Individual patient data meta-analysis shows a significant association between the *ATM* rs1801516 SNP and toxicity after radiotherapy in 5,456 breast and prostate cancer patients

C Nicolaj Andreassen[1], Barry S Rosenstein[2], Sarah L Kerns[3,2], Harry Ostrer[4], Dirk De Ruysscher[5], Jamie A Cesaretti[6], Gillian C Barnett[7,8], Alison M Dunning[7,8], Leila Dorling[8], Catharine ML West[9], Neil G Burnet[7], Rebecca Elliott[7], Charlotte Coles[7], Emma Hall[10], Laura Fachal[11], Ana Vega[11], Antonio Gómez-Caamaño[12], Christopher J Talbot[13] R. Paul Symonds[14], Kim De Ruyck[15], Hubert Thierens[15], Piet Ost[16], Jenny Chang-Claude[17,18], Petra Seibold[17], Odilia Popanda[19], Marie Overgaard[1], David Dearnaley[20], Matthew R Sydes[21], David Azria[22], C Anne Koch[23], Matthew Parliament[24], Michael Blackshaw[24], Michael Sia[25], Maria J. Fuentes-Raspall[26], Teresa Ramon y Cajal[27], Agustin Barnadas[27], Danny Vesprini[28], Sara Gutiérrez-Enríquez[29], Meritxell Mollà[30], Orland Díez[31], John R Yarnold[20], Jens Overgaard[1], Søren M Bentzen[32] and Jan Alsner[1]. On behalf of the International Radiogenomics Consortium (RgC)

[1]Department of Experimental Clinical Oncology, Aarhus University Hospital, Aarhus, Denmark
[2]Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, New York, USA
[3]Department of Radiation Oncology, University of Rochester Medical Center, Rochester, New York, USA
[4]Departments of Pathology and Pediatrics, Albert Einstein College of Medicine, New York, New York, USA.
[5]Department of Radiotherapy (Maastro Clinic), Maastricht University Medical Center, Maastricht, the Netherlands
[6]Southpoint Cancer Center, Jacksonville, Florida, USA
[7]Department of Oncology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
[8] Centre for Cancer Genetic Epidemiology, Strangeways Research Laboratory, University of Cambridge, Cambridge, UK.
[9]Institute of Cancer Sciences, University of Manchester, the Christie NHS Foundation Trust, Manchester, UK
[10]Clinical Trials & Statistics Unit (ICR-CTSU), The Institute of Cancer Research, London, UK
[11]Fundacion Publica Galega de Medicina Xenomica-SERGAS, Grupo de Medicina Xenomica-USC, IDIS, CIBERER, Santiago de Compostela, Spain
[12]Department of Radiation Oncology, Complexo Hospitalario Universitario de Santiago, SERGAS, Santiago de Compostela
[13]Department of Genetics, University of Leicester, Leicester, UK
[14]Department of Cancer Studies, University of Leicester, Leicester, UK
[15]Department of Basic Medical Sciences, Ghent University, Ghent, Belgium
[16]Department of Radiotherapy, Ghent University Hospital, Ghent, Belgium
[17]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[18]University Cancer Center Hamburg, University (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[19] Division of Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Heidelberg, Germany
[20]The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, UK
[21]MRC Clinical Trials Unit at UCL, London, UK
[22]Department of Radiation Oncology and Medical Physics, Institut regional du Cancer Montpellier, Montpellier, France
[23]Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada
[24]Department of Oncology, Cross Cancer Institute, University of Alberta, Edmonton, Alberta, Canada
[25]Department of Radiation Oncology, British Columbia Cancer Agency Abbotsford Clinic, British Columbia, Canada
[26]Department of Radiation Oncology, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
[27]Medical Oncology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
[28]Department of Radiation Oncology, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada
[29]Oncogenetics Group, Vall d'Hebron Institute of Oncology (VHIO), Universitat Autònoma de Barcelona, Barcelona, Spain
[30]Department of Radiation Oncology, Vall d'Hebron University Hospital, Barcelona, Spain
[31]Oncogenetics Group, Vall d'Hebron Institute of Oncology (VHIO); Area of Clinical and Molecular Genetics, Vall d'Hebron University Hospital, Barcelona, Spain
[32]Greenebaum Cancer Center and Department of Epidemiology & Public Health, University of Maryland School of Medicine, Baltimore, Maryland, USA

Corresponding author:

Christian Nicolaj Andreassen, MD, PhD, Associate professor
Department of Experimental Clinical Oncology
Aarhus University Hospital
Noerrebrogade 44
8000 Aarhus C
Denmark
nicolaj@oncology.au.dk
Phone: +45 7846 2620
Fax: +45 86197109

Running head: *ATM* rs1801516 SNP and toxicity after radiotherapy

Keywords: normal tissue toxicity; radiosensitivity; ataxia telangiectasia mutated; p.Asp1853Asn; c.5557G>A

Manuscript category: Original paper
Word count abstract: 184
Word count main text: 3.891
Number of references: 42
Number of tables: 2
Number of figures: 3
Number of supplementary files: 8

Conflicts of interest: None

We confirm that neither the submitted manuscript nor any similar manuscript, in whole or in part, other than two meetings abstracts, is under consideration, in press, published, or reported elsewhere.

Individual patient data meta-analysis shows a significant association between the *ATM* rs1801516 SNP and toxicity after radiotherapy in 5,456 breast and prostate cancer patients

Abstract

Purpose: Several small studies have indicated that the *ATM* rs1801516 SNP is associated with risk of normal tissue toxicity after radiotherapy. However, the findings have not been consistent. In order to test this SNP in a well-powered study, an individual patient data meta-analysis was carried out by the International Radiogenomics Consortium.

Material and methods: The analysis included 5,456 patients from 17 different cohorts. 2,759 patients were given radiotherapy for breast cancer and 2,697 for prostate cancer. Eight toxicity scores (overall toxicity, acute toxicity, late toxicity, acute skin toxicity, acute rectal toxicity, telangiectasia, fibrosis and late rectal toxicity) were analyzed. Adjustments were made for treatment and patient related factors with potential impact on the risk of toxicity.

Results: For all endpoints except late rectal toxicity, a significantly increased risk of toxicity was found for carriers of the minor (Asn) allele with odds of approximately 1.5 for acute toxicity and 1.2 for late toxicity. The results were consistent with a co-dominant pattern of inheritance.

Conclusion: This study convincingly showed a significant association between the *ATM* rs1801516 Asn allele and increased risk of radiation-induced normal tissue toxicity.

Background

Since 2003, more than 100 published studies have tried to link single nucleotide polymorphisms (SNPs) to the risk of normal tissue toxicity after radiotherapy [1]. Except for 7 studies [2-8], all took a candidate gene approach. One of the main purposes of this research is to identify genetic factors that can be used to establish a predictive assay for normal tissue radiosensitivity [9]. The vast majority of these studies has been underpowered to detect the small effect sizes usually found for SNPs [10]. Even though numerous significant associations were reported [10-12], only very few of these have been independently confirmed [13,14]. The gene product of *ATM* plays a crucial role in biological response to ionizing radiation. It is involved in the detection of DNA double strand breaks and initiation of pathways that lead to

cycle arrest followed by DNA repair or apoptosis [15]. Patients having a truncating mutation in both copies of the *ATM* gene suffer from the rare syndrome ataxia telangiectasia [16]. It has been reported that these patients develop devastating normal tissue toxicity if given radiotherapy [17]. At the turn of the millennium, *ATM* mutations were among the first to be investigated for a possible association with normal tissue radiosensitivity. Several studies looked for an association between heterozygocity for truncating *ATM* mutations and risk of normal tissue toxicity after radiotherapy (reviewed in [18]). These studies did not show any significant association. However, the limited size of the studies should be taken into account in the interpretation of these results. Numerous studies have investigated the impact of SNPs in *ATM* upon normal tissue toxicity. The SNP designated rs1801516 (c.5557G>A, p.Asp1853Asn) is among the *ATM* SNPs that have been most extensively investigated. This SNP results in a non-conservative amino acid substitution from an aspartic acid to an asparagine in exon 37 (NM_000051.3). A number of relatively small studies indicated that the minor allele of this SNP increases the risk of normal tissue toxicity after radiotherapy (reviewed in [19-21]). Nevertheless, the findings have not been entirely consistent. Three literature based meta-analyses have addressed the impact of the rs1801516 SNP upon radiation-induced normal tissue toxicity. Two of these meta-analyses demonstrated a significantly increased risk of acute toxicity (N=1,588)[20] and radiation-induced fibrosis (N=2,000) [21] respectively among carriers of the minor (Asn) allele whereas the largest of these meta-analyses (N=2,127) did not show a significant association between the SNP and risk of normal tissue toxicity in broader terms [19]. In order to bring this SNP to a test in the setting of a well-powered investigation, the International Radiogenomics Consortium (RgC) [22] conducted an individual patient data meta-analysis. With the intention to minimize the effect of publication bias, the analysis included published as well as unpublished data. In this paper, we intended the best possible adherence to the STROGAR guidelines for reporting results of genetic association studies in radiogenomics [23].

Material and methods

Study cohorts

Between January 2013 and September 2014, members of the RgC were encouraged to submit patient series (published as well as unpublished) for a meta-analysis addressing the impact of the *ATM* rs1801516 SNP upon normal tissue radiosensitivity. As part of this process, 17 cohorts (6 for prostate cancer and 11 for breast cancer) were

submitted and all of these were included in the present meta-analysis. The cohorts summed up to 5,624 patients. However, 168 patients were excluded from the analysis due to missing genotype data (n=31) or missing toxicity data (n=137) leaving 5,456 patients to be included in the analysis (Table 1). Consort diagram of the analysis is provided in Supplementary Figure S1. For three of the cohorts, the rs1801516 SNP was indirectly genotyped by means of the rs4988023 SNP (Table 1). This SNP is in strong linkage with the rs1801516 SNP (r^2=1) [11] and we therefore consider it an appropriate surrogate for the rs1801516 SNP. A brief description of the cohorts, genotyping procedures and scoring of the normal tissue reactions is given in Supplementary Text File 1 and in [5,6,8,13,14] and [24-30]. In addition to genotype and toxicity data, recordings of covariates with potential influence on radiation response were collected.

Statistical analysis

Overall analytical strategy

The strategy for statistical analysis was determined prior to data analysis. It was decided to analyze the dataset with regard to so called STAT scores (see below) for overall toxicity, acute toxicity and late toxicity. In addition, it was decided to analyze the dataset with regard to separate endpoints for which data were available from at least 1,000 patients. This was the case for acute skin toxicity (n=1,357), acute rectal toxicity (n=1,005), telangiectasia (n=2,404), fibrosis (n=2,457) and late rectal toxicity (n=2,215).

The following criteria were used to determine which data should be included in the analysis:

- Toxicity recorded during treatment or within a period of 3 months after treatment was included in the STAT scores for acute toxicity
- Toxicity recorded 2 years or more from the end of treatment was included in the STAT scores for late toxicity
- For cohorts that were scored more than once, the recordings closest to the end of treatment were used for the analysis of acute toxicity and the recordings at least 2 years after treatment that had the highest proportion of non-missing values were used for late toxicity
- All toxicity recordings were included in the STAT score calculations except for endpoints with a frequency of events below 5% (an event was defined as a toxicity score different from zero)
- For the 5 ‘separate endpoints’, data series with event frequencies below 5% were accepted

Using these criteria, 31,006 individual toxicity measurements were eligible for the meta-analysis. An overview of the data that were included in the analysis is provided in Table 1 and a more elaborate description is given in Supplementary Table S1. Recordings of covariates with potential influence on the normal tissue outcome were available for 3,666 patients from 13 cohorts (Supplementary Table S2). A total of 50,458 covariate observations were eligible for the analysis.

Z score conversion

In order to obtain a summary measure of toxicity across different endpoints, scoring systems and institutions, all analyzed toxicity recordings were converted into Z scores using the equation

$Z_k = (S_k - \text{mean}) / \text{standard deviation}$

in which Z_k designates the Z score for a patient k having the toxicity score S_k and the mean and standard deviation taken over all cases in the cohort with non-missing scores for that endpoint. This procedure produces a standardized normal variable, i.e. a toxicity score with a mean of zero and a standard deviation of one. By doing so, a measure is obtained of where the patient's toxicity score fits into the sensitivity spectrum of his or her original study cohort. Patients having a toxicity score that was in the right-hand tail of the distribution would be assigned a relatively high Z score whereas a patient with a score close to the population average for the endpoint in question would be assigned a Z score close to zero.

Calculation of STAT and Z STAT scores

Standardized Total Average Toxicity (STAT) scores [31] were calculated for each patient by taking the mean of the (non-missing) Z scores for different endpoints, thereby providing an overall measure of the patient's radioresponsiveness. STAT scores were calculated for overall toxicity (STAT global), acute toxicity (STAT acute) and late toxicity (STAT late).

For some of the cohorts, the STAT scores were based on only one or a few toxicity recordings whereas for others numerous recordings were used for the calculation. The distribution of the mean STAT score depends on the number of toxicity items used in forming the average (i.e. STAT scores derived from a large number of observations tend to have smaller standard deviations than STAT scores derived from a single or few toxicity recordings). To produce a standardized variable, we converted the STAT scores into Z scores (hereafter termed Z STAT). In this way, all Z scores and Z STAT scores have a common format characterized by a mean of zero and a standard deviation of one, making them directly comparable across different endpoints and cohorts.

Statistical test of the association between genotype and normal tissue toxicity

Analysis based on Z STAT scores and Z scores

For each of the three possible *ATM* rs1801516 genotypes we calculated the means of Z STAT global, Z STAT acute, Z STAT late, Z acute skin, Z acute rectal, Z telangiectasia, Z fibrosis and Z late rectal. Statistical significance was evaluated using the non-parametric Mann–Whitney *U* test (Wilcoxon rank-sum test) comparing carriers of one or two minor alleles with the common allele homozygotes. In addition, statistical significance was tested using Spearman's rank correlation coefficient as a test for trend over all three possible SNP genotypes.

Analysis of odds ratios for having a score in the upper quartile

In order to provide a measure of radiosensitivity that is easier to translate into a clinical context, we calculated odds ratios for having a Z STAT or Z score in the upper quartile comparing carriers of one or two minor alleles with the common allele homozygotes. Since the data typically had many ties, the proportion of patients having a score in the upper quartile occasionally deviated from 25%. The data were meta-analyzed and Forest plots created using the DerSimonian-Laird random effects-model [32]. The I-squared test was used to test for heterogeneity between cohorts [33]. Funnel plots [34] and Peters test for 'small study effects' [35] were used to test for indications of publication bias.

Multivariate analysis, calculation of residuals and analysis of odds ratios for having a residual in the upper quartile

As mentioned previously, information on covariates with potential impact on normal tissue outcome was available for 13 of the cohorts. For each of these cohorts, ordinal logistic regression was used to test for associations between the recorded patient and treatment related factors and the 8 toxicity scores. The tested covariates together with regression coefficients and p-values are listed in Supplementary Table S2. Multivariate analysis of the Z STAT and Z scores with the rs1801516 genotype was then performed on each cohort, including covariates associated with the toxicity score on univariate analysis with a p-value <0.2. For covariates that were closely related (e.g. various dose volume parameters for the same organ), the most significant one was chosen for the multivariate analysis. After multivariate analysis, residuals were calculated for each patient, for the STAT and Z scores (hereafter termed R-STAT and R-Z scores respectively). The analysis war carried out using STATA (StataCorp, Chicago, USA). A residual is the difference between the observed and the estimated toxicity score and provides a measure of the toxicity not explained by available patient and treatment related factors. Patients with residuals of zero have toxicity entirely accounted for by these factors. Patients with negative or positive residuals have less or greater toxicity respectively than is explained by known factors. In parallel with the analysis of the Z STAT and Z scores, we calculated odds ratios for having an R- STAT or R-Z score in the upper quartile comparing carriers of one or two minor alleles with the common allele homozygotes. The data were meta-analyzed and Forest plots created using the Der-Simonian-Laird random effects-model [32].

Results

Genotype distribution

The genotype distribution of the patients included in the analysis is shown in Supplementary Table S3. Approximately 76% were common allele homozygotes, 22% were heterozygotes whereas just below 2% were minor allele homozygotes. This corresponds to a MAF of 12.9%. These results are consistent with those reported in other studies and the genotype distribution did not deviate from Hardy Weinberg equilibrium using a significance threshold of 0.01.

Analysis based on Z STAT scores and Z scores

Figure 1 shows the mean Z STAT and Z scores as per genotype for the eight toxicity endpoints. Generally, patients with the variant allele had higher toxicity scores. The results were mostly consistent with a co-dominant pattern of inheritance. Furthermore, the strongest associations were observed for acute endpoints. Using the Mann–Whitney *U* test (Wilcoxon rank-sum test) as well as the Spearman rank correlation test, significant results were obtained for Z STAT global, Z STAT acute, Z STAT late, Z acute skin and Z acute rectal. For Z telangiectasia and Z fibrosis, the results did not reach statistical significance with p-values around 0.1. However, the results had the same direction as those reaching statistical significance. For late rectal toxicity, the analysis did not provide any indication that the SNP affects the normal tissue complication risk.

Analysis of odds ratios for having a score in the upper quartile

Figure 2 shows the Forest and funnel plots for the meta-analysis of the odds ratios for having a Z STAT global score in the upper quartile comparing patients with one or two minor alleles with the common allele homozygotes. Supplementary Figure S2 shows the Forest plots and funnel plots for the other endpoints. Table 2 provides a summary of the meta-analysis odds ratios for the 8 endpoints. In this analysis, significant associations were found for five of these endpoints but not for Z STAT late, Z acute rectal and Z late rectal. Generally, the associations were stronger for acute endpoints (with odds ratios around 1.5) than for the late endpoints (with odds ratios around 1.2). The I-squared test did not provide indications of significant heterogeneity between the cohorts for any of the endpoints. The funnel plots and Peters test for 'small study effects' did not provide indications of publication bias (Figure 2 and Supplementary Figure S2).

Multivariate analysis, calculation of residuals and analysis of odds ratios for having a residual in the upper quartile

Figure 3 shows the Forest and funnel plots for R-STAT global. Supplementary Figure S3 shows the Forest plot and funnel plots for the other endpoints. Table 2 provides an overview of the meta-analysis results. The adjustment for the included covariates only slightly altered the results compared to the unadjusted scores. Only the associations for R-STAT global, R-STAT acute and R-Z acute skin reached statistical significance.

Discussion

Methodological aspects

The primary strength of the present study is its very large sample size. Furthermore, the data were analyzed at the level of the individual patient and adjustments were made for treatment related factors with potential impact on toxicity risk. With more than 5,000 patients included in the analysis and a total number of toxicity recordings exceeding 30,000, the study is by far the largest of its kind. Nevertheless, this came at the expense of the cohorts included in the analysis being rather heterogeneous. Particularly for the STAT scores, very diverse data were pooled for the analysis. In order to allow for a comparison across different endpoints, scoring systems, treatment regimens and institutions, we converted the toxicity data into Z-scores. This method is generally accepted as a pragmatic solution to the problems that relate to heterogeneous study cohorts (31) and has often been used in collaborative radiogenomics studies (e.g. [6,13,14] and [30]). Furthermore, the method mitigates the so-called centre effect [36] wherein the institution at which the patient was treated may act as a confounder. Concerning the multivariate analysis, it was reassuring that the covariates included in the multivariate linear regression model were often factors known to affect the risk of normal tissue toxicity (such as body mass index/breast size, radiation dose and fractionation). Nevertheless, the adjustments made for these covariates only altered the results marginally. Since recordings of treatment and patient related covariates were only available for some of the cohorts, the number of patients included in the multivariate analysis was somewhat reduced. Presumably due to this, fewer of the findings reached statistical significance. The present meta-analysis included published as well as unpublished patient series. The funnel plots showed no obvious signs of publication bias, and the Peters test for 'small study effects' was negative for all the meta-analyses conducted. It is therefore unlikely that the results are significantly affected by publication bias. One base downstream of the rs1801516 a SNP designated rs1801673 (c.5558A>T, p.Asp1853Val) is located. This SNP is relatively rare with a minor allele frequency below 1% and was not taken into account in our meta-analysis. Our meta-analysis involved multiple comparisons. Despite this, we did not adjust for multiple testing. In most instances, the tests were not statistically independent and a standard adjustment for multiple comparisons would therefore be overly conservative. Furthermore, there was a high degree of consistency of the results across endpoints that were

not scored in the same patients. Consequently, we consider it highly unlikely that the positive findings of our meta-analysis could be the result of random fluctuations due to multiple testing.

Interpretation of the results

The present meta-analysis shows, with a high level of confidence, that the *ATM* rs1801516 SNP is associated with risk of normal tissue toxicity after radiosensitivity. Furthermore, the analysis indicates that the SNP has a stronger association with acute than late toxicity. The results are indicative of a co-dominant pattern of inheritance. With an odds ratio around 1.3, the phenotypic impact of this SNP is fairly small. This observation is in agreement with the results of many genome wide association studies (GWAS) addressing various biomedical phenotypes in which most of the identified SNPs only had a modest impact on phenotype [1,37,38]. The effect size of the rs1801516 SNP demonstrated by our meta-analysis is below the detectable effect size in the radiogenomics GWASs published so far (2-8). Furthermore, the small effect size explains why the candidate gene studies conducted did not report consistent findings for this SNP [19-21]. As mentioned in the introduction, three literature based meta-analyses have addressed the impact of the rs1801516 SNP upon risk of normal tissue toxicity after radiotherapy. One of these included 2,127 patients from 12 cohorts and addressed normal tissue toxicity in a broader sense [19]. The patients were treated for breast, prostate and head and neck cancer and investigated a mixture of different acute and late normal tissue reactions. This study did not show any significant association between the rs1801516 SNP and normal tissue outcome. Nevertheless, with a per-allele odds ratio of 1.11 (95% CI 0.77-1.59) for the Asn allele, our value is within the 95% confidence interval of the meta-analysis. Another meta-analysis included 1,588 patients from 5 different cohorts and addressed acute reactions after radiotherapy for prostate and breast cancer [20]. Consistent with our findings, this study reported an odds ratio of 1.33 (95%CI 1.04-1.77) for carriers of the minor (Asn) allele. The most recent meta-analysis included 2,000 patients from 9 different cohorts [21]. The patients were given radiotherapy for breast cancer, prostate cancer, head and neck cancer and brain tumours. This study addressed radiation-induced fibrosis. In agreement with our analysis, it showed a significantly increased risk of fibrosis (OR = 1.78, 95% CI: 1.07, 2.94) among carriers of the Asn allele. It should, however, be noted that all these meta-analyses had a number of cohorts in common with our study.

The rs1801516 SNP is located in a highly conserved element in the *ATM* gene (LOD score 250)(Supplementary Text File 2). It results in a non-conservative amino acid substitution from the acidic aspartic acid to the polar asparagine at codon 1853 in exon 37 (NM_000051.3). Thus, the SNP may potentially have a functional impact. Nevertheless, the SNP is classified as 'benign' by several different missense prediction algorithms (Align GVGD, SIFT, NCBI ClinVar and PROVEAN) (Supplementary Text File 2). Likewise, no change in natural splice sites is predicted by several splicing prediction tools (Splice Sequence Finder, MaxEntScan, NNSplice, and Human Splicing Finder). A cryptic and alternative acceptor splice site, located at the nucleotide 5558, is predicted to be removed by the rs1801516 SNP minor allele (Supplementary Text File 2). However, the consequence of this on alternative splicing is unknown. To our knowledge, only one study has investigated the functional impact of the rs1801516 SNP [39]. In this study, no significant differences were observed in the constitutive ATM protein level, cell survival or TP53 protein induction after radiation *in vitro* of lymphoblastoid cell lines from patients with each of the three possible rs1801516 SNP genotypes. Nevertheless, these investigations only included 'a limited number of patients' [39] and were therefore probably not powered to detect small differences in these parameters. The rs1801516 SNP is in linkage disequilibrium with numerous other SNPs [40](Supplementary Text File 2). Consequently, it cannot be ruled out that the SNP acts as a tagging SNP for another sequence variant with functional impact. Heterozygosity for a truncating mutation in *ATM* is a well-known risk factor for breast cancer [38]. Given that the rs1801516 SNP has a functional impact (or tags a variant with functional impact), one could expect this SNP to affect breast cancer risk. However, even though breast cancer susceptibility has been investigated in several very large GWASs that were powered to detect small effect sizes, the rs1801516 SNP was not identified as a risk factor [38]. Nor did a large meta-analysis (N=7,971) provide evidence that the rs1801516 affects breast cancer risk [40]. A search in the NHGRI-EBI GWAS database (http://www.ebi.ac.uk/gwas/) identifies only one hit for the rs1801516 SNP. This involves a large multi-centre GWAS with several replication cohorts addressing risk of malignant melanoma [41]. This study included a total of 10,422 subjects and identified an association for the rs1801516 SNP that reached genome wide significance with a p-value of 3.4×10^{-9}. Nevertheless, in this study the minor allele was protective with an odds ratio of 0.84. A more recent meta-analysis of data from 11 GWASs addressing malignant melanoma with a total of 15,990 cases and 26,409 controls confirmed the association with a p value of 6.6×10^{-12} [42]. As mentioned previously, the gene product of *ATM* is involved in the detection of DNA double strand breaks and initiation of pathways that leads to cell cycle arrest followed by DNA repair or apoptosis [15]. Thus, *ATM* plays a crucial role in the biological response to ionizing radiation and influences the fate of the cell following irradiation [16]. Acute radiation toxicity typically

takes place in skin or mucosal membranes and its pathogenesis is assumed to involve depletion of rapidly dividing cells in the basal layer [12]. In our meta-analysis, the rs1801516 SNP had a stronger impact on acute than late toxicity. Based on these observations, it can be hypothesized that the Asn allele in codon 1853 (or another variant tagged by this allele) may result in an increased tendency to cell cycle arrest or apoptosis. This would possibly lead to increased acute toxicity but at the same time be protective against radiation-induced malignancy. Our meta-analysis shows that it is possible to extract useful information from study cohorts that are very heterogeneous. However, it does at the same time illustrate the methods and large sample sizes required to detect a relatively weak signal in a noisy dataset. Thus, our study warrants further cooperative research initiatives in radiogenomics.

Conclusion

This very large individual patient data meta-analysis conducted by the RgC provides convincing evidence of an association between the Asn allele in the *ATM* rs1801516 SNP and an increased risk of normal tissue toxicity after radiotherapy with and odds ratio of approximately 1.5 for acute toxicity and 1.2 for late toxicity. The findings of this study warrant research into the functional impact of this SNP (or another sequence variant tagged by this SNP). In addition, further large studies quantifying the effect size of this SNP in the setting of a less heterogeneous study population should be encouraged.

Acknowledgements:

C Nicolaj Andreassen and Jan Alsner received funding from The Danish Cancer Society.

David Dearnaley, Emma Hall and John R Yarnold acknowledge NHS funding to the Royal Marsden Hospital and The Institute of Cancer Research NIHR Biomedical Research Centre.

Alison M Dunning was supported by a grant [C8197/A16565] from Cancer Research UK (CRUK).

Leila Dorling received funding from the UK Medical Research Council.

Charlotte Coles was supported by the Cambridge National Institute of Health Research Biomedical Research Centre.

Sara Gutiérrez-Enríquez was recipient of an Instituto de Salud Carlos III (ISCIII) grant FIS (PI05/2181) and is currently funded by a Miguel Servet contract awarded by the ISCIII of the 'Ministerio Español de Economía y Competitividad'.

The RADIOGEN project was partially funded by grants from the Instituto de Salud Carlos III (FIS PI10/00164 and PI13/02030) and the Fondo Europeo de Desarrollo Regional (FEDER 2007-2013).

Laura Fachal was supported by Xunta de Galicia and the European Social Fund (POS-A/2013/034).

Barry S Rosenstein was supported by grants from the National Institutes of Health (1R01CA134444), the Department of Defense (PC074201 and PC140371) and the American Cancer Society (RSGT-05-200-01-CCE).

References

1) Andreassen CN, Schack LM, Laursen LV, Alsner J. Radiogenomics - current status, challenges and future directions. Cancer Lett 2016. pii: S0304-3835(16)30005-2. doi: 10.1016/j.canlet.2016.01.035. [Epub ahead of print]. PMID:26828014
2) Kerns SL, Ostrer H, Stock R et al. Genome-wide association study to identify single nucleotide polymorphisms (SNPs) associated with the development of erectile dysfunction in African-American men after radiotherapy for prostate cancer. Int J Radiat Oncol Biol Phys 2010;78:1292-300.
3) Kerns SL, Stock R, Stone N et al. A 2-stage genome-wide association study to identify single nucleotide polymorphisms associated with development of erectile dysfunction following radiation therapy for prostate cancer. Int J Radiat Oncol Biol Phys 2013;85:e21-8.
4) Kerns SL, Stone NN, Stock RG, Rath L, Ostrer H, Rosenstein BS. A 2-stage genome-wide association study to identify single nucleotide polymorphisms associated with development of urinary symptoms after radiotherapy for prostate cancer. J Urol 2013;190:102-8.
5) Kerns SL, Stock RG, Stone NN et al. Genome-wide association study identifies a region on chromosome 11q14.3 associated with late rectal bleeding following radiation therapy for prostate cancer. Radiother Oncol 2013;107:372-6.
6) Barnett GC, Thompson D, Fachal L 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.

7) Fachal L, Gómez-Caamaño A, Barnett GC et al. A three-stage genome-wide association study identifies a susceptibility locus for late radiotherapy toxicity at 2q24.1. Nat Genet 2014;46:891-4.
8) Bourgier C, Kerns S, Gourgou S et al. Concurrent or sequential letrozole with adjuvant breast radiotherapy: Final results of the CO-HO-RT phase II randomized trial. Ann Oncol 2016;27:474-80.
9) Kerns SL, West CM, Andreassen CN et al. Radiogenomics: the search for genetic predictors of radiotherapy response. Future Oncol 2014;10:2391-406.
10) Andreassen CN. Searching for genetic determinants of normal tissue radiosensitivity--are we on the right track? Radiother Oncol 2010;97:1-8.
11) Barnett GC, Coles CE, Elliott RM 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.
12) Herskind C, Talbot CJ, Kerns SL, Veldwijk MR, Rosenstein BS, West CM. Radiogenomics: A systems biology approach to understanding genetic risk factors for radiotherapy toxicity? Cancer Lett 2016. pii: S0304-3835(16)30103-3. doi: 10.1016/j.canlet.2016.02.035. [Epub ahead of print] PMID: 26944314
13) Talbot CJ, Tanteles GA, Barnett GC et al. A replicated association between polymorphisms near TNFα and risk for adverse reactions to radiotherapy. Br J Cancer 2012;107:748-53.
14) Seibold P, Behrens S, Schmezer P et al. XRCC1 Polymorphism Associated With Late Toxicity After Radiation Therapy in Breast Cancer Patients. Int J Radiat Oncol Biol Phys 2015;92:1084-92.
15) van Gent DC, Hoeijmakers JH, Kanaar R. Chromosomal stability and the DNA double-stranded break connection. Nat Rev Genet 2001;2:196-206.
16) Chaudhary MW, Al-Baradie RS. Ataxia-telangiectasia: future prospects. Appl Clin Genet 2014;7:159-67.
17) Pollard JM, Gatti RA. Clinical radiation sensitivity with DNA repair disorders: an overview. Int J Radiat Oncol Biol Phys 2009;74:1323-31.
18) Andreassen CN. Can risk of radiotherapy-induced normal tissue complications be predicted from genetic profiles? Acta Oncol 2005;44:801-15.
19) Su M, Yin ZH, Wu W, Li XL, Zhou BS. Meta-analysis of associations between ATM Asp1853Asn and TP53 Arg72Pro polymorphisms and adverse effects of cancer radiotherapy. Asian Pac J Cancer Prev 2014;15:10675-81.
20) Dong L, Cui J, Tang F, Cong X, Han F. Ataxia telangiectasia-mutated gene polymorphisms and acute normal tissue injuries in cancer patients after radiation therapy: a systematic review and meta-analysis. Int J Radiat Oncol Biol Phys 2015;91:1090-8.
21) Zhang Y, Liu Z, Wang M et al. Single Nucleotide Polymorphism rs1801516 in Ataxia Telangiectasia-Mutated Gene Predicts Late Fibrosis in Cancer Patients After Radiotherapy: A PRISMA-Compliant Systematic Review and Meta-Analysis. Medicine (Baltimore) 2016;95:e3267.
22) West C, Rosenstein BS, Alsner J et al. Establishment of a Radiogenomics Consortium. Int J Radiat Oncol Biol Phys 2010;76:1295-6.
23) Kerns SL, de Ruysscher D, Andreassen CN et al. STROGAR - STrengthening the Reporting Of Genetic Association studies in Radiogenomics. Radiother Oncol 2014;110:182-8.

24) Andreassen CN, Overgaard J, Alsner J et al. ATM sequence variants and risk of radiation-induced subcutaneous fibrosis after postmastectomy radiotherapy. Int J Radiat Oncol Biol Phys 2006;64:776-83.
25) Andreassen CN, Alsner J, Overgaard M, Sørensen FB, Overgaard J. Risk of radiation-induced subcutaneous fibrosis in relation to single nucleotide polymorphisms in TGFB1, SOD2, XRCC1, XRCC3, APEX and ATM--a study based on DNA from formalin fixed paraffin embedded tissue samples. Int J Radiat Biol 2006;82:577-86.
26) Iannuzzi CM, Atencio DP, Green S, Stock RG, Rosenstein BS. ATM mutations in female breast cancer patients predict for an increase in radiation-induced late effects. Int J Radiat Oncol Biol Phys 2002;52:606-13.
27) Andreassen CN, Alsner J, Overgaard J et al. TGFB1 polymorphisms are associated with risk of late normal tissue complications in the breast after radiotherapy for early breast cancer. Radiother Oncol 2005;75:18-21.
28) Fachal L, Gómez-Caamaño A, Peleteiro P et al. Association of a XRCC3 polymorphism and rectum mean dose with the risk of acute radio-induced gastrointestinal toxicity in prostate cancer patients. Radiother Oncol 2012;105:321-8.
29) De Langhe S, De Ruyck K, Ost P et al. Acute radiation-induced nocturia in prostate cancer patients is associated with pretreatment symptoms, radical prostatectomy, and genetic markers in the TGFβ1 gene. Int J Radiat Oncol Biol Phys 2013;85:393-9.
30) Barnett GC, Elliott RM, Alsner J et al. Individual patient data meta-analysis shows no association between the SNP rs1800469 in TGFB and late radiotherapy toxicity. Radiother Oncol 2012;105:289-95.
31) Barnett GC, West CM, Coles CE 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;82:1065-74.
32) DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.
33) Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.
34) Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.
35) Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. JAMA 2006;295:676-80.
36) Andreassen CN, Barnett GC, Langendijk JA et al. Conducting radiogenomic research--do not forget careful consideration of the clinical data. Radiother Oncol 2012;105:337-40.
37) Manolio TA. Bringing genome-wide association findings into clinical use. Nat Rev Genet 2013;14:549-58.
38) Fachal L, Dunning AM. From candidate gene studies to GWAS and post-GWAS analyses in breast cancer. Curr Opin Genet Dev 2015;30:32-41.
39) Angèle S, Romestaing P, Moullan N et al. ATM haplotypes and cellular response to DNA damage: association with breast cancer risk and clinical radiosensitivity. Cancer Res 2003;63:8717-25.
40) Gao LB, Pan XM, Sun H et al. The association between ATM D1853N polymorphism and breast cancer susceptibility: a meta-analysis. J Exp Clin Cancer Res 2010;29:117.

41) Barrett JH, Iles MM, Harland M et al. Genome-wide association study identifies three new melanoma susceptibility loci. Nat Genet 2011;43:1108-13.
42) Law MH, Bishop DT, Lee JE et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. Nat Genet 2015;47:987-95.

Table 1. Data included from each cohort with total number of observations (Obs) and number of patients (Ptn). (b) = breast cohort. (p) = prostate cohort. Ref. indicates the reference in which the cohort is described. †indicates cohorts from which data on covariates with potential impact on the normal tissue outcome were available. *indicates that the rs1801516 SNP was indirectly genotyped by means of the rs4988023 SNP.

			STAT global		STAT acute		acute skin		Acute rectal		STAT late		Telengiectasia		Fibrosis		Late rectal	
Cohort	Ref.	N (all)	Obs	Ptn	Obs	Ptn	Obs	Ptn	Obs	Ptn	Obs	Ptn	Obs	Ptn	Obs	Ptn	Obs	Ptn
CIHR1 (b)†	14	158	632	158							632	158	158	158	158	158		
CIHR2 (b) †	14	78	465	78							465	78	78	78	78	78		
DBCG I (b) †	24	41	123	41	41	41	41	41			82	41	41	41	41	41		
DBCG II (b) †	25	234	702	234	234	234	234	234			468	234	234	234	234	234		
GenePARE (b)	8,26	179	358	179							358	179	179	179	179	179		
LeND (b)	13	583	1,165	583							1,165	583	583	583	582	582		
MARIE RAD (b) †	14	273	545	273							545	273	273	273	272	272		
Pre START (b)	27	52	104	52							104	52			52	52		
RAPPER (b) †*	6	940	11,651	940	925	925	925	925			10,726	929	858	858	737	737		
SANT PAU (b) †	30	101	390	101	99	99	99	99			291	101			96	96		
VHEBRON (b) †		58	86	58	58	58	58	58			28	28			28	28		
AERT RADIOGEN (p) †	28	724	4,304	724	2,896	724			724	724	1,408	470					468	468
CCI (p) †	5	155	465	155							465	155					155	155
GenePARE (p)	5	976	1,920	976							1,920	976					970	970
Ghent (p) †	29	281	1,964	281	1,964	281			1.964	281								
RAPPER CHHiP (p) †*	6	411	4,483	411							4.483	411					411	411
RAPPER RT01 (p) †*	6	212	1,649	212							1.649	212					211	211
All breast		2,697	16,221	2,697	1,357	1,357	1,357	1,357			14,864	2,656	2,404	2,404	2,457	2,457		
All prostate		2,759	14,785	2,759	4,860	1,005			2,688	1,005	9,925	2,224					2,215	2,215
All		5,456	31,006	5,456	6,217	2,362	1,357	1,357	2,688	1,005	24,789	4,880	2,404	2,404	2,457	2,457	2,215	2,215

Table 2. Overview of the meta-analysis odds ratios for having a Z or Z STAT score or a residual of these scores in the upper quartile (Asp/Asp and Asp/Asn vs. Asn/Asn). Significant findings are marked in bold.

Endpoint	N	Meta-analysis OR	95% CI	Endpoint	N	Meta-analysis OR	95% CI
Z STAT global	5,456	1.20	**1.04-1.38**	R-STAT global	3,339	1.24	**1.03-1.48**
Z STAT acute	2,362	1.49	**1.17-1.88**	R-STAT acute	2,300	1.51	**1.21-1.88**
Z acute skin	1,357	1.71	**1.11-2.66**	R-Z acute skin	1,304	1.60	**1.06-2.41**
Z acute rectal	1,005	1.38	0,93-2.03	R-Z acute rectal	996	1.25	0.70-2.24
Z STAT late	4,880	1.16	0.99-1.34	R-STAT late	2,773	1.18	0.97-1.45
Z telangiectasia	2,004	1.31	**1.05-1.65**	R-Z telangiectasia	1,593	1.23	0.82-1.84
Z fibrosis	2,457	1.27	**1.02-1.58**	R-Z fibrosis	1,602	1.18	0.86-1.62
Z late rectal	2,215	1.12	0.77-1.64	R-Z late rectal	1225	1.24	0. 90-1.72

Figure legends

Figure 1

Mean Z STAT and Z scores as per genotype for the eight normal tissue endpoints. Error bars indicate standard error of the mean (SEM).

Figure 2

Forest plot and funnel plot showing the meta-analysis result for having a Z STAT global score in the upper quartile (Asp/Asp and Asp/Asn vs. Asn/Asn).

Figure 3

Forest plot and funnel plot showing the meta-analysis result for having an R-STAT global score in the upper quartile (Asp/Asp and Asp/Asn vs. Asn/Asn).

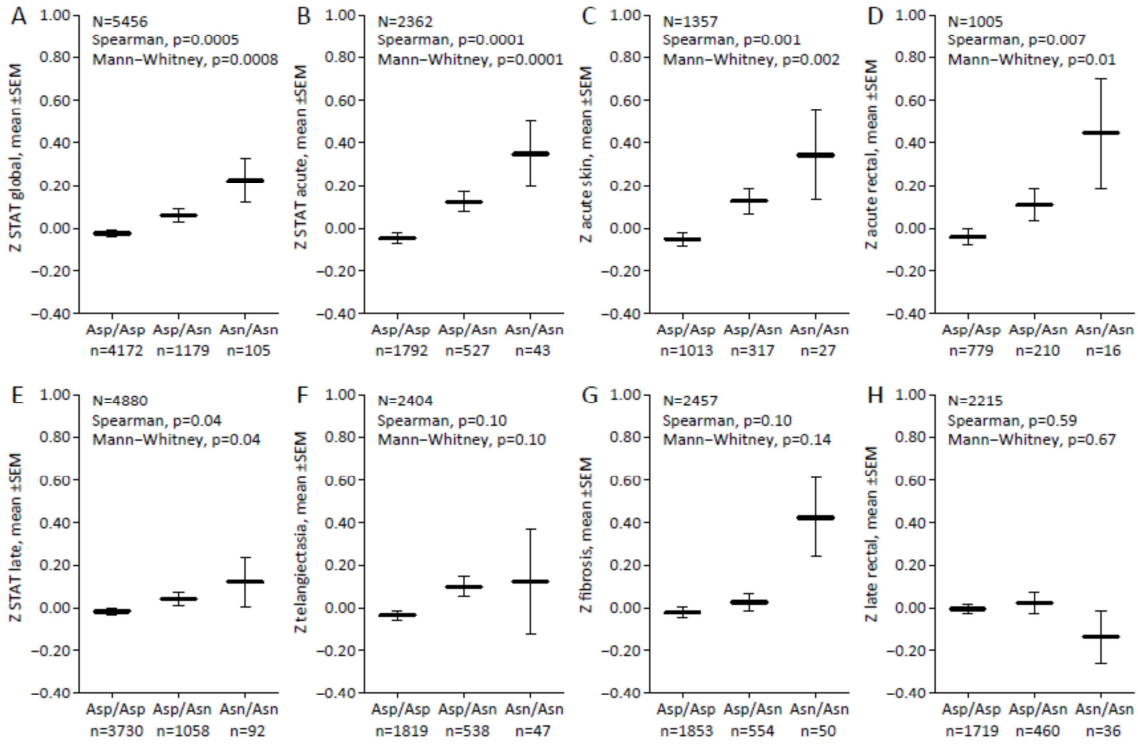

Figure 1

Mean Z STAT and Z scores as per genotype for the eight normal tissue endpoints. Error bars indicate standard error of the mean (SEM).

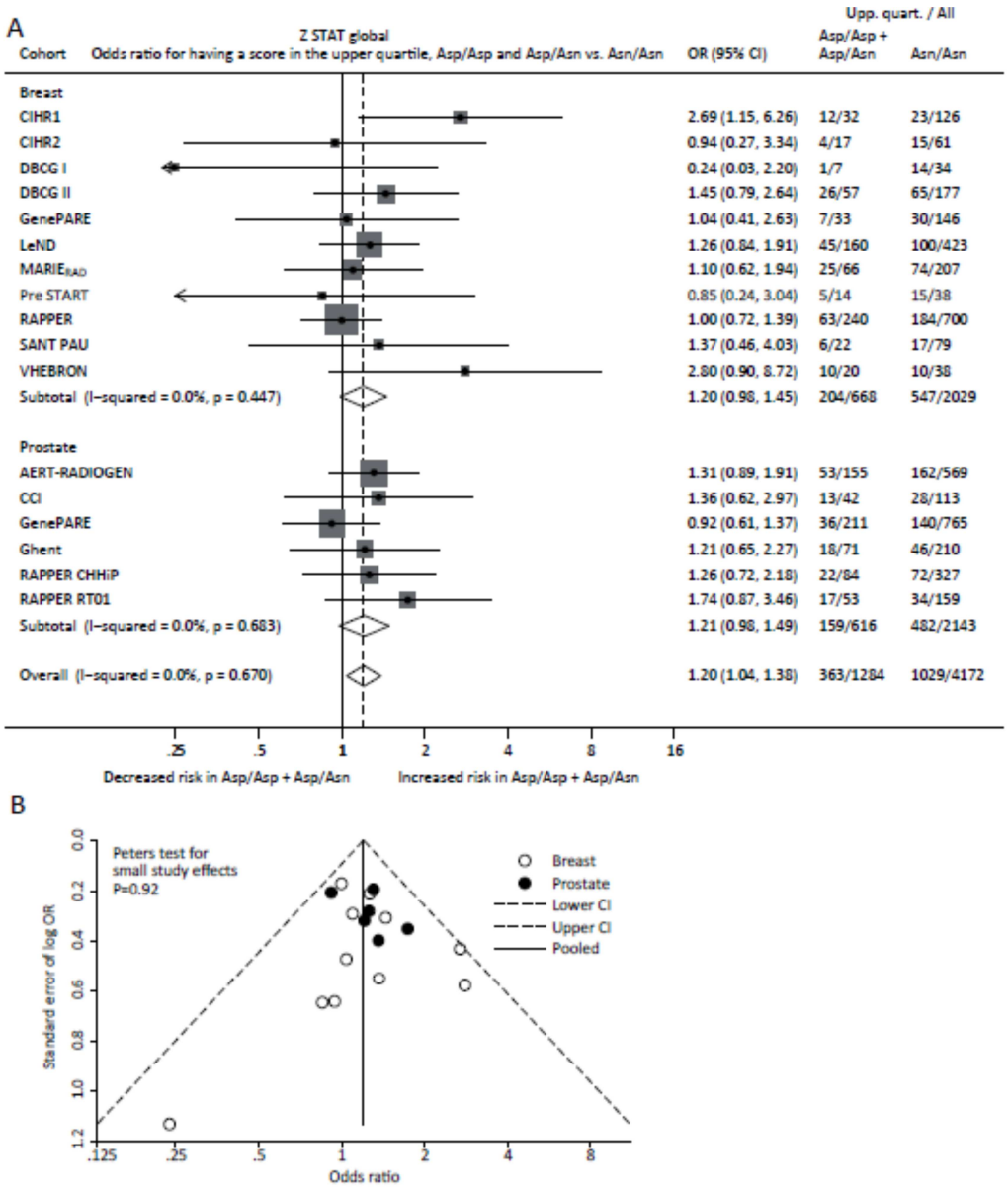

Figure 2

Forest plot and funnel plot showing the meta-analysis result for having a Z STAT global score in the upper quartile (Asp/Asp and Asp/Asn vs. Asn/Asn).

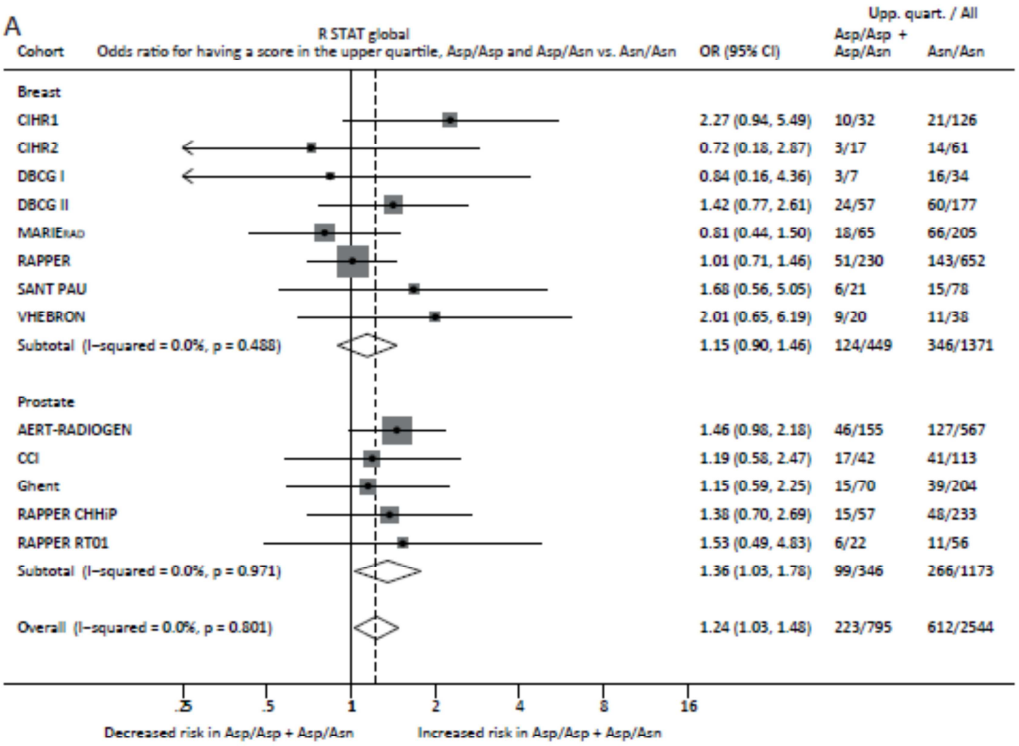

Figure 3

Forest plot and funnel plot showing the meta-analysis result for having an R-STAT global score in the upper quartile (Asp/Asp and Asp/Asn vs. Asn/Asn).