Tumor control and normal tissue toxicity: The two faces of radiotherapy
van Oorschot, B.

Citation for published version (APA):
van Oorschot, B. (2016). Tumor control and normal tissue toxicity: The two faces of radiotherapy
CHAPTER 6

Inefficient DNA DSB Repair via the Homologous Recombination Pathway in Prostate Cancer Patients with Late Radiation-Toxicity

Bregje van Oorschot, Lon Uitterhoeve, Rosemarie ten Cate, Ilja Oomen, Jan Paul Medema, Harry Vrieling, Lukas J.A. Stalpers, Perry D. Moerland and Nicolaas A.P. Franken

Under Review
ABSTRACT

Purpose: Severe late normal tissue damage limits radiotherapy treatment regimens. This study aims to validate γ-H2AX foci decay ratios and induced expression levels of DNA double strand break (DSB) repair genes, found in a retrospective study, as possible genetic predictors for late radiation toxicity.

Methods and Materials: Prospectively, decay ratios (initial/residual γ-H2AX foci numbers) and genome-wide expression profiles were examined in ex vivo irradiated lymphocytes of 198 prostate cancer patients. All patients were followed ≥2 years after radiotherapy, clinical characteristics were assembled and toxicity was recorded using the Common Terminology Criteria (CTCAE) v4.0.

Results: No clinical factors were correlated with late radiation toxicity. Analysis of γ-H2AX foci uncovered a negative correlation between the foci decay ratio and toxicity grade. Significantly smaller decay ratios were found in grade≥3 compared to grade 0 patients (p=0.02), indicating less efficient DNA-DSB repair in radio-sensitive patients. Moreover, utilizing a foci decay ratio threshold determined in our previous retrospective study correctly classified 23 of the 28 grade≥3 patients (sensitivity, 82%) and 9 of the 14 grade 0 patients (specificity, 64%). Grade of toxicity also correlated with a reduced induction of the homologous recombination (HR) repair gene-set. The difference in average fold induction of the HR gene-set was most pronounced between grade 0 and grade≥3 patients (p=0.008).

Conclusions: Reduced responsiveness of HR-repair genes to irradiation and inefficient DSB repair correlate with an increased risk of late radiation toxicity. Using a decay ratio classifier, we could correctly classify 82% of the patients with grade≥3 toxicity. This study is of utmost importance for the identification of patients who are sensitive to radiation and likely to have a genetic predisposition to develop late radiation toxicity.
INTRODUCTION

Radiotherapy is a widely used anti-cancer treatment, applied in approximately 50% of all cancer patients [1]. Prostate cancer is well controlled by external beam radiotherapy, leading to high survival rates [2]. However, the development of severe late side effects induced by radiotherapy remains a heavy burden. Severe late radiation toxicity occurs in approximately 10% of the treated patients, and only a small group of patients undergo treatment with no or very few complications. Early identification of patients at high risk for late radiation toxicity may help selection for alternatives, such as surgery or brachytherapy.

Clinical factors like age, radiation-dose or volume can only partly explain the risk of late radiation toxicity [3, 4]. In addition, previous abdominal surgery [5-7], diabetes mellitus [8] and cardiovascular disease [6, 9] have been incriminated as possible predictors but there is a lack of confirmation in independent studies. Patients with radio-resistant normal tissue (toxicity grade 0) are thought to have an altered radiation response compared to radio-sensitive normal tissue patients (toxicity grade≥3). Therefore, it has been suggested that there is a genetic predisposition for the risk of severe late side effects [10]. Several studies investigated genetic variations of genes involved in apoptosis [11, 12], DNA repair [13, 14] or fibrosis [15] and their role in the development of late radiation toxicity. Genome-wide association studies (GWAS) in large patient cohorts uncovered other possible genetic risk factors [16, 17]. Apart from mutations or polymorphisms, also gene expression profiles have been explored as prognostic markers for late radiation toxicity [18, 19]. However, overall results are conflicting[20] and no reproducible or reliable prognostic markers [21-23] associated with late radiation toxicity have been identified yet. In 2013, we retrospectively showed reduced induction of DNA repair gene expression levels after radiation in lymphocytes of patients with late complications compared to those without [24].

Furthermore, monitoring the induction and repair of individual DNA double strand breaks (DSBs) in lymphocytes has been proposed as a method to predict normal tissue radiation sensitivity [25-27]. The assay uses the immuno-fluorescent detection of γ-H2AX, a well-established biomarker for radiation-induced DNA DSBs [28, 29]. DSBs represent the most lethal DNA damage induced by radiation treatment and may lead, if unrepaired, to cell death. The two major pathways involved in the repair of DNA-DSBs are Homologous Recombination (HR) and Non-Homologous End Joining (NHEJ). In our retrospective study more residual γ-H2AX foci were found in ex vivo irradiated lymphocytes of radio-sensitive patients compared to radio-resistant patients, indicating less efficient repair of DNA-DSBs in patients with radiation toxicity.

This study aims to validate the findings of our retrospective study [24], where DNA-DSB foci decay ratios and induction levels of HR DSB repair genes were significantly different between lymphocytes of radio-sensitive and radio-resistant prostate cancer patients. Prospectively, 200 prostate cancer patients were included and the application of foci decay ratios and induction levels of DSB repair genes to predict late toxicity was investigated.
MATERIALS & METHODS

Patient inclusion and sample collection
Between 2009 and 2013, we accrued 200 patients diagnosed with prostate cancer receiving curative external beam radiotherapy in combination with hormonal therapy at the Academic Medical Center (AMC) of the University of Amsterdam, with a follow-up of ≥2 years. We analyzed data from 198 patients with full information about PSA, age, T-classification, and comorbidities. After written informed consent, 40 ml whole blood was collected of all patients before start of treatment. Peripheral lymphocytes were isolated using Ficoll (Ficoll-Paque PLUS, GE Healthcare) gradient separation, and stored in liquid nitrogen. Development of late toxicities was monitored over more than 2 years after treatment using the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) version 4.0. Toxicity grade was determined mainly focusing on late gastrointestinal (GI) and genitourinary (GU) toxicities.

Immunohistochemistry for γ-H2AX foci
Peripheral lymphocytes were thawed and ex vivo irradiated with 1Gy γ-rays using a 137Cs source with a dose rate of approximately 0.5Gy/min. Induction and decay of radiation induced γ-H2AX foci was measured in unstimulated G(0) cells. At 30 min and 24 h post irradiation, lymphocytes were dropped on poly-D-lysine coated slides and fixed in 4% paraformaldehyde. After 25 min, slides were washed with PBS and ready for immunostaining. The γ-H2AX foci immunostaining was performed as previously described[30].

γ-H2AX foci scoring
The number of radiation-induced γ-H2AX foci was determined in fluorescent stack images using Image-Pro Plus software. Slices of 20 stacks with a 200-nm interval were obtained using a Leica-DM-RA-HC Upright Microscope. Stacks were deconvolved as 1 photomicrograph and the number of foci per nucleus was scored[31]. Decay ratios were calculated by dividing the number of γ-H2AX foci found at 30 min by the number of γ-H2AX foci found at 24 h post irradiation. For every patient the kinetics of γ-H2AX foci decay was determined in at least 100 cells.

Foci decay threshold determination
In our previous retrospective study [24], foci decay ratios were determined in lymphocytes of 24 patients, 11 patients with Grade 0 and 13 patients with Grade 3 late radiation toxicity (See Supplementary Figure 1). Receiver-operator characteristics (ROC) and area under the curve (AUC) were computed in GraphPad-Prism 6 (California, USA) using the method of Hanley & McNeil and resulted in an AUC=0.970±0.03, 95% CI 0.91–1.00 and p<0.0001. A threshold ratio to separate both groups was determined at the highest sensitivity value with a 100% specificity. In this way, all grade 0 patients are correctly identified as negative for radiation toxicity as
Late Radiation Toxicity due to Reduced DSB Repair

treatment regimens should not be adjusted in this group of patients. The threshold ratio was calculated as the mean of the lowest decay ratio in the grade 0 and the highest in the grade 3 group (excluding the decay ratio that was higher than the lowest grade 0 decay ratio). This resulted in a threshold decay ratio of 3.41, which separated both groups with 92.1% sensitivity and 100% specificity.

Microarray analysis
Similarly to the retrospective study, lymphocytes were cultured stimulated by phytohemagglutinin (concentration of 1μg/ml). After two weeks, half of the cells were irradiated at room temperature with 2Gy gamma rays from a $^{137}$Cs source, dose rate of approximately 0.5Gy/min and the other half of the cells was left untreated. Isolation of RNA was executed 24h post radiation using RNeasy Mini Kit (Qiagen) according to manufacturer’s protocol. RIN value and RNA quality of both treated as untreated RNA was assessed with a BD Bioanalyzer. Biotin-labeled cRNA probes were generated and RNA was hybridized to HT HG-U133+ PM GeneChip® array (Affymetrix, Santa Clara CA, USA) according to the manufacturer’s protocol. Scanning of the array was conducted by the MicroArray Department of the University of Amsterdam, and images were processed to obtain an intensity value for each oligonucleotide probe. Standardized microarray data quality control was performed using the R/Bioconductor package arrayQualityMetrics. Data were normalized and summarized to the probe set level using the robust multi-array average (RMA) [32].

Genes that respond differently to irradiation between patient groups were determined using a linear model with a fixed main effect for grade, and nested interactions of grade with patient and radiation (2Gy or 0Gy). Significant differences in response between grade 0 and the other grades were determined using the appropriate contrasts and empirical Bayes moderated F- and t-statistics (R/Bioconductor package limma). Resulting p-values were corrected for multiple testing using the Benjamini-Hochberg False Discovery Rate adjustment. See supplementary Information B for a detailed description of the microarray analysis and additional results.

Gene-set enrichment analysis
To validate the retrospective study, the HR gene-set was a used to determine differences in radiation response between patients of the four toxicity groups. Probesets for the HR gene-set were chosen based on our retrospective study [24]. Induction levels were assessed with a ROAST gene set test [33] (limma package). ROAST P-values were calculated using the nested interaction design described above for three possible alternative hypotheses using 50,000 rotations and default parameters. The alternative ‘up’ tests whether the genes in a gene-set tend to be up-regulated, the alternative ‘down’ tests whether the genes in a gene-set tend to be down-regulated, the alternative ‘mixed’ tests whether the genes in the set tend to be differentially expressed, without regard for direction. The two-sided directional p-value is reported.
Figure 1. Calculation of γ-H2AX foci decay ratios after 1Gy radiation and classification model. A: Visualization of γ-H2AX foci at 30min and 24h post irradiation in a patient with grade 0 and a patient with grade 3 toxicities. Bar is 5μm. B: Foci decay ratios of all assessed patient per toxicity group as scored by CTCAE 4.0 grading system. A significant negative correlation is detected between toxicity grade and foci decay ratios (dashed line). Foci decay ratio: number of foci at 30 min/number of foci at 24h. At least 100 cells per patient per condition are counted. C: Receiver operating characteristic curve for foci decay ratios of patients with toxicity grade 3 versus toxicity grade 0. Based on decay ratios, the diagnostic accuracy as quantified by the area under the curve is 73%. D: Foci decay ratios of grade 0 and grade 3 patients; every point represents an individual patient. Geometric mean±95% confidence interval within each group is shown. A threshold of 3.41 determined in our retrospective study (pink dashed line) correctly classified 82.1% of grade 3 patients with 64.3% specificity.

Statistical Analysis
Clinical data was analyzed in IBM SPSS statistics 22. Numbers of induced γ-H2AX foci and decay ratios were analyzed in GraphPad Prism version 5.0. Ordinal and categorical data were assessed with a cross tab chi-square analysis. Continuous variables were first tested for normality using the Shapiro-Wilk test (p>0.05), followed by either one-way ANOVA (normally distributed; post hoc analysis using Holm-Sidak test) or non-parametric Kruskal-Wallis (post hoc analysis using
Late Radiation Toxicity due to Reduced DSB Repair

Dunn’s multiple comparisons test). ROC, AUC and confidence interval were computed using the method of Hanley and McNeil. Significant P values are given: * indicates P<0.05, ** indicate P<0.01 and *** indicate P<0.001, ns indicates not statistically significant.

RESULTS

Toxicity score determination
Out of 198 patients, 28 patients (14%) displayed grade≥3 toxicity to the bladder and/or rectum at more than one time point within two years post treatment. Grade 2 was recorded in the majority of the patients (109 patients, 55%), 47 (24%) patients developed grade 1 and 14 (7%) patients were recorded as grade 0. As expected, the incidence of both extreme patient groups, either no side effects at all (grade 0) or severe late side effects (grade≥3) are approximately 10%.

Clinical characteristics
General information, baseline characteristics and medical history were recorded of all patients and correlated to the incidence and severity of normal tissue damage (Table 1). Interestingly, a higher dose of radiation did not influence the development of severe late side effects. Only 12 patients received 70Gy instead of 77Gy, and they were evenly distributed among patient groups. Furthermore, no significant differences in baseline characteristics or medical history were detected between grade 0, 1, 2 and 3 respectively. In conclusion, the four patients groups were well-matched and no clinical prognostic events for the development of late radiation toxicity were observed.

γ-H2AX foci decay ratios
The induction and decay of radiation induced DNA–DSB breaks was monitored in ex vivo irradiated lymphocytes by detecting γ-H2AX foci 30min and 24h post radiation as depicted in Fig. 1A. Similarly to the retrospective study, radio-sensitive patients appear to have a less efficient DNA-DSB damage response. Calculation of foci decay ratios revealed significant differences between grade≥3 and grade 0 patients (P=0.02). Subsequent linear regression analysis confirmed a significant negative correlation between foci decay ratios and the severity of normal tissue damage (R²=0.03, P=0.01; Fig. 1B). Furthermore, based on their foci decay ratios, patients with severe radiation toxicity (grade≥3) can be discriminated from those without (grade=0) fairly accurately (AUC= 0.73, 95% CI 0.55-0.92; Fig. 1C). A threshold foci decay ratio of 3.41 was determined in the retrospective patient cohort, where only grade 0 and grade 3 patients were included [24] (Materials & Methods, Supplementary Figure 1). Based on this threshold 82% (23 out of 28) of grade≥3 patients could be correctly classified and 64% (9 out of 14) of grade 0 patients (Fig. 1D).
However, an overlap between the foci decay ratios of all 4 patients groups can be observed. Using the threshold of 3.41 incorrectly classified 36.2% of the grade 1 and 29.4% of the grade 2 patients as grade 0.

Figure 2. Comparison of the radiation response of HR gene-set between grade 3 and grade 0 patients after 2Gy irradiation. A: Volcano plot shows that all HR genes have a negative fold change, meaning that the gene was induced to a lesser degree in grade≥3 patients compared to grade 0 patients. B: Receiver operating characteristic curve for the mean $\log_2$ fold induction of HR gene-set. Diagnostic accuracy to discriminate grade 3 from grade 0 patients based on HR expression levels is 74%, as quantified by the area under the curve. C: Average induction levels of HR-set as single value show a significant difference between grade 0 and grade 3 patients, indicating a less active DNA-DSB HR repair in radiosensitive patients. Geometric mean±95% confidence interval within each group is shown; every point represents an individual patient.
Differences in gene expression of HR DNA repair genes among patient groups

Gene expression levels of our previous established gene-set for the DNA-DSB repair pathway HR were examined for all patients. Fig. 2A shows the differences in radiation response for the HR genes between the two groups. Concordant with the retrospective study, the HR repair gene-set was significantly down-regulated after radiation (2Gy versus 0Gy) in grade≥3 patients compared to grade 0 patients (P=0.008; Fig. 2C). Although HR induction levels distinguished radio-sensitive patients from radio-resistant patients (AUC=0.75, 95% CI 0.60-0.89; Fig. 2B), reliable classification based on gene expression levels was not possible. A linear support vector machine classifier fit on the HR induction levels of the retrospective cohort correctly classified only 13 of the 28 patients developing grade 3 toxicity in the prospective cohort (sensitivity 46%; Supplementary Information B, section 5).

DISCUSSION

Radiotherapy aims to precisely target the tumor, while sparing the normal healthy tissue[34]. However, using external beam therapy the surrounding normal tissue is unavoidably exposed to radiation. The possible appearance of normal tissue complications is therefore a limiting factor for radiation treatment. In concordance with our retrospective results [24], our present study prospectively confirms a significant correlation between the severity of late radiation toxicity in prostate cancer patients and the ability of their normal cells to repair DNA-DSBs. Both gene expression profiling and γ-H2AX foci decay data after in vitro irradiation of patient lymphocytes show that inefficient repair of radiation-induced DSB is highly associated with the development of severe late normal tissue damage.

This study uses the ratio of γ-H2AX foci numbers 24h after radiation treatment to γ-H2AX foci numbers 30 min after radiation treatment as a measure for the efficiency of DNA-DSB repair. Whether disappearance of foci corresponds to the actual repair of a DSB is still subject of discussion[35]. Several studies have shown that residual, persisting γ-H2AX foci are associated with DNA-repair deficiencies or radio-sensitivity [25, 28], indicating that the γ-H2AX foci assay can be used to study DNA-DSB repair. Here we show that foci decay ratios correlate well with late radiation toxicity. Significantly lower foci decay ratios were found in patients with severe late radiation toxicity (grade≥3) compared to patients developing no or mild complications (grade 0; P=0.02). In addition, less efficient DNA damage response is reflected by a reduced transcriptional responsiveness of HR genes to irradiation. Gene expression levels of HR-set are overall down-regulated in radio-resistant grade 0 patients compared to radio-sensitive patients (grade≥3; P=0.008). An increased complication rate seems to correlate with lower activity of HR genes.
Karnofsky Performance Status (KPS), an index used to classify patients to their functional impairment.

Gleason grading system based on microscopic analysis of tumor samples evaluates the prognosis of patients (the higher the Gleason score, the more aggressive the tumor is and the worse prognosis).

Prostate-specific antigen, an important biochemical marker for prostate cancer; normal value <4ng/ml.

T-classification: clinical classification for primary tumor extension; T1, non-palpable; T2, palpable within prostate; T3, extension beyond prostate; T4, invasion of adjacent organs.

TURP: transurethral resection of the prostate, urological operation to reduce urinary symptoms if medical treatment fails.

---

**Table 1: Patient Characteristics of 198 prostate cancer patients.** Toxicity score is determined according to CTCAE 4.0 system.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sub-category</th>
<th>Grade 0 (n=14)</th>
<th>Grade 1 (n=47)</th>
<th>Grade 2 (n=109)</th>
<th>Grade 3 (n=28)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>mean</td>
<td>range</td>
<td>n</td>
<td>mean</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>69</td>
<td>59 - 76</td>
<td>70</td>
<td>54 - 86</td>
<td>69</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td>26.1</td>
<td>21.3 - 30.4</td>
<td>27.2</td>
<td>19.6 - 38.3</td>
<td>27.1</td>
</tr>
<tr>
<td>KPS*</td>
<td></td>
<td>97.5</td>
<td>80 - 100</td>
<td>97.1</td>
<td>80 - 100</td>
<td>97.5</td>
</tr>
<tr>
<td>Gleason score†</td>
<td></td>
<td>7.3</td>
<td>6 - 9</td>
<td>7.4</td>
<td>6 - 9</td>
<td>7.3</td>
</tr>
<tr>
<td>PSA†† before radiotherapy</td>
<td></td>
<td>38.8</td>
<td>6.2 - 110</td>
<td>35.5</td>
<td>2.9 - 457.3</td>
<td>33.5</td>
</tr>
<tr>
<td>Radiotherapy</td>
<td>70 Gy</td>
<td>13</td>
<td>1</td>
<td>42</td>
<td>4</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>77 Gy</td>
<td>5</td>
<td>7</td>
<td>13</td>
<td>19</td>
<td>49</td>
</tr>
<tr>
<td>T-classification¶</td>
<td>T1</td>
<td>4</td>
<td>7</td>
<td>13</td>
<td>5</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>T2</td>
<td>5</td>
<td>15</td>
<td>49</td>
<td>5</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>T3</td>
<td>19</td>
<td>44</td>
<td>12</td>
<td>1</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>T4</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>Abdominal surgery</td>
<td>Yes</td>
<td>3</td>
<td>15</td>
<td>23</td>
<td>5</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>11</td>
<td>32</td>
<td>14</td>
<td>7</td>
<td>95</td>
</tr>
<tr>
<td>TURP†</td>
<td>Yes</td>
<td>2</td>
<td>5</td>
<td>14</td>
<td>7</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>12</td>
<td>42</td>
<td>95</td>
<td>21</td>
<td>0.34</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Yes</td>
<td>1</td>
<td>7</td>
<td>17</td>
<td>4</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>13</td>
<td>40</td>
<td>92</td>
<td>24</td>
<td>0.97</td>
</tr>
<tr>
<td>Intestinal disease</td>
<td>Yes</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>14</td>
<td>46</td>
<td>106</td>
<td>27</td>
<td>0.22</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>Yes</td>
<td>4</td>
<td>13</td>
<td>34</td>
<td>14</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>10</td>
<td>34</td>
<td>75</td>
<td>14</td>
<td>0.22</td>
</tr>
</tbody>
</table>

*Karnofsky Performance Status (KPS), an index used to classify patients to their functional impairment. †Gleason grading system based on microscopic analysis of tumor samples evaluates the prognosis of patients (the higher the Gleason score, the more aggressive the tumor is and the worse prognosis).

††Prostate-specific antigen, an important biochemical marker for prostate cancer; normal value <4ng/ml. †T-classification: clinical classification for primary tumor extension; T1, non-palpable; T2, palpable within prostate; T3, extension beyond prostate; T4, invasion of adjacent organs. †TURP: transurethral resection of the prostate, urological operation to reduce urinary symptoms if medical treatment fails.
None of the patients’ characteristics that have been described as prognostic marker for late radiation toxicity were predictive in the present prospective study. We could not confirm the observations by Peeters et al[36] that the risk of radiation-related late rectal bleeding increased with a higher radiation dose. In addition, no association was found between previous abdominal surgery, diabetes mellitus or cardiovascular disease and a higher risk for radiation toxicity. This might be due to the fact that procedures to determine the toxicity score of patients differ considerably. Valdagni et al. [5, 6, 37] used a patient-administered RTOG/EORTC questionnaire to score rectal and intestinal toxicities whereas we used the clinician-based CTCAE system. Similar classification of patient status is very important for meaningful comparison of different studies and results. Most clinical trials track patient toxicity symptoms by the CTCAE system, however no standardized implementation of the system exists and interpretation of symptoms varies among clinicians. It has been reported that patients can effectively report toxicity using patient-reported outcomes (PRO) [38, 39], but the results are more subjective compared to clinician-based scoring and reflect daily health status [40].

Over the years, significant progress has been made toward the effectiveness of radiation treatment and tumor control. The resulting increase of cure rates in cancer patients also increases the importance of minimizing the radiation impact to normal tissues. Prevention or reduction of late normal tissue toxicity can lead to an improved quality of life after treatment. Patients at high risk to develop late radiation toxicity might alternatively be treated by surgery or brachytherapy. In combination with hyperthermia, patients can be treated with lower radiation dose while maintaining clinical effectiveness [41]. On the other hand, patients who are less sensitive to radiation toxicity might be treated with higher doses. Especially for other types of cancer, with worse survival rates than prostate cancer, grade 0 patients might benefit from a higher radiation dose. However, care must be taken with right classification of the patients and possible change of treatment plan. Patients who develop late toxicities should not be irradiated with higher dose, while on the contrary grade 0 patients will not be disadvantaged from a different treatment regime if it results in similar outcomes.

In conclusion, this study confirmed that there are marked differences between prostate cancer patients in their DNA damage response in normal tissues. An altered or less efficient DNA-DSB response correlates with severe late radiation toxicity. Therefore, it is worthwhile to further investigate the DNA damage response or other DSB markers to elucidate the mechanisms which underlie the differences in radiation response. The ability to separate patients suffering from severe late radiation toxicities (grade≥3) from those without toxicities (grade 0) by γ-H2AX foci decay ratio measurement is already promising. Since the γ-H2AX assay is a robust and highly reproducible technique, it should be possible to develop a standardized method for clinical use. However, due to the large overlap with patients experiencing milder toxicities, results need to be further validated before implementation in the clinic is warranted.
SUPPLEMENTARY INFORMATION

Figure S1. Information file B

ACKNOWLEDGEMENTS

The authors thank the participating patients and nurses of the AMC Radiotherapy department. The Dutch Cancer Foundation (nos. UVA 2008-4019) and the Stichting Vanderes are acknowledged for financing personnel support. We are grateful to the Maurits and Anna Kock and the Nijbakker Morra foundations for sponsoring laboratory equipment.

REFERENCES

Late Radiation Toxicity due to Reduced DSB Repair


Supplementary Figure S1. Receiver operating characteristic (ROC) curve (A) and foci decay ratios (B) from our retrospective study. This study included a total of 24 patients, 11 grade 0 and 13 grade 3 patients. A: ROC curve and a diagnostic accuracy of 97% as quantified by the area under the curve. Threshold was determined at the highest sensitivity value with a specificity of 100%. This resulted in a threshold of 3.41. B: Foci decay ratios are significantly different between patients with (grade 3) and without (grade 0) severe late radiation toxicities (p=0.0001). The threshold of 3.41 correctly identified all grade 0 patients (pink dashed line). Every point represents an individual patient. Geometric mean±95% confidence interval within each group is show.
SUPPLEMENTARY INFORMATION B

CONTENTS
1. Reproducibility 126
2. Microarray pre-processing 126
3. Radiation Sensitivity 127
4. Differences in response to irradiation between patient groups 128
5. Classification 130

1. REPRODUCIBILITY
All analyses were performed using R/Bioconductor packages in the statistical software environment R(v3.1.2). Versions of the main packages used were: affy (1.44.0), arrayQualityMetrics (3.22.0), e0171(1.6-7), hgu133plus2.db (3.0.0), limma (3.22.7), MCRestimate (2.22.0), org.Hs.eg.db (3.0.0), and pROC (1.8).

2. MICROARRAY PRE-PROCESSING
Raw data were extracted from the CEL files using the affy package. Data was normalized and summarized at the probeset level using robust multiarray averaging (RMA) with default settings (function rma, package affy) [1]. The package arrayQualityMetrics was used to assess the quality of the microarray data both before and after normalization. In the raw data both the Relative Log Expression (RLE) and Normalized Unscaled Standard Error (NUSE) showed slightly deviating distributions for some of the arrays. After normalization, there were no clear indications for arrays of lower quality. We therefore decided to not remove any arrays based on quality control.

Samples were hybridized in two batches, the first batch consisting of 104 arrays on five Affymetrix HT HG-U133+ PM array plates (2x16, 3x24) and the second batch of 336 arrays on five plates (2x24, 3x96). Clear differences between the two batches and, to a lesser extent, between the plates within batches were still present after normalization. For all patients the irradiated and the control sample had by design been hybridized in the same batch and, for the second and largest batch, on the same plate. Since in all analyses the pairing of the patient samples was explicitly taken into account, differences between batches and plates were automatically corrected for. The most dominant signal in the data was patient-specific and in general the irradiated and the control sample of the same patient clustered closely together (data not shown).

Twelve technical replicates arrays had been included in the first batch. Replicated samples were in general very similar and were therefore replaced with their averages. Twenty eight arrays were used for samples from our retrospective study [2] and were left out. Furthermore, four arrays were left out for two patients for whom the toxicity score was missing. In total, the prospective cohort therefore consisted of 396 arrays for 198 prostate cancer patients. Probesets were annotated using the package hgu133plus2.db.
3. RADIATION SENSITIVITY

We first determined the genes that most changed their expression levels upon irradiation in any of the four patient groups, corresponding to the four toxicity score levels. Overall a large radiation response was observed with 18,480 genes differentially expressed (adjusted $P<0.05$). The most differentially expressed genes consisted of well-known radiation-sensitive genes (for example, FDXR, MDM2, XPC, and DDB2; Table S1).

Table S1: Top-25 probesets differentially expressed between irradiated (2Gy) and control (0Gy) lymphocytes.

For each grade (0, 1, 2, ≥3) the differentially expressed probesets were determined by extracting the appropriate coefficients from the linear model fit with the nested interaction design described in the main text. Fold changes are given on a log$_2$ scale. $P$ values were calculated on the basis of the moderated F-statistics for the four comparisons. Correction for multiple testing was performed using the Benjamini-Hochberg false discovery rate (adj. $P$ value).

We used the systematic upregulation of three genes (FDXR, MDM2, and TRIM22) to check for possible sample swaps. For this purpose, we calculated the log$_2$ fold change between the expression levels at 2Gy and 0Gy within each patient. For all patients except one, expression of the three selected genes was indeed upregulated. Induction levels after correcting for this sample swap are depicted in Figure S1.

A geneset analysis of the response to radiation confirmed that the signature identified in the current experiment is largely concordant with previously identified radiation signatures and related pathways, such as the P53 signaling pathway (Table S2, Figure S2).
4. DIFFERENCES IN RESPONSE TO IRRADIATION BETWEEN PATIENT GROUPS

In our retrospective study, we found that expression of a set of genes involved in homologous recombination (HR) was less induced upon irradiation in 9 patients with Grade 3 than in 7 patients with Grade 0 late radiation toxicity [2]. In the main text the prospective validation of this observation is described. Here, these and related results are presented in more detail.

We first determined genes that respond differently to irradiation between the four toxicity score-based patient groups. After correction for multiple testing, no genes were differentially expressed (adj. \( P \geq 0.7 \), Table S3). Differences in fold induction were in general small and the fold induction was highly variable within patient groups. Limiting the analysis to the 14 patients with Grade 0 and the 28 patients with Grade 3 late radiation toxicity, the most differentially expressed probesets were 230409_PM_at (MAGI3, adj. \( P = 0.16 \)), 1556817_PM_a_at (adj. \( P = 0.16 \)) and 235059_PM_at (RAB12, adj. \( P = 0.23 \)). We next investigated the response to irradiation of the HR genes. The large majority of HR genes was less induced upon irradiation in Grade 3 patients compared to Grade 0 patients (Table S4) in line with the findings in our retrospective study. A ROAST geneset test [4] (see main text for details) confirmed that although the effects at the individual gene level were modest at best, the HR geneset was clearly downregulated in Grade 3 versus 0 (\( P = 0.008 \); Figure S3). For Grade 2 and Grade 1 versus Grade 0, the HR geneset was not significantly downregulated (\( P = 0.17 \) and \( P = 0.11 \), respectively; Figure S3). In our retrospective study a trend towards downregulation was observed for the genes involved in non-homologous end-joining (NHEJ) and its back-up pathway (BEJ) [2]. In the current study, such a trend could not be observed (Table S5). This was confirmed by ROAST geneset tests: Grade 3 versus 0, \( P = 0.38 \); Grade 2 versus 0, \( P = 0.83 \), Grade 1 versus 0, \( P = 0.17 \).

![Figure S1: Induction of gene expression upon irradiation calculated within each individual patient (indicated by the index on the x-axis) for three selected radiation-sensitive genes (Table S1). Note that the induction levels were sorted separately for each gene and that therefore expression levels for the same value of the sample index should not be compared.](image-url)
Figure S2: Enrichment of the set of 439 probesets included in the HALLMARK P53 PATHWAY geneset that was strongly induced upon irradiation (Table S2). The lower panel shows the moderated t-statistics on the x-axis, vertical bars indicate the position of the probesets included in the geneset. The upper panel contains a worm plot illustrating the enrichment of the geneset relative to random ordering.

Table S2: A CAMERA geneset test was performed to identify sets consisting of genes that are highly ranked in terms of differential expression upon irradiation relative to genes not in the set [3]. CAMERA was applied using genesets from the Hallmark and C2 collections from the Molecular Signatures Database (MSigDB v5.0, Entrez Gene ID version). P values were calculated for each gene set for two alternative hypotheses (up, down) and corrected for multiple testing using the Benjamini-Hochberg false discovery rate. For more information about the genesets, see http://software.broadinstitute.org/gsea/msigdb/genesets.jsp. Note that these results are based on a simple additive model using patients as a blocking variable, effectively performing a paired comparison between irradiated versus not irradiated lymphocytes.
5. CLASSIFICATION

Next we asked whether a classification model could be built for prediction of late radiation toxicity from HR gene expression levels. First, the patient-specific fold inductions (on log$_2$ scale) of the HR genes were calculated for the prospective cohort. The receiver operating characteristic (ROC) curve for the mean log$_2$ fold induction has an area under the curve (AUC) of 0.75 (95% CI: 0.6 – 0.89; calculated using DeLong’s method for determining AUC variance). At 100% specificity, such that all Grade 0 patients are correctly identified as negative for radiation toxicity, the sensitivity is 50% (Figure S4). Next a linear support vector machine (SVM) classifier was constructed for prediction of late radiation toxicity from the patient-specific HR log$_2$ fold inductions. Predictive performance was estimated using nested cross-validation (5-fold inner cross-validation, 3-fold outer cross-validation, with ten repeats). The parameter class.weights was used to set weights inversely proportional to the class frequencies, in order to compensate for the unbalanced class distributions. Optimal values for the ‘cost’ parameter were selected in the inner cross-validation loop (cost={2$^{-3}$; 2$^{-2}$;...; 2$^{10}$}) and predictive performance was calculated in the outer cross-validation loop (package MCRestimate). In the retrospective cohort 5 of 9 (sensitivity, 56%) of Grade 3 patients were correctly classified and 4 of 7 (specificity, 57%) of Grade 0 patients. In the prospective cohort 19 of 28 (sensitivity, 68%) of Grade 3 patients were correctly classified and only 5 of 14 (specificity, 36%) of Grade 0 patients.

Finally, a linear SVM classifier was fit on the retrospective cohort using the function ‘tune’ (package e1071) with 5-fold inner cross-validation to select optimal values for the ‘cost’ parameter. The resulting SVM was applied to the prospective cohort with 13 of 28 (sensitivity, 46%) Grade 3 patients correctly classified and 5 of 14 (specificity, 36%) Grade 0 patients.

<table>
<thead>
<tr>
<th>Probeset</th>
<th>Gene ID</th>
<th>Gene symbol</th>
<th>log$_2$(FC) (Grade 1 vs. 0)</th>
<th>log$_2$(FC) (Grade 2 vs. 0)</th>
<th>log$_2$(FC) (Grade 3 vs. 0)</th>
<th>P value</th>
<th>adj. P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>239689_PM.s</td>
<td>201745</td>
<td>RAB12</td>
<td>-0.14</td>
<td>-0.20</td>
<td>-0.43</td>
<td>2.2E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>214921_PM.s</td>
<td>2908</td>
<td>GYS2</td>
<td>0.09</td>
<td>0.19</td>
<td>0.07</td>
<td>5.8E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>239068_PM.s</td>
<td>209425</td>
<td>MAGI3</td>
<td>-0.09</td>
<td>0.00</td>
<td>0.15</td>
<td>6.3E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>155817_PM.s</td>
<td>2315</td>
<td>MLANA</td>
<td>0.21</td>
<td>0.23</td>
<td>0.29</td>
<td>7.8E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>230469_PM.s</td>
<td>2908</td>
<td>GYS2</td>
<td>0.21</td>
<td>0.23</td>
<td>0.29</td>
<td>7.8E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>230469_PM.s</td>
<td>2908</td>
<td>GYS2</td>
<td>0.21</td>
<td>0.23</td>
<td>0.29</td>
<td>7.8E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>230469_PM.s</td>
<td>2908</td>
<td>GYS2</td>
<td>0.21</td>
<td>0.23</td>
<td>0.29</td>
<td>7.8E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>230469_PM.s</td>
<td>2908</td>
<td>GYS2</td>
<td>0.21</td>
<td>0.23</td>
<td>0.29</td>
<td>7.8E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>230469_PM.s</td>
<td>2908</td>
<td>GYS2</td>
<td>0.21</td>
<td>0.23</td>
<td>0.29</td>
<td>7.8E-05</td>
<td>7.0E-01</td>
</tr>
<tr>
<td>230469_PM.s</td>
<td>2908</td>
<td>GYS2</td>
<td>0.21</td>
<td>0.23</td>
<td>0.29</td>
<td>7.8E-05</td>
<td>7.0E-01</td>
</tr>
</tbody>
</table>

Table S3: Top-10 probesets with a different response to irradiation between the four patient groups. For grade (1, 2, ≥ 3) the probesets differentially induced with respect to Grade 0 were determined by extracting the appropriate coefficients from the linear model fit with the nested interaction design described in the main text. Fold changes are given on a log2 scale. P values were calculated on the basis of the moderated F-statistics for the three comparisons. Correction for multiple testing was performed using the Benjamini-Hochberg false discovery rate (adj. P Value).
Late Radiation Toxicity due to Reduced DSB Repair

Figure S3: Volcano plots for the comparison of response to irradiation between Grade 3, 2, 1 and Grade 0 patients. Fold change is indicated on log₂ scale on the x-axis. The nominal P value is indicated on log₁₀ scale on the y-axis. Probesets that are part of the HR geneset (Table S4) are highlighted in red and labelled with their corresponding gene symbol. For LIG1, the most differentially expressed gene between Grade 3 and Grade 0 patients, the patient-specific fold inductions (on log₂ scale) are shown.

Table S4: Differential response to irradiation between Grade 3 and Grade 0 patients for the probesets part of the HR geneset. Fold changes are given on a log2 scale. P values were calculated on the basis of the moderated t-statistics. Correction for multiple testing was performed using the Benjamini-Hochberg false discovery rate (adj. P Value).

<table>
<thead>
<tr>
<th>Probeset</th>
<th>Gene ID</th>
<th>Gene symbol</th>
<th>log₂(FC)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>211022_PM.s_at</td>
<td>546</td>
<td>ATRX</td>
<td>-0.35</td>
<td>0.017</td>
</tr>
<tr>
<td>202726_PM.s_at</td>
<td>3978</td>
<td>LIG1</td>
<td>-0.18</td>
<td>0.009</td>
</tr>
<tr>
<td>205023_PM.s_at</td>
<td>5888</td>
<td>RAD51</td>
<td>-0.14</td>
<td>0.110</td>
</tr>
<tr>
<td>216299_PM.s_at</td>
<td>7517</td>
<td>XRCC3</td>
<td>-0.11</td>
<td>0.212</td>
</tr>
<tr>
<td>208368_PM.s_at</td>
<td>675</td>
<td>BRCA2</td>
<td>-0.10</td>
<td>0.408</td>
</tr>
<tr>
<td>207598_PM.x_at</td>
<td>7516</td>
<td>XRCC2</td>
<td>-0.09</td>
<td>0.456</td>
</tr>
<tr>
<td>211851_PM.x_at</td>
<td>672</td>
<td>BRCA1</td>
<td>-0.02</td>
<td>0.866</td>
</tr>
<tr>
<td>203061_PM.s_at</td>
<td>9656</td>
<td>MDC1</td>
<td>-0.00</td>
<td>0.983</td>
</tr>
<tr>
<td>210630_PM.s_at</td>
<td>5893</td>
<td>RAD52</td>
<td>0.01</td>
<td>0.938</td>
</tr>
</tbody>
</table>

Table S5: Differential response to irradiation between Grade 3 and Grade 0 patients for the probesets part of the NHEJ/BEJ geneset. Fold changes are given on a log2 scale. P values were calculated on the basis of the moderated t-statistics. Correction for multiple testing was performed using the Benjamini-Hochberg false discovery rate (adj. P Value).

<table>
<thead>
<tr>
<th>Probeset</th>
<th>Gene ID</th>
<th>Gene symbol</th>
<th>log₂(FC)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>232633_PM.s_at</td>
<td>7520</td>
<td>XRCC5</td>
<td>-0.32</td>
<td>0.017</td>
</tr>
<tr>
<td>210542_PM.s_at</td>
<td>5591</td>
<td>PRKDC</td>
<td>-0.10</td>
<td>0.363</td>
</tr>
<tr>
<td>219418_PM.s_at</td>
<td>7518</td>
<td>XRCC1</td>
<td>-0.06</td>
<td>0.300</td>
</tr>
<tr>
<td>208644_PM.at</td>
<td>142</td>
<td>PARP1</td>
<td>-0.05</td>
<td>0.210</td>
</tr>
<tr>
<td>204123_PM.at</td>
<td>3980</td>
<td>LIG3</td>
<td>-0.04</td>
<td>0.670</td>
</tr>
<tr>
<td>215308_PM.at</td>
<td>2547</td>
<td>XRCC6</td>
<td>-0.00</td>
<td>0.992</td>
</tr>
<tr>
<td>227766_PM.at</td>
<td>3981</td>
<td>LIG4</td>
<td>0.02</td>
<td>0.841</td>
</tr>
<tr>
<td>203655_PM.at</td>
<td>7515</td>
<td>XRCC1</td>
<td>0.03</td>
<td>0.703</td>
</tr>
<tr>
<td>205071_PM.x_at</td>
<td>7518</td>
<td>XRCC4</td>
<td>0.17</td>
<td>0.087</td>
</tr>
</tbody>
</table>
Chapter 6

Figure S4: Receiver operating characteristic curve for the mean log2 fold inductions of the HR geneset. Diagnostic accuracy was evaluated by determining the area under the receiver operating characteristic curve (AUC, pROC package). DeLong’s method for determining AUC variance was used for the calculation of the AUC 95% confidence interval.

REFERENCES


