replication studies (**Supplementary Table e5**). Unfortunately, data on the *APEX1* SNP
27 was not available in our replication phase studies.

1 The *XRCC1* SNP rs2682585 is highly coinherited with the non-synonymous *XRCC1* SNPs
2 rs25487 ($r^2=0.17$, $D'=1.0$) and rs25489 (His→Arg; $r^2=0.01$, $D'=1.0$), based on HapMap data,
3 yet it fails to predict the latter two SNPs due to the low r^2 . It is, however, correlated with an
4 intronic *XRCC1* SNP (rs3213334) ($r^2=0.85$). In a recent study (7), this intronic *XRCC1* SNP
5 failed to be validated for radiation-induced late toxicity in 976 breast and 637 prostate cancer
6 patients, with breast cancer patients predominantly from the H trial, which was also included
7 in this investigation. However, sample size was smaller and follow-up time of only 2 years
8 was shorter compared to our data, and photographic assessment tools were used. On the
9 other hand, rs3213334 was recently shown to be associated with basal DNA damage levels
10 in lymphocytes of lung cancer patients (vs. cancer-free controls), assessed by comet assay
11 technology (29). In addition, *XRCC1* polymorphisms were clearly correlated with *in vitro*
12 radiosensitivity in the lung cancer patients. Our findings thus support evidence on *XRCC1*
13 being associated with radiosensitivity in cancer patients, in particular in breast and lung
14 cancer.

15 While the replication effort to demonstrate robustness of the observed association is a clear
16 strength of the investigation, data collection was not standardized across studies. We
17 accounted for this limitation through thorough data harmonization and use of a random-
18 effects meta-analysis approach. In addition, differences in patient characteristics and
19 treatment factors were adjusted for in the statistical analysis to reduce any bias due to study
20 heterogeneity. Despite different treatment modalities and different allele frequencies, our
21 findings for rs2682585 seem to be robust also across different ethnic groups (Caucasians vs.
22 Japanese). However, the heterogeneity in reported toxicities between the studies could have
23 led to reduced statistical power and our ability to validate the association with some of the
24 selected SNPs in the replication studies. To overcome such problems in the future, large
25 prospective cohorts with standardized data and biosample collection are required, one of
26 which is being established by XXX Consortium members (www.xxx.eu) (33).

1 Further analyses need to determine whether *XRCC1*_rs2682585 is a causal variant or is in
2 linkage disequilibrium with the causal SNP. As the biology of radiosensitivity is complex, it is
3 unlikely that only a few genetic variants will be involved in affecting individual radiosensitivity
4 and risk of toxicity. If genetic variation in *XRCC1* together with other biomarkers for
5 radiosensitivity can be validated in further independent cohorts, these genetic biomarkers
6 could be used alongside patient and treatment parameters to individualize patient treatment.
7 Irradiation doses could be increased for non-sensitive patients to improve their prognosis.
8 For patients with an increased risk of toxicity, alternative therapeutic strategies could be
9 offered (e.g. mastectomy, proton therapy), so that the quality of life of cancer patients might
10 be substantially enhanced.

11
12 In conclusion, using a staged design with replication, we identified a genetic variant in
13 *XRCC1* associated with a significantly decreased risk of late toxicities after radiotherapy in
14 breast cancer patients. Additional variants need to be identified to develop a test with a
15 sufficiently large effect size that it can be used clinically for prediction of late radiation toxicity.

References

1. Stone, H.B., Coleman, C.N., Anscher, M.S., et al. Effects of radiation on normal tissue: consequences and mechanisms. *Lancet Oncol.* 2003; 4: 529-536
2. Bentzen, S.M. Preventing or reducing late side effects of radiation therapy: radiobiology meets molecular pathology. *Nat Rev Cancer.* 2006; 6: 702-713
3. Al-Ghazal, S.K., Fallowfield, L., Blamey, R.W. Does cosmetic outcome from treatment of primary breast cancer influence psychosocial morbidity? *Eur J Surg Oncol.* 1999; 25: 571-573
4. Andreassen, C.N. Can risk of radiotherapy-induced normal tissue complications be predicted from genetic profiles? *Acta Oncol.* 2005; 44: 801-815
5. Popanda, O., Marquardt, J.U., Chang-Claude, J., et al. Genetic variation in normal tissue toxicity induced by ionizing radiation. *Mutat Res.* 2009; 667: 58-69
6. AAA., et al. Establishment of a xxx Consortium. *Int J Radiat Oncol Biol Phys.* 2010; 76: 1295-1296
7. AAA, et al. Independent validation of genes and polymorphisms reported to be associated with radiation toxicity: a prospective analysis study. *Lancet Oncol.* 2012; 13: 65-77
8. AAA, et al. Individual patient data meta-analysis shows no association between the SNP rs1800469 in TGFB and late radiotherapy toxicity. *Radiother Oncol.* 2012; 105: 289-295
9. AAA, et al. A replicated association between polymorphisms near TNFalpha and risk for adverse reactions to radiotherapy. *Br J Cancer.* 2012; 107: 748-753
10. AAA, et al. A three-stage genome-wide association study identifies a susceptibility locus for late radiotherapy toxicity at 2q24.1. *Nat Genet.* 2014; 46: 891-894
11. AAA, 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: 178-185

12. Kerns, S.L., de Ruyscher, D., Andreassen, C.N., et al. STROGAR - STrengthening the Reporting Of Genetic Association studies in Radiogenomics. *Radiother Oncol.* 2014; 110: 182-188
13. AAA, et al. Risk of different histological types of postmenopausal breast cancer by type and regimen of menopausal hormone therapy. *Int J Cancer.* 2008; 123: 933-941
14. AAA, et al. Polymorphisms in oxidative stress-related genes and mortality in breast cancer patients--potential differential effects by radiotherapy? *Breast.* 2013; 22: 817-823
15. AAA, et al. Predictive factors for late normal tissue complications following radiotherapy for breast cancer. *Breast Cancer Res Treat.* 2007; 106: 143-150
16. AAA, et al. Genetic polymorphisms in DNA repair and damage response genes and late normal tissue complications of radiotherapy for breast cancer. *BrJCancer.* 2009; 100: 1680-1686
17. AAA, et al. Standardized Total Average Toxicity score: a scale- and grade-independent measure of late radiotherapy toxicity to facilitate pooling of data from different studies. *Int J Radiat Oncol Biol Phys.* 2012; 82: 1065-1074
18. Fan, J.B., Oliphant, A., Shen, R., et al. Highly parallel SNP genotyping. *Cold Spring Harb Symp Quant Biol* 2003; 68: 69-78
19. Firth, D. Bias reduction of maximum likelihood estimates. *Biometrika.* 1993; 80: 27-38
20. Lebesque, J.V., Keus, R.B. The simultaneous boost technique: the concept of relative normalized total dose. *Radiother Oncol.* 1991; 22: 45-55
21. Robertson, A.B., Klungland, A., Rognes, T., et al. DNA repair in mammalian cells: Base excision repair: the long and short of it. *Cell Mol Life Sci.* 2009; 66: 981-993
22. Horton, J.K., Watson, M., Stefanick, D.F., et al. XRCC1 and DNA polymerase beta in cellular protection against cytotoxic DNA single-strand breaks. *Cell Res.* 2008; 18: 48-63
23. Giotopoulos, G., Symonds, R.P., Foweraker, K., et al. The late radiotherapy normal tissue injury phenotypes of telangiectasia, fibrosis and atrophy in breast cancer patients have distinct genotype-dependent causes. *Br J Cancer.* 2007; 96: 1001-1007

24. Mangoni, M., Bisanzi, S., Carozzi, F., et al. Association between genetic polymorphisms in the XRCC1, XRCC3, XPD, GSTM1, GSTT1, MSH2, MLH1, MSH3, and MGMT genes and radiosensitivity in breast cancer patients. *Int J Radiat Oncol Biol Phys.* 2011; 81: 52-58
25. Brem, R., Cox, D.G., Chapot, B., et al. The XRCC1 -77T->C variant: haplotypes, breast cancer risk, response to radiotherapy and the cellular response to DNA damage. *Carcinogenesis.* 2006; 27: 2469-2474
26. Zschenker, O., Raabe, A., Boeckelmann, I.K., et al. Association of single nucleotide polymorphisms in ATM, GSTP1, SOD2, TGFB1, XPD and XRCC1 with clinical and cellular radiosensitivity. *Radiother Oncol.* 2010; 97: 26-32
27. AAA, et al. Association between polymorphisms in the DNA repair genes, XRCC1, APE1, and XPD and acute side effects of radiotherapy in breast cancer patients. *Clin Cancer Res.* 2005; 11: 4802-4809
28. Yin, M., Liao, Z., Liu, Z., et al. Functional polymorphisms of base excision repair genes XRCC1 and APEX1 predict risk of radiation pneumonitis in patients with non-small cell lung cancer treated with definitive radiation therapy. *Int J Radiat Oncol Biol Phys.* 2011; 81: e67-73
29. Hornhardt, S., Rossler, U., Sauter, W., et al. Genetic factors in individual radiation sensitivity. *DNA Repair (Amst).* 2014; 16: 54-65
30. AAA, et al. Haplotype-based analysis of genes associated with risk of adverse skin reactions after radiotherapy in breast cancer patients. *Int J Radiat Oncol Biol Phys.* 2007; 69: 685-693
31. AAA, et al. Concurrent or sequential adjuvant letrozole and radiotherapy after conservative surgery for early-stage breast cancer (CO-HO-RT): a phase 2 randomised trial. *Lancet Oncol.* 2010; 11: 258-265
32. AAA, et al. Test of association between variant tgbeta1 alleles and late adverse effects of breast radiotherapy. *Radiother Oncol.* 2010; 97: 15-18

33. AAA, et al. The xx Project: Validating Predictive Models and Biomarkers of Radiotherapy Toxicity to Reduce Side-effects and Improve Quality of Life in Cancer Survivors. *Clin Oncol (R Coll Radiol)*. 2014 Sep 27. pii: S0936-6555(14)00341-0. doi: 10.1016/j.clon.2014.09.008. [Epub ahead of print]
34. AAA, et al. Genetic polymorphisms and radiation sensitivity in breast cancer patients. *Proceedings of the Canadian Breast Cancer Research Alliance*, p. 263. May, 2006.

Figure legends

Figure 1.

Flow chart of study design showing the number of eligible patients with genotype and toxicity data and follow-up ≥ 1 year in the discovery and replication phase.

Figure 2.

Forest plots for associations of *XRCC1* rs2682585 G>A with radiotherapy-related late toxicity endpoints of skin toxicities (A), fibrosis (B) using ordinal logistic regression and the combined STAT scores (C) using linear regression. “Cases” in (A) and (B) are the number of patients with toxicity scores ≥ 1 . A dominant genetic model was used (AG/AA vs. GG)

List of Table(s) in the main text:

Table 1. Summary of the breast cancer study populations.

Table 2. Top hits selected for replication and their associations with late toxicity in breast cancer patients in the discovery A and B cohorts using multivariate binary logistic regression models.

List of Figure(s) in the main text:

Figure 1. Study design.

Figure 2. Forest plots for associations of *XRCC1* rs2682585 and radiotherapy-related late toxicity.

Table 1. Summary of the breast cancer study populations.

Characteristics	Discovery studies		Replication studies					
	MARIE ^{RAD}	ISE-2	BRRG	CIHR	COHORT	IMRT	LeND	RACE
Reference	(9)	(15, 16)	(30)	(34)	(31)	(7)	(9)	(32)
Patients*	363	390	301	106	105	839	361	171
Country	Germany	Germany	Japan	Canada	France	UK	UK	UK
Follow-up [†] (years)	5.7 (3.6-7.5)	4.3 (3.0-6.4)	2.0 (1.0-3.0)	3.1 (1.7-4.7)	6.0 (1.2-8.1)	2.0 (all)	5.2 (2.4-21)	5.0 (2.0,5.0)
Age at cancer diagnosis [†] (years)	64 (50-75)	60 (27-88)	53 (31-83)	57 (33-81)	63 (51-82)	59 (26-83)	58 (32-87)	56 (34-76)
Whole breast dose (Gy) [†]	50 (38-56)	50 (44-56)	50 (45-60)	42 (39-50)	50 (45-60)	40 (all)	50 (39-50)	50 (39-50)
Boost yes	72%	90%	21%	36%	30%	62%	10%	82%
Skin toxicities grade 1	75	74	104	7	21	66	28	39
grade 2	37	117	3	4	6	53	50	5
grade 3	0	1	0	1	0	8	0	4+1
≥Grade 2 skin toxicity	10%	30%	1%	5%	6%	7%	14%	6%
Fibrosis grade 1	240	214	NA	6	47	316	47	2
grade 2	23	22	NA	2	16	164	29	42
grade 3	5	4	NA	0	4	3	0	0
≥Grade 2 fibrosis	8%	7%	NA	2%	19%	20%	8%	25%
Overweight/Cup size [‡]	58%	49%	NA	48%	50%	66%	27%	NA
Ever smoking	40%	30%	17%	8%	31%	NA	24%	NA
Chemotherapy	0%	0%	26%	40%	18%	16%	33%	23%

* With genotype and late toxicity data available and follow-up time of ≥ 1 year.

† Median (min-max);

‡ Overweight = >25 kg/m² and cup size $\geq$ D.

|| Maximum toxicity grade after a minimum of 1 year of follow-up.

NA=not available

Table 2. Top SNPs selected for replication and their associations with late toxicity (skin toxicities, fibrosis) in breast cancer patients in the discovery cohorts MARIE_{RAD} and ISE-2 using multivariate logistic regression models.

Outcome	Gene	SNP	MARIE _{RAD} study		ISE-2 study	
			OR (95% CI)*	n _{total} /n _{Events}	OR (95% CI)*	n _{total} /n _{Events}
Skin toxicities	<i>NQO1</i>	rs10517	0.27 (0.08-0.79)	359 / 37	0.44 (0.24-0.77)	385 / 116
	<i>NQO1</i>	rs2917667	0.30 (0.07-0.93)	345 / 34	0.45 (0.24-0.81)	382 / 116
	<i>TXN</i>	rs2776	0.27 (0.10-0.69)	341 / 34	0.93 (0.56-1.55)	381 / 117
	<i>XRCC1</i>	rs2682585	0.45 (0.19-0.99)	360 / 37	0.64 (0.39-1.04)	367 / 111
	<i>XRCC1</i>	rs3213356	0.29 (0.12-0.66)	344 / 34	0.92 (0.57-1.49)	384 / 117
Fibrosis of the breast	<i>TDG</i>	rs10861152	0.48 (0.21-1.07)	355 / 27	0.72 (0.32-1.61)	387 / 26
	<i>TDG</i>	rs2723877	0.16 (0.02-0.66)	358 / 28	0.51 (0.13-1.48)	385 / 26
	<i>TNF</i>	rs1800629 [†]	1.23 (0.49-2.84)	359 / 28	3.30 (1.48-7.68)	385 / 26
	<i>TXN</i>	rs2418076	1.39 (0.60-3.27)	342 / 24	2.75 (1.21-6.74)	380 / 26
	<i>XRCC1</i>	rs25487	2.03 (0.90-4.88)	358 / 28	1.04 (0.47-2.43)	379 / 26
	<i>XRCC1</i>	rs3213266	2.62 (1.06-6.17)	358 / 28	2.71 (1.12-6.22)	385 / 26

* Multivariate (binary) logistic regression (toxicity grade ≥ 2 vs. grade 0+1), adjusted for age, boost energy type, follow-up time, irradiation dose, BMI.

Please note: Due to slightly different models used in the discovery and replication phases, respectively, the estimates somewhat differ to the ones presented in forest plots in Figure 2.

† Results on this SNP are already published in (9)(9).

Figure 1.

1. Discovery Phase: 305 SNPs in 59 oxidative stress-related genes

2 breast cancer cohorts from Germany:

- MARIE_{RAD} (N=363)
- ISE-2 (N=390)

2. Replication Phase: 10 SNPs in 4 genes

6 breast cancer cohorts (≥100 patients and at least one year of follow-up) from the Radiogenomics Consortium

- BRRG (Japan; N=301)
- CIHR (CAN; N=106)
- COHORT (France; N=105)
- IMRT trial (UK; N=839)
- LeND (UK; N=361)
- RACE (UK; N=171)

Figure 2.

A. Skin toxicities

B. Fibrosis of the breast

C. STAT scores