REVIEW

Radiogenomics helps to achieve personalized therapy by evaluating patient responses to radiation treatment

Zhen Guo¹, Yan Shu², Honghao Zhou¹, Wei Zhang¹,* and Hui Wang³,*

¹Department of Clinical Pharmacology, Xiangya Hospital, Central South University and Institute of Clinical Pharmacology, Central South University; Hunan Key Laboratory of Pharmacogenetics, Changsha 410008, P.R. China; ²Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201, USA and, ³Department of Radiation Oncology, Hunan Provincial Tumor Hospital & Affiliated Tumor Hospital of Xiangya Medical School, Central South University, Changsha 410013, P.R. China

*To whom correspondence should be addressed. Tel: +86 13973168774; Fax: +86 731 8235 4476; Email: yjsd2003@163.com. Correspondence may also be addressed to Hui Wang. Tel: +86 13973135460; Fax: +86 731 8865 1999; Email: wanghui710327@163.com

Abstract

Radiogenomics is the whole genome application of radiogenetics, which focuses on uncovering the underlying genetic causes of individual variation in sensitivity to radiation. There is a growing consensus that radiosensitivity is a complex, inherited polygenic trait, dependent on the interaction of many genes involved in multiple cell processes. An understanding of the genes involved in processes such as DNA damage response and oxidative stress response, has evolved toward examination of how genetic variants, most often, single nucleotide polymorphisms (SNPs), may influence interindividual radioresponse. Many experimental approaches, such as candidate SNP association studies, genome-wide association studies and massively parallel sequencing are being proposed to address these questions. We present a review focusing on recent advances in association studies of SNPs to radiotherapy response and discuss challenges and opportunities for further studies. We also highlight the clinical perspective of radiogenomics in the future of personalized treatment in radiation oncology.

Introduction

Radiation therapy is a key modality in the treatment of cancer. Moreover, when radiotherapy is combined with chemotherapy, surgery or other targeted therapies, treatment efficiency is improved and recurrence and cancer death rates are reduced (1). The main aim of radiotherapy is to maximize local control of the tumor while minimizing the damage to patient’s normal tissues. There is a spectrum of side effects (toxicities) in the surrounding normal tissues associated with radiotherapy (2). Acute/early toxicities, usually occur within 90 days after treatment, are transient and healing. However, late toxicities are more persistent and troublesome, which may include chronic inflammatory processes, vascular changes, fibrosis and atrophy (3). Tumor radiation response is the determining factor of the radiotherapeutic effect. How to elevate radiation response in tumor cells while decrease radiosensitivity in adjacent normal tissues has become a core issue in the tumor radiotherapeutic field (4). Previous interest has primarily focused on tumor radio-curability, but a shift toward cancer survivorship in recent years has seen growing interest in understanding radiation-induced complications in normal tissues (5).

A large patient-to-patient variability exists in radiotherapy response despite uniform treatment protocols (6). Although some of this may be ascribed to comorbidities, body habitus and stochastic factors (7–9), it has become increasingly believed that genetic background takes a leading role in the interindividual variation in radiosensitivity (10–13). Heritability analysis suggested that 58–78% of the variance can be attributed to genetic factors (14,15). The initial evidence for the genetic basis of radiosensitivity came from the observation that certain individuals with specific genetic disorders, such as ataxia telangiectasia, Nijmegen breakage syndrome, Fanconi anemia and DNA ligase IV deficiency, were hypersensitive to radiation (16). Around the
(G > A) may influence radiation sensitiv-
26–28 were more nal transduction pathways (damage repair, apoptosis and other responses too much damage by manipulating cell cycle checkpoints, DNA radioresistance or radiosensitivity which decides the cell’s fate ways to handle DNA lesions caused by ionizing radiation. The ability to predict a predisposition for severe radiotherapy-induced adverse effects in normal tissues could potentially aid treatment decision making and improve therapy effects. Avoiding or reducing radiation in these patients could lessen the likelihood of radiotoxicity-related morbidities and may also potentially reduce the burden of healthcare costs incurred for the supportive care required for these conditions (24). There has been intense interest in the phenomenon of inter-

**Biological pathways associated with radiosensitivity**

Radiosensitivity can be defined as the relative susceptibility of cells, tissues or organs to the effects of radiation (4). To preserve genomic stability, cells have developed elaborate response path-
ways to handle DNA lesions caused by ionizing radiation. The DNA damage response is a core determining factor of tumor radioresistance or radiosensitivity which decides the cell’s fate either to repair DNA damage or to undergo apoptosis if there is too much damage by manipulating cell cycle checkpoints, DNA damage repair, apoptosis and other responses via multiple signal transduction pathways (26–28). DNA damage-induced activation of cell cycle checkpoints at either G/S or G/M give cells time to employ DNA repair machineries to resolve DNA lesions, thus contributing to radioresistance (29,30). Multiple DNA repair pathways, including single-strand break repair, nonhomologous end joining repair, homologous recombination repair, base excisi-

**Candidate gene approach**

Genetic association studies have been employed to identify causal functional SNPs in normal tissue radiotoxicities. Most association studies have based on a candidate gene approach and investigated SNPs in genes involved in the complex radi-

toxicity-related pathways as mentioned above (Figure 2). Through a number of case-control studies, in which SNPs at candidate loci were genotyped across patients with or without radiother-

**DNA damage repair genes**

XRCC1, XRCC2, XRCC3, XRCC4, XRCC5

Among the DNA damage repair genes, XRCC1 is the most fre-

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC</td>
<td>breast cancer</td>
</tr>
<tr>
<td>GWAS</td>
<td>genome-wide association study</td>
</tr>
<tr>
<td>HNC</td>
<td>head and neck cancer</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NPC</td>
<td>nasopharyngeal carcinoma</td>
</tr>
<tr>
<td>OS</td>
<td>overall survival</td>
</tr>
<tr>
<td>RP</td>
<td>radiation pneumonitis</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
</tbody>
</table>

turn of the century, there was a growing consensus that the genetics behind individual radiosensitivity is more compli-
cated and must be considered as a so-called polygenic or com-

mRNA         mRNA         | messenger RNA       |
NPC         NPC         | nasopharyngeal carcinoma |
OS         OS         | overall survival   |
RP         RP         | radiation pneumonitis |
SNP        SNP        | single nucleotide polymorphism |

plex trait dependent on the combined influence of risk alleles with different frequencies and different relative risks (10). It is of importance to increase our understanding of the molecular pathogenesis of radiotherapy toxicity, find ways of predicting those patients likely to suffer with longer side effects and develop new approaches for their amelioration.

With the evolution of DNA sequencing and bioinformatics, radiogenomics has emerged as a new research field which studies the influence of genetic variations on radiation response (17–20). One aim of radiogenomics studies is to develop a risk model including genetic factors and clinical covariants capable of predicting individual radiosensitivity and likelihood of developing side effects after radiotherapy. Another aim is to improve the radiotherapeutic effect with individualized radiotherapy, through prevention or intervention of radiation-induced mor-

dities, by increasing our understanding of the biological mecha-
nisms behind them (4,21–24). Many cancer patients achieve survivorship at the cost of treatment complications occurring in normal tissues. The ability to predict a predisposition for severe radiotherapy-induced adverse effects in normal tissues could potentially aid treatment decision making and improve therapy effects. Avoiding or reducing radiation in these patients may lessen the likelihood of radiotoxicity-related morbidities and may also potentially reduce the burden of healthcare costs incurred for the supportive care required for these conditions (25). There has been intense interest in the phenomenon of inter-

**Candidate gene approach**

Genetic association studies have been employed to identify causal functional SNPs in normal tissue radiotoxicities. Most association studies have based on a candidate gene approach and investigated SNPs in genes involved in the complex radi-

**DNA damage repair genes**

XRCC1, XRCC2, XRCC3, XRCC4, XRCC5

Among the DNA damage repair genes, XRCC1 is the most fre-

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC</td>
<td>breast cancer</td>
</tr>
<tr>
<td>GWAS</td>
<td>genome-wide association study</td>
</tr>
<tr>
<td>HNC</td>
<td>head and neck cancer</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NPC</td>
<td>nasopharyngeal carcinoma</td>
</tr>
<tr>
<td>OS</td>
<td>overall survival</td>
</tr>
<tr>
<td>RP</td>
<td>radiation pneumonitis</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
</tbody>
</table>

quently researched candidate. Contrasting results have been reported on the impact of XRCC1 on radiation-induced normal tissue toxicities. Evidence from lung cancer patients showed that XRCC1 rs25487 (G > A) may influence radiation sensitiv-

35) and contribute to the development of grade ≥2 radia-
tion pneumonitis (RP) (36). As for nasopharyngeal carcinoma (NPC) patients, rs25487 A allele was conferred a reduced risk of late normal tissue complications after radiotherapy (37–39). Nevertheless, another independent study demonstrated that head and neck cancer (HNC) patients with the rs25487 AA/AG genotypes have a higher probability of radiation-induced mucositis (40). When it comes to breast cancer (BC) patients, no positive association has been found between rs25487 and acute toxicity (41).

Previous findings demonstrated that XRCC2 rs3218336 (G > A) was correlated with overall survival (OS) in non-small-cell lung cancer (NSCLC) patients receiving radiotherapy (42). In addition, results from classification and regression tree analysis displayed male patients with TT genotype of XRCC4 rs6869366 (G > T) and female patients with AG/AA genotypes of XRCC5 rs3835 (G > A) were also at increased risk of severe RP (43). Alsbeih et al. (37,38) have initially indicated the role of XRCC5 rs1051677 (T > C) allele in radiation toxicity among NPC patients. More recently, another study in Chinese NPC patients demonstrated a trend of possible association between rs861539 (C > T) of XRCC3 and radiation-induced fibrosis, but not rs25487 of XRCC1 (44). Additionally, BC patients with the variant allele of XRCC3 rs861539 were more likely to experience skin toxicity (45).

**RAD51**

The RAD51 gene family consists of several proteins that show DNA-stimulated ATPase activity and play a central role in the homologous recombination repair activation. A functional SNP
rs1801320 (G>C), located in the promoter region of RAD51, resulting in upregulated gene expression level through an increased promoter activity by substituting G for C allele, displayed a predictive value for RP development in NSCLC patients and dysphagia among HNC patients after radiotherapy (40), which is also an independent prognostic factor for OS in NSCLC patients (42). Moreover, a recent study firstly confirmed that RAD51 rs1801321 (G>T) T allele carriers may have a better OS as for cervical cancer patients upon radiotherapy (46).

ATM

ATM protein presents kinase activity induced by ionizing radiation, and it is involved in DNA damage detection and cell cycle progression (47). A number of studies have evaluated the role of ATM polymorphisms in RP developing for NSCLC patients upon radiotherapy. Two independent studies have consistently confirmed the rs18037 (G>A) A allele as a risk allele for RP (48-50). In both in vitro and in vivo studies demonstrated that rs18037 may affect ATM expression by reducing transcriptional activity and interfering nuclear protein binding (49). Another two SNPs of ATM [rs573759 (G>A) and rs228590 (C>T)] were also found to be correlated with RP (49,50). In addition, haplotype analysis showed that the rs18037/rs228590/rs1801516 (G-T-G) haplotype have a significantly lower risk of severe RP (hazard ratio = 0.52, 95% confidence interval: 0.35–0.79, P = 0.002) (50). The variant A allele of ATM rs1801516 (G>A) was correlated with increased
risk of severe fibrosis in NPC patients (37,38). For prostate cancer patients (51) and BC patients (52), however, there seems to be no association between rs1801516 and radiation toxicity.

**NBN**

NBN is a component of MRE complex (MRE11-RAD50-NBN) which is involved in damage sensing, signaling and responding to DNA double-strand breaks (53). HNC patients experiencing severe oral mucositis after radiotherapy were correlated to NBN rs1805794 (G > C) with an odds radio of 4.72 (54). However, negative results have also been found between rs1805794 and radiation toxicity in BC patients (55) and NSCLC patients (42).

**Oxidative stress response genes**

Radiation treatment exert its antineoplastic effects either directly by attacking cellular macromolecules, including DNA, or indirectly by generating reactive oxygen species and their by-products. TXNRD2 is a key player in the defense against oxidative damage, by reducing the oxidoreductase enzyme thioredoxin involved in reactive oxygen species scavenging (56). rs1139793 (C > T) was found significantly associated with messenger RNA (mRNA) expression level of TXNRD2 in blood. A recent study with a relatively long median follow-up time of >17 years has reported the association between radiation-induced subcutaneous fibrosis and rs1139793 in a cohort of 92 BC survivors, which was then confirmed in an independent cohort of 283 BC survivors (57). Subjects carrying polymorphic variant rs1695 (A > G) of GSTP1, an antioxidant enzyme, have shown a higher significant risk of developing fibrosis or fat necrosis, but not acute skin toxicity after radiation treatment in BC patients (58,59). The endothelial nitric oxide synthase participates in the oxidative stress response by catalyzing the production of the free radical nitric oxide. An increased risk of acute skin toxicity has been observed in BC patients with endothelial nitric oxide synthase rs1799983 (G > T) (41).

**Apoptosis-related genes**

Apoptosis occurs through two main pathways, the intrinsic pathway that is triggered by various intracellular stimuli, including DNA damage and oxidative stress, and the extrinsic pathway that is stimulated by binding of death ligands of the tumor necrosis factor superfamily to their cell surface death receptors. FASL, a member of the tumor necrosis factor superfamily, by interacting with its receptor FAS, can initiate the extrinsic apoptotic pathway. The expression of FAS and FASL is increased after irradiation and has been associated with radiation-induced cell death (60). FASL rs763110 (C > T) is located in the binding motif for the transcription factor CAAT/enhancer-binding protein β and has been shown to affect the biological activity of FASL promoter and FASL expression. Thurner et al. (61) have provided the first evidence that the variant allele of rs763110 may have a protective effect against the development of high-grade late rectal and/or urinary side effects after prostate cancer radiotherapy.

Polymorphisms of TP53 gene have been investigated as possible causes of normal tissue radiation sensitivity in numerous cancers. rs1042522 (G > C), the most frequently studied SNP of TP53 to date, is associated with the ability to induce cell apoptosis (62). Mountainous evidence have demonstrated that rs1042522 has a potential in predicting radiation responses, such as radiation-induced telangiectasia for BC patients (63), RP for lung cancer patients (48) and disease progression (local recurrence and distant metastases) for NPC patients (64). A study with a relatively small sample size of 48 prostate cancer patients noted a new SNP of TP53, rs35117667 (C > T), which was reported for the first time and has a predictive value for developing acute skin adverse effects (51). Evidence also showed that the intronic polymorphism at TP53 rs17883323 (C > A) was linked to high-grade urinary toxicity among prostate cancer patients (51). Additionally, TP73 rs3765701 (A > G) was associated with survival in stage III-IV NSCLC patients receiving chemoradiation therapy (65).

HDM2 is an essential negative regulator in the p53 pathway that directly binds to the p53 transactivation domain and inhibits its transcriptional activity. Initial results from Alsbeih et al. (37,38) have indicated that the variant allele of HDM2 rs2279744 (T > G) and rs1196333 (T > A) were associated with reduced risk to develop late normal tissues complications among NPC patients. The functional polymorphic variant in the promoter at rs2279744 has been suggested to increase the
transcriptional activator SP1 binding, resulting in higher levels of HDM2 mRNA and protein and the subsequent attenuation of the p53 pathway (66). On the other hand, Cintra et al. (51) noted that rs2279744 did not play a role in the incidence of radiotherapy-induced acute and chronic effects in prostate cancer patients.
Fibroblast proliferation gene

TGFβ1 is a versatile cytokine involved in inflammation, cell proliferation and fibrosis, which has been convincingly linked to the development of radiation-induced fibrosis (67). Multiple studies have observed an association among TGFβ1 SNPs and radiation toxicities, including erythema (52), fibrosis (59), erectile dysfunction (68), esophageal toxicity (69), radiation pneumonitis (70) and miscellaneous severe complications (38). Among them, the two most commonly studied SNPs are rs1800469 (C>T) and rs1982073 (T>C), which have been reported to be correlated with serum level of TGFβ1 (71,72). Unlike the typical studies that use the clinical radiation toxicities as research endpoint, Kelsey et al. (73) applied an objective radiologic endpoint, the slope of the dose–response curve, and confirmed that rs1800469 may affect radiation sensitivity. In addition, the rs1800469 might act as an independent predictor for radiotherapy outcomes as the CC genotype carriers showed better OS, whereas CT/TT genotypes carriers displayed a higher probability of severe radiation esophagitis in lung cancer patients (74,75). Meanwhile, certain attention should be given to ethnic differences as the evidence noted that rs1982073, associated with RP risk in whites, was not correlated with RP in Chinese patients (76,77).

Although a large number of exploratory studies and a small number of confirmatory studies have claimed the positive associations between TGFβ1 SNPs and the probability of developing late adverse effects of radiotherapy, it is remarkable that few studies with relatively large sample size failed to find any association. Strikingly, results from 124 BC patients’ lymphocytes and 15 human fibroblast strains clearly demonstrated that rs1800469 genotype has no effect on cellular radiosensitivity, TGFβ1 protein secretion or TGFβ1 expression (78). This is of particular importance, because it implies that none of the SNPs or SNP networks that affect TGFβ1 expression seem to be involved in the regulation of cellular radiosensitivity. Consistently, the absence of a link with radiation toxicity has just been confirmed for rs1800469 in BC patients and prostate cancer patients from different research groups with quite large sample size (79–82). So, even for TGFβ1 polymorphisms which are among the most extensively examined SNPs in normal tissue radiobiology, it is difficult to draw any definite conclusion.

Genome-wide approach

The recent development of GWASs has provided a radical alternative to the candidate gene approach. GWASs offer the opportunity to conduct a hypothesis-free survey for SNP associations without any need for a prior understanding of the biology underlying the phenotype of interest. For at least two reasons GWASs represent an important breakthrough in the attempts to unravel the genetics of complex traits. Above all, GWASs have dramatically increased the number of compelling SNP associations reported in the study of various phenotypes (83). Secondly, they shed important new light on the overall properties of the allelic architecture presumably underlying most complex phenotypes (84,85).

The first GWAS in radiogenomics aimed at identifying SNPs associated with erectile dysfunction following radiotherapy for prostate cancer in a specific African-American ancestry with a modest sample size and noted a significant SNP rs2268363 (86). With a two-stage design (discovery and replication), the same research group expanded the sample size and further conducted the GWAS in a cohort composed predominantly of men of European ancestry. Twelve SNPs, which lie in or near genes involved in controlling erectile function or cell adhesion and signaling, have been addressed to be correlated with erectile dysfunction (87). Meanwhile, they demonstrated that a region on chromosome 11q14.3 was associated with late rectal bleeding following radiotherapy for prostate cancer patients (88).

Cell line-based GWAS to identify genes associated with radiation sensitivity also yields certain novel biomarkers. Through the initial screening in 227 human lymphoblastoid cell lines and further functional validation using small interfering RNA knockdown in multiple tumor cell lines, it has been demonstrated that C13orf54, MAD2L1, PLK4, TPD52, DEPDC1B, were all significantly linked with radiosensitivity (89). More recently, Barnett et al. (90) have performed a phase-designed GWAS (1850 discovery + 1733 replication) with a replication phase in three independent cohorts who had radiotherapy for prostate or BC. Interestingly, none of the SNPs with potential associations to radiotherapy toxicity in this study have previously considered to be candidates. It is noteworthy to point that the study provided strong evidence that the identified SNPs were correlated with tumor site-specific toxicity (90).

As mentioned above, a striking finding in most GWASs is that the identified SNPs are not involved in any of the genes related to the major pathways assumed to be of importance for the investigated phenotype. Many of the identified SNPs are located in noncoding regions without any known function (87,90). This experience seriously questions the value of the candidate gene approach and broadens our understanding of the mechanisms underlying radiation-induced normal tissue complications at the same time. The ‘intrinsic’ genetics and biology have renewed enthusiasm for exploring functional implications as it may lead to alternative splicing or microRNAs regulation (91,92). Noncoding mutations with roles in radiosensitivity will likely open new doors to understand genome biology and gene regulation in radiation-induced side effects.

Even though GWASs have certainly provided several interesting insights and proven their worth, they do at the same time display a number of challenges. First of all, the SNPs detected through GWASs are mostly limited to ‘index SNPs’, which are a subset of common variants derived from haplotype blocks, and that will typically be inefficient to capture the impact of rare variants with minor allele frequency below 5% (93). In other words, it is unable to uncover the genetics of a complex polygenic trait once and for all. This is an issue of major importance since the genetic background of normal tissue radiosensitivity is made up by a spectrum ranging from rare highly penetrant alleles to common low-risk alleles (93). Deeper sequencing-based characterization of genomic variation, fine mapping, imputation and denser SNP array are needed as they can extend the reach of GWAS to ever lower ranges of minor allele frequency (94,95).

Although GWAS arrays are extraordinarily efficient at genotyping SNPs, they are less effective at capturing structural variations, such as insertions, deletions, inversions and copy number variants, which may also have a role in interindividual variation in radiation response (96,97). In addition, GWAS is a double-edged sword as it facilitates the initial identification of a region but makes it difficult to distinguish causal mutation(s), of which the challenge will be to separate true associations from the blizzard of false positives (91). Moreover, in order to keep the risk of false positives at a reasonable level and still obtain acceptable statistical power, most GWASs have been based on very large patient cohorts (98). For example, achieving 90% power to detect an allele with 5% frequency and a factor of 1.2 effect at a statistical significance of 10^{-4} requires 24,994 samples. For further discovery studies, larger GWAS will be necessary to include rare sequence variants through next generation sequencing. Besides,
rigorous re-testing in an independent validation set is also of utmost importance.

Challenge and opportunity

Though many positive results have been reported, the findings have always been inconsistent and lack the ability to replicate. Thus, none of the associations reported so far can by any means be regarded as unambiguously proven. This can presumably be attributed to certain methodological shortcomings. Based on theoretical considerations and experiences from other scientific fields, we address a number of critical points in order to successfully unravel the genetics of normal tissue radiosensitivity.

Clearly defined endpoint

The search for clinical and biologic biomarkers associated with late radiotherapy toxicity is hindered by the use of multiple and different endpoints from a variety of scoring systems, which make it tanglesome for comparisons across studies and pooling of data. In addition, although simple and useful in clinical, subjective scoring systems (e.g. mild, moderate, severe) for mixed endpoints (like breast appearance or urinary quality of life) cannot sufficiently distinguish the underlying biological mechanisms, which probably are determined by different genetic components. Thus, more quantifiable, biologically relevant phenotypes must be clearly and specifically defined, both to pinpoint the variable genetic components and to compare data across institutions. Barnett et al. (99) have proposed such a novel metric, the Standardized Total Average Toxicity score, to try to overcome these difficulties.

Confounding factors

Another major issue relating to the phenotypes is confounding factors. Briefly, not all studies pay equal attention to the potential confounding effects of differences in tumor site, target volume, radiotherapy technique, radiation dose, juxtaposed skin surface, overall treatment time, concomitant chemotherapy, body habitus and comorbidities (8,9). For example, Barnett et al. (100) have noted that a larger breast volume was the greatest nongenetic risk factor for the development of late toxicity in a cohort of 1014 BC patients. A decision tree considered clinicopathological variables and two SNPs demonstrated that mean lung dose was the strongest risk factor for RP among NSCLC patients (36). For late morbidities, the length of follow-up is a critical issue because the frequency and severity of late reactions tend to progress over time (101). Additionally, the differences between clinical specimen sources (e.g. blood, tissue, plasma), the time of specimen collection (e.g. before, during, after treatment) and the assays used for analysis (e.g. genomics, transcriptomics, proteomics) may also have subtle impact on outcomes (22). It is essential to account for all dosimetric-, patient- and treatment-related factors that may predispose to increased late effects in order to estimate the residual or unexplained toxicity likely due to genetics. As individualized radiotherapy is increasingly used, many centers have made great efforts to develop standardized procedures for storing and evaluating treatment plans and clinical outcomes.

The need for replication study

A very important aspect of any association study, whether it is derived from a candidate gene approach or whole genome analysis, is the need for replication studies validating the initial findings. With a large (N = 1613), well-powered prospective study, Barnett et al. (102) have failed to replicate previously reported individual radiation toxicity associations with candidate gene SNPs, implying that the published associations were either false positives or true weak positives. It is worth to highlight that a three-staged internal replication study, involving 1938 patients from three independent cohorts, has positively addressed that polymorphisms near tumor necrosis factor-α were associated with overall radiotherapy toxicities in BC patients (103).

Many instances have arisen in which initial findings have not been reproduced in follow-up studies (104). Considering that small sample size can provide imprecise or incorrect estimate of the magnitude of the observed effect, sufficient sample size is necessary in replication study to convincingly distinguish the proposed effect from no effect (105). Heterogeneity in clinical confounding factors and endpoint phenotypes between initial and replication studies can undermine the opportunity to compare among them. Because many initial studies have been reported in populations of European descent, the challenge remains to extend the studies to other populations (106,107). Data ‘dredging’ is also a major problem, especially when criteria for defining phenotypes are altered to achieve statistical significance worthy of publication (108). Efforts are underway to pool homogenous patient cohorts with well-documented radiotoxicity endpoints to provide sufficient statistical power to detect what will probably be modest functional effects. The National Cancer Institute-National Human Genome Research Institute working group on replication in association studies has published a comprehensive set of guidelines, providing a number of essential criteria for establishing positive replication studies (108).

Predictive models construction

Development of user-friendly tools for the prediction of radiation-induced complications support the patient and the physician in selecting the best therapeutic approach and avoid unnecessary worsening of quality of life. In the last years, numerous efforts have been performed to construct well-calibrated, integrated predictive models by combining genotyping profiles, clinical data and treatment parameters, which are suitable in clinical practice to identify patients at risk for developing radiotherapy-induced toxicities.

Multimetric normal tissue complication probability models, such as Lyman-Kutcher-Burman model (109), nearest-neighbor risk prediction model (110), relative seriality model (111) and Logit-equivalent uniform dose model (111), have been proposed to evaluate radiation-induced toxicities based on clinical and dosimetric parameters. By incorporating SNP information into the Lyman-Kutcher-Burman model based on mean lung dose, the predictive ability for severe radiation pneumonitis among patients with NSCLC has been significantly improved (109).

It is worth to note that a robust statistical method, the EMLasso technique, has been increasingly used for model selection in predictive radiogenetics. The EMLasso is a penalized logistic regression method with a stochastic expectation-maximization algorithm handling missing data and 10-fold cross-validation avoiding overfit (112). In light of this method, a multicomponent prediction model involving clinical, treatment and genetic parameters for grade ≥2 acute esophagitis postradiotherapy toxicity in lung cancer patients with a sensitivity of 84% and a specificity of 75.3% has been constructed by De Ruyck et al. (113). Clinical usefulness assessment of the model demonstrated that 69.4% patients predicted to suffer from acute esophagitis will actually develop this toxicity and will benefit from therapy modification. Following this, a number of prediction models achieved by EMLasso have been developed recently,
for example, late genitourinary complaints predictive model for prostate cancer (114), and dysphagia predictive model for HNC (115).

The development of EMLasso is appealing in constructing predictive models as it selects models with the smallest number of parameters while preserving predictive value, which really benefits clinical practice in consideration of time and cost efficiency. Cross-validation is another fortunate property of the EMLasso that avoids over/under-fitting. Besides, EMLasso is very suited for datasets with missing values which are common in radiogenetics studies due to the use of high-throughput genotyping techniques lacking a 100% call rate (115). Nevertheless, validation and assessment of clinical usefulness are necessary before implementing these models in the clinic for routine use. Clinical usefulness can be quantified by decision analytic methods, such as net benefit and decision curve analysis (116,117).

The biological mechanism of genotype–phenotype associations

The strong association between a genotype and a phenotype does not necessarily imply that the genetic variation is itself the cause. The true variants affecting a given phenotype can be located within genes, or they can be located in regions that have not yet been assigned any functional elements. Thus, understanding the biological mechanism is of utmost importance in genotype–phenotype association studies.

Since expression levels can be regulated by genetic variants in regulatory elements, examining the gene expression patterns after irradiation in vitro or in vivo with microarray to map the complex traits may shed new light on the pathways activated by ionizing radiation and help to identify novel candidate genes (118,119). For example, by comparing the expression profiles of radiation-resistant and radiation-sensitive NPC patient biopsy specimens with DNA microarray, Yang et al. (118) have demonstrated that 26 pathways (such as cell ion homeostasis, cell proliferation, receptor protein signaling, membrane system, humoral immune response, as well as cytokines and inflammation) were probably linked to radiation resistance with biostatistics and bioinformatics analysis. Although most studies have focused on mRNA profiles, proteomic studies will also be important because mRNA expression is not always correlated with protein expression or posttranslational modification (120,121). Experiments based on animal models, gene knockout and small interfering RNAs techniques could provide interesting insights into functional candidates (122,123). Besides SNPs, structural variations, such as insertions, deletions, inversions and copy number variants, may also have important implications on radiosensitivity (92). Furthermore, epigenetic changes such as DNA methylation, microRNA and histone modification, may represent an additional layer of complexity to understanding cancer cell biology and response to radiation treatment by posttranslational modifications (124–126). A recent study analyzing DNA methylation changes at >450000 loci after ionizing radiation has identified that the most differentially methylated genes were enriched in gene ontology categories relating to cell cycle, DNA repair and apoptosis pathways, with a radiation dose and postirradiation time-dependent manner (127). Evaluating the link among genetic variants, epigenetic modifications and disease predispositions is currently a very active area in research.

Large cohort, multicentered collaboration

Recent advances in technology and emerging insights into genetic variations provide unprecedented opportunities. Nonetheless, to fully exploit these new possibilities, international cooperation is essential. Several national and international cooperation genetic association studies have been initiated, such as GenePARE (128), RadGenomics (129), RAPPER (130) and REQUITE project (131). The International Radiogenomics Consortium has been established in 2009 in order to foster large-scale collaborative research projects (132). The future of the field is to create large patient cohorts for multiple cancer types, to validate the genetic loci and build reliable predictive models.

Conclusion

The identified predictors of radiosensitivity are aimed to ultimately contribute to an algorithm for guiding oncology treatment decisions. Patients at high risk of developing radiation-induced toxicities may be offered an alternative treatment approach, or, for patients who have received radiotherapy, advanced planning corrections can be introduced to better individualized radiation treatment. In addition to predictive and prognostic testing, the products of the identified genes could become targets for innovative therapies in susceptible individuals. Although we are still at the foot of the mountain, we have faith to believe that research on radiogenomics will help us achieve individualized radiotherapy protocols and broadly personalized medicine, leading to safer and more effective outcomes.

Funding

National Natural Science Fund (81273595); National High Technology Research and Development Program (2012AA02A518, 2012AA02A517).

Acknowledgement

We thank F.He, M.Liu and Z.Luo for helpful comments on the manuscript.

Conflict of Interest Statement: None declared.

References


