EDITORIAL

Radiogenomics: Radiobiology Enters the Era of Big Data and Team Science

Barry S. Rosenstein, PhD,*,†,‡ Catharine M. West, PhD,§
Søren M. Bentzen, PhD, DSc,‖ Jan Alsner, PhD,¶
Christian Nicolaj Andreassen, MD,¶ David Azria, MD, PhD,#
Gillian C. Barnett, MD, PhD,**,†† Michael Baumann, MD, PhD,‡‡
Neil Burnet, MD,§§ Jenny Chang-Claude, PhD, MHS,‖‖
Eric Y. Chuang, ScD,¶¶ Charlotte E. Coles, PhD,** Andre Dekker, PhD,##
Kim De Ruyck, PhD,*** Dirk De Ruysscher, MD, PhD,†††
Karen Drumea, MD,‡‡‡ Alison M. Dunning, PhD,††
Douglas Easton, PhD,§§§ Rosalind Eeles, PhD, FRCP, FRCR,‖‖‖
Laura Fachal, PhD,¶¶¶ Sara Gutiérrez-Enríquez, PhD,###
Karin Haustermans, MD, PhD,†††
Luis Alberto Henríquez-Hernández, PhD,**** Takashi Imai, PhD,††††
George D.D. Jones, PhD,‡‡‡‡ Sarah L. Kerns, PhD, MPH,*
Zhongxing Liao, MD,§§§§ Kenan Onel, MD, PhD,‖‖‖‖
Harry Ostrer, MD,¶¶¶¶ Matthew Parliament, MD,####
Paul D.P. Pharoah, PhD,***** Timothy R. Rebbeck, PhD,†††††
Christopher J. Talbot, PhD,‡‡‡‡‡ Hubert Thierens, PhD,***
Ana Vega, PhD,¶¶¶ John S. Witte, PhD,§§§§§ Philip Wong, MD,‖‖‖‖‖
and Frederic Zenhausern, PhD, MBA,¶¶¶¶¶, for The Radiogenomics
Consortium

*Department of Radiation Oncology and †Departments of Genetics and Genomic Sciences,
Dermatology, and Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, New
York; ‡Department of Radiation Oncology, New York University School of Medicine, New York, New
York; §Translational Radiobiology Group, Institute of Cancer Sciences, The University of Manchester,
Manchester Academic Health Science Centre, Christie Hospital, Manchester, UK; ‖Division of
Biostatistics and Bioinformatics, University of Maryland Greenebaum Cancer Center, and

*Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, Maryland; ¶Department of Experimental Clinical Oncology, Aarhus University Hospital, Aarhus, Denmark; #Department of Radiation Oncology, Institute of Cancer of Montpellier (INSERM), Center for Cancer Research, Montpellier, France; **Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK; ††Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Strangeways Research Laboratory, Cambridge, UK; ‡‡Department of Radiation Oncology and OncoRay National Center for Radiation Research in Oncology, Technische Universität Dresden, and Helmholtz-Zentrum Dresden-Rossendorf Dresden, Germany; §§University of Cambridge Department of Oncology, Cambridge Biomedical Campus, Addenbrooke's Hospital, Cambridge, UK; ‖‖Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany; ¶¶Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan; ##Department of Radiation Oncology (MAASTRO), GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands; ***Department of Basic Medical Sciences, Ghent University, Ghent, Belgium; †††Radiation Oncology, University Hospitals Leuven, and Department of Oncology, KU Leuven, Belgium; ‡‡‡Department of Oncology, Rambam Health Care Campus, Haifa, Israel; §§§Department of Public Health and Primary Care and Department of Oncology, University of Cambridge, Cambridge, UK; ‖‖‖Division of Cancer Genetics and Epidemiology, The Institute of Cancer Research, Sutton, UK; ¶¶¶Fundación Pública Galega de Medicina Xenómica-SERGAS, Grupo de Medicina Xenómica-USC, CIBERER, IDIS, Santiago de Compostela, Spain; ###Oncogenetics Group, Vall d'Hebron Institute of Oncology (VHIO), Universitat Autònoma de Barcelona, Barcelona, Spain; ****Department of Radiation Oncology, Hospital Universitario de Gran Canaria Dr. Negrin, Las Palmas de Gran Canaria, Spain; ††††Advanced Radiation Biology Research Program, Research Center for Charged Particle Therapy, National Institute of Radiological Sciences, Chiba, Japan; ‡‡‡‡Department Cancer Studies and Molecular Medicine, University of Leicester, Leicester, UK; §§§§Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas; ‖‖‖‖Department of Pediatrics, The University of Chicago, Chicago, Illinois; ¶¶¶¶Departments of Pathology, Genetics and Pediatrics, Albert Einstein College of Medicine at Yeshiva University, New York, New York; ####Department of Radiation Oncology, Cross Cancer Institute, University of Alberta, Edmonton, Alberta, Canada; *****Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK; †††††Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, and Abramson Cancer Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; ‡‡‡‡‡Department of Genetics, University of Leicester, Leicester, UK; §§§§§Department of Epidemiology and Biostatistics, Institute for Human Genetics, University of California, San Francisco, California; ‖‖‖‖‖Department of Radiation Oncology, Centre Hospitalier de l'Université de Montréal, Montréal, Quebec, Canada; and ¶¶¶¶¶Department of Basic Medical Sciences, Center for Applied Nanobioscience and Medicine, College of Medicine Phoenix, University of Arizona, Phoenix, Arizona

Radiogenomics is the study of the link between germ line genotypic variations and the large clinical variability observed in response to radiation therapy. The radiogenomics hypothesis is that a proportion of the variance in the phenotype of interest—radiation toxicity—is explained by genotypic variation. Thus, the aim of radiogenomics is to identify the alleles that underlie the inherited dissimilarities in phenotype. However, this hypothesis does not assume that all of the phenotypic differences are due to germ line genetic alterations, but it acknowledges that epigenetic changes (inherited and acquired) and other factors could also be important.

To foster collaborative research, the Radiogenomics Consortium (RGC) was established in 2009 (1). The RGC currently has 174 members from 90 institutions in 20 countries and is a cancer epidemiology consortium supported by the National Cancer Institute (NCI) (2). The goal of the RGC is to facilitate large-scale collaborative research assessing gene—radiation effect relationships, including genome-wide association studies (GWAS). The aim of this research is to produce assays for use in the clinic to predict risk of toxicities after radiation therapy, given alone or in multimodality treatments. The results of this research could also lead to the identification of novel strategies for prevention or mitigation of toxicities.

Currently, the radiogenomics landscape is rapidly changing. This is partly a result of research advances within the field of radiogenomics itself and partly due to progress in biotechnology and medical informatics that facilitate the pursuit of novel discovery strategies. Radiogenomics is now rapidly advancing from an effort to screen a limited number of candidate genes toward an open discovery

approach or, from relatively small studies conducted by a few scientists to the potentially powerful, but challenging, era of big data and team science.

However, the substantial progress in radiogenomics may be largely unnoticed by the broader radiation therapy community, leading to underestimation of its potential for improving patient outcomes. Therefore, it is critical to address the following questions at this stage in the development of research in radiogenomics.

What Have We Learned From Candidate Gene Studies?

The pioneering radiogenomics studies conducted over the past decade were hypothesis driven. The underlying hypothesis took advantage of a considerable body of knowledge about genes involved in the recognition and repair of DNA damage and the molecular biology of normal tissue response and remodeling after exposure to ionizing radiation. The candidate gene hypothesis was that single-nucleotide polymorphisms (SNPs) could cause a patient-specific variability in the level or function of proteins involved in radiation damage processing. The early studies were hypothesis generating and initially found numerous SNPs associated with variability in radiosensitivity. However, it has not been possible to validate the association with radiation toxicity for any of the approximately 100 SNPs that were identified by use of the candidate gene approach (3, 4). Most researchers therefore concur that the candidate gene studies tell us that common sequence alternations in the most obvious genes are unlikely to explain a significant fraction of patient-to-patient variability in radiation response. However, any study will have statistical power to detect only an association of a given minimum effect size. This means that we cannot rule out that SNPs in candidate genes have some effect on the risk of toxicity, but it appears safe to say that there are now sufficient data to rule out large effect sizes in the most obvious candidate genes.

Have the Goals of Radiogenomics Been Achieved Through the Performance of GWAS?

The GWAS approach being pursued in radiogenomics provides a new path toward the creation of predictive assays that can be incorporated into routine clinical care. However, this goal has not yet been achieved in radiogenomics because: (1) cohorts have only recently become large enough to have the power to identify genetic variants with a minor allele frequency above 5% and associated with clinically relevant odds ratios for the development of a given effect in the range of 1.2 to 1.5 at stringent levels of statistical significance ($P<5 \times 10^{-8}$); and (2) the associations identified have not been tested rigorously in independent validation studies. Ongoing RGC research is

addressing these 2 important limitations by moving from single-institution studies to multinational team science. Recent studies show that many genetic associations will be identified with the use of large multisite cohorts and contemporary techniques of tagging SNPs, imputation, and meta-analysis (3). It is noteworthy that the NCI has embraced this multisite, multicohort GWAS approach as the key to identifying genetic risks for cancer and treatment outcomes through their support of the Genetic Associations and Mechanisms in Oncology (GAME-ON) initiative (5). Radiogenomics research follows a similar approach.

Has the Validity of GWAS for Identifying Genetic Variations Responsible for Human Disease Been Proved?

GWAS provide a comprehensive and unbiased assessment of SNPs across the genome and test for their association with disease phenotypes. However, the validity of GWAS in identifying genetic variations responsible for human disease is not universally accepted (6), and there are divergent views about the clinical relevance and utility of the associated SNPs (7). However, there is no doubt that GWAS have revolutionized the identification of common genetic variations reported as SNPs associated with human traits and disease susceptibility. Since 2005, GWAS have identified approximately 2000 new human trait/disease susceptibility loci (8). For example, analysis of approximately 50,000 patients with prostate cancer and control individuals identified more than 70 cancer susceptibility loci, explaining approximately 30% of the familial risk for this disease; the top 1% of the risk distribution having a 4.7-fold increased risk of prostate cancer (9). Another illustration of the success of GWAS was the identification of a variant associated with an increased risk for the development of second malignant neoplasms in survivors of Hodgkin lymphoma treated with radiation therapy. The variant is common and penetrant, but only in the context of radiation; thus, it has potential to contribute to clinical decision making (10). In parallel, RGC investigators are making substantial progress to identify SNPs associated with radiation toxicity, particularly because the sample sizes and therefore statistical power of radiogenomics studies are increasing markedly (3).

Are GWAS Likely to Identify Biologically Relevant Causal Relationships?

A frequently aired misconception is that GWAS have primarily identified common variants that are statistically correlated with various conditions but that have no established biological function and may even be located in noncoding regions of the genome and therefore not be of

biological importance (6, 11). The argument typically continues along the line that these genetic alterations represent false positive results, whereas rare variants with a high level of phenotypic penetrance are responsible for most inherited disease susceptibility. It may be correct that rare mutations could be responsible for a significant proportion of human disease, but this does not preclude an important effect of common variants. Although the "missing heritability" not explained by identified variants following the common disease—common variant model leaves room for the common disease-rare allele model, it is generally considered that both models are valid and complementary (12). As a demonstration that these 2 concepts are not mutually exclusive, it should be noted that once the common functional variants driving a significant SNP association have been discovered through a GWAS, rarer variants with similar modes of action are frequently found within the same locus (13). Furthermore, there is no fundamental problem in the fact that most SNP associations were identified without prior knowledge of biological function (8). Also, the argument that common variants associated with disease are inconsequential because they lie in noncoding regions of DNA is countered by the ENCODE project, which has clearly demonstrated that this portion of the genome is functionally significant (14). In addition, the functional mechanisms for specific SNPs located in noncoding regions have now been confirmed (15). New post-GWAS studies are beginning to show that many SNPs associated with complex diseases act in a trans manner, wherein the SNP affects the expression of a gene that lies in a completely different locus (16).

Should the Emphasis of Research in Radiogenomics Be on the Performance of Mechanistic Studies?

A question can be posed as to whether the focus of research in radiogenomics should now be shifted to functional studies to determine how SNPs affect the radiation responses of normal tissues. Although mechanistic studies are important, this direction in research is premature until the SNPs associated with radiosensitivity have been validated with a high level of certainty and fine mapping has been performed to identify the causative variants. Also, before the achievement of a complete understanding of the mechanism underlying an effect, prediction in clinical medicine can still be of great importance. Although it could take many years to complete the functional studies needed to characterize the biological mechanism underlying SNPs found to be associated with radiation therapy toxicity, this piece of the puzzle is not necessarily required for the purpose of developing predictive models. In traditional epidemiology, many known risk factors play an important role in prevention and screening programs, even though the precise mechanisms linking them with disease are poorly understood.

Can We Conduct the Large Clinical Studies Required?

GWAS require very large sample sizes to ensure resolution of clinically relevant effect sizes while keeping the probability of false positive findings low. Virtually all data for patients included in published radiogenomics studies were obtained prospectively by the use of validated instruments for recording and grading toxicities. One challenge is pooling data collected using different toxicity recording dictionaries, radiation therapy delivery platforms, and patient populations (17). This requires large collaborative teams of clinicians, nurses, biologists, medical physicists, data managers, biostatisticians, and bioinformaticians—ie, team science! However, an issue that could be raised is whether this heterogeneity confounds the results of GWAS based on such studies (18) and whether the predictive ability of the genetic markers is generalizable for different radiation therapy approaches. In this light, it is reassuring to note that despite these variations, SNPs have been discovered that are associated with toxicities and replicate in independent cohorts involving treatment with multiple forms of radiation therapy (3, 19). The discovery of genetic markers that apply to patients independently of radiation therapy schedules is particularly important because it demonstrates that the assays being developed by RGC investigators will be applicable across radiation therapy modalities. With this in mind, it can be argued that the diversity of radiation therapy approaches in the RGC cohorts needs to be embraced, rather than avoided, to ultimately decide which treatment is best for each individual patient (20).

It should be noted that even though SNPs have been successfully identified in cohorts treated under multiple protocols, the RGC is working toward the standardization of toxicity data collection through publication of a recommended common set of guidelines for reporting radiogenomic studies (21). In addition, through a project funded by the European Union, RGC investigators are collecting detailed treatment, comorbidity, and toxicity data prospectively for 5300 patients with prostate, breast, or lung cancer undergoing radiation therapy in a multicenter study using identical treatment and toxicity data collection forms.

What Is the Potential Clinical Utility of Radiogenomics?

Tumor eradication is the primary aim of radiation therapy. However, with an increasing number of long-term cancer survivors—many of whom received radiation therapy—the impact of radiation-induced injuries on health-related quality of life is a growing concern. Moreover, research on the use of patient decision aids supports the concept that where choices for therapy exist, information that supports a

patient's understanding of the potential harms and benefits of treatment options encourages more active involvement of patients in decision making, resulting in decreased anxiety and decision regret.

In addition, the risk of severe toxicity in a minority of patients limits the potentially curative doses prescribed to the majority. Modeling studies support the hypothesis that identification of patients with a low risk for the development of adverse effects should allow for dose escalation and improvements in tumor control. Radiogenomics has therefore focused initially on the identification of SNPs associated with radiation therapy toxicity.

It is clear that the development and application of a set of genetic biomarkers is indeed the epitome of patient-centered oncology care. However, better alignment of patient needs and provision of information by physicians will be required to achieve this goal. Shared decision making between patients and their radiation oncologists that balances the probabilities of freedom from disease progression with freedom from toxic normal tissue effects is increasingly desired by patients, providers, and insurers alike (22). A set of validated biomarkers should therefore become part of a patient decision aid that incorporates patient values into a personalized radiation treatment approach (23). However, as has been recognized (6), there are significant obstacles to the routine clinical use of such biomarkers, inasmuch as their implementation would require clear interest and support from physicians, patients, and insurance companies; unfortunately, the incentive structure is not clearly aligned with achieving this goal.

In summary, collaborative efforts in radiogenomics fostered through the RGC are underpinning progress in identifying genetic variants that predict the risk of toxicity after radiation therapy. Radiogenomics investigators have already proved that we have the team to deliver the large data sets required. It is the nature of research that the end result can never be guaranteed, but what we can guarantee is the quality of our science. The research currently under way, which uses state-of-the-art techniques used successfully in genetic studies of other complex phenotypes, may well lead to the development of predictive assays ready for implementation in the routine clinical care of cancer patients. Thus, in the future, it is anticipated that genetic profiling will play a critical role in helping to guide cancer patients and their doctors to select the best treatment approach or radiation therapy dose for each individual.

References

1. West C, Rosenstein BS, Alsner J, et al. Establishment of a radiogenomics consortium. *Int J Radiat Oncol Biol Phys* 2010;76:1295-1296.

2. Radiogenomics Consortium (RGC): Epidemiology and genomics research in NCI's Division of Cancer Control and Population Sciences.

Available at http://epi.grants.cancer.gov/radiogenomics. Updated 2014. Accessed February 1, 2014.

3. Kerns SL, Ostrer H, Rosenstein BS. Radiogenomics: Using genetics to identify cancer patients at risk for development of adverse effects following radiotherapy. *Cancer Discov* 2014;4:155-165.

4. 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.

5. Genetic Associations and Mechanisms in Oncology (GAME-ON): A network of consortia for post-genome wide association (post-GWA) research. Available at http://epi.grants.cancer.gov/gameon/. Updated 2014. Accessed February 1, 2014.

6. ASTRO Cancer Biology/Radiation Biology Taskforce. Current status and recommendations for the future of research, teaching, and testing in the biological sciences of radiation oncology. Available at http://www.astro.org/uploadedFiles/Content/Research/Biologytaskforce-report.pdf. Updated 2014. Accessed February 1, 2014.

7. Devilee P, Rookus MA. A tiny step closer to personalized risk prediction for breast cancer. *N Engl J Med* 2010;362:1043-1045.

8. Manollio TA. Bringing genome-wide association findings into clinical use. *Nat Rev Genet* 2013;14:549-558.

9. Eeles RA, Olama AA, Benlloch S, et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. *Nat Genet* 2013;45:385-391.

10. Best T, Li D, Skol AD, et al. Variants at 6q21 implicate PRDM1 in the etiology of therapy-induced second malignancies after Hodgkin's lymphoma. *Nat Med* 2011;17:941-943.

11. McClellan J, King MC. Genetic heterogeneity in human disease. *Cell* 2010;141:210-217.

12. Gibson G. Rare and common variants: Twenty arguments. *Nat Rev Genet* 2012;13:135-145.

13. Edwards SL, Beesley J, French JD, et al. Beyond GWASs: Illuminating the dark road from association to function. *Am J Hum Genet* 2013;93:779-797.

14. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489:57-74.

15. French JD, Ghoussaini M, Edwards SL, et al. Functional variants at the 11q breast cancer risk loci regulate cyclin D expression through long-range enhancers. *Am J Hum Genet* 2013;92:489-503.

16. Smirnov DA, Morley M, Shin E, et al. Genetic analysis of radiation-induced changes in human gene expression. *Nature* 2009;459:587-591.

17. Andreassen CN, Barnett GC, Langendijk JA, et al. Conducting radiogenomic research: Do not forget careful consideration of the clinical data. *Radiother Oncol* 2012;105:337-340.

18. Wheeler HE, Maitland ML, Dolan ME, et al. Cancer pharmacogenomics: Strategies and challenges. *Nat Rev Genet* 2013;14:23-34.

19. Fachal L, Gomez-Caamaño A, Barnett G, et al. A three stage genome wide association study reveals susceptibility for late radiotherapy toxicity at the 2q24 *(TANC1)* locus. *Nat Genet* (in press)

20. Lambin P, Roelofs E, Reymen B, et al. Rapid Learning health care in oncology: An approach towards decision support systems enabling customised radiotherapy. *Radiother Oncol* 2013;109:159-164.

21. Kerns SL, Ruysscher DD, Andreassen CN, et al. STROGAR - STrengthening the Reporting Of Genetic Association studies in Radiogenomics. *Radiother Oncol* 2014;110:182-188.

22. Feldman-Stewart D, Brundage MD, Tong C. Information that affects patients' treatment choices for early stage prostate cancer: A review. *Can J Urol* 2011;18:5998-6006.

23. Feldman-Stewart D, Tong C, Siemens R, et al. The impact of explicit values clarification exercises in a patient decision aid emerges after the decision is actually made: Evidence from a randomized controlled trial. *Med Decis Making* 2012;32:616-626.