Special Review

Beyond Correlations, Sensitivities, and Specificities:

A Roadmap for Demonstrating Utility of Advanced Imaging in Oncology Treatment and Clinical Trial Design

Erich P. Huang, PhD, Frank I. Lin, MD, Lalitha K. Shankar, MD, PhD

ACRIN
American College of Radiology Imaging Network

CT
computed tomography

ER
estrogen receptor

FDG
[18F]-fluorodeoxyglucose

FDG-PET
positron emission tomography with [18F]-fluorodeoxyglucose

FES
[18F]-fluoroestradiol

FMISO
[18F]-fluoromisonidazole

FMISO-PET
positron emission tomography with [18F]-fluoromisonidazole

HNSCC
head and neck squamous cell carcinoma

MRI
magnetic resonance imaging

NPV
negative predictive value

OS
overall survival

Despite the widespread belief that advanced imaging should be very helpful in guiding oncology treatment decision and improving efficiency and success rates in treatment clinical trials, its acceptance has been slow. Part of this is likely attributable to gaps in study design and statistical methodology for these imaging studies. Also, results supporting the performance of the imaging in these roles have largely been insufficient to justify their use within the design of a clinical trial or in treatment decision making. Statistically significant correlations between the imaging results and clinical outcomes are often incorrectly taken as evidence of adequate performance. Assessments of whether the imaging can outperform standard techniques or meaningfully supplement them are also frequently neglected. This paper provides guidance on study designs and statistical analyses for evaluating the performance of advanced imaging in the various roles in treatment decision guidance and clinical trial conduct. Relevant methodology from the imaging literature is reviewed; gaps in the literature are addressed using related concepts from the more extensive genomic and in vitro biomarker literature.

Key Words: Clinical trial; cancer imaging; study design; statistical analysis.

Acad Radiol 2017; ■:■■–■■

From theBiometric Research Program, Division of Cancer Treatment, Diagnosis National Cancer Institute, NIH, 9609 Medical Center Drive, MSC 9735, Bethesda, MD 20892-9735 (E.P.H.); Cancer Imaging Program, Division of Cancer Treatment, Diagnosis National Cancer Institute, NIH, Bethesda, Maryland (F.I.L., L.K.S.). Received September 20, 2016; revised January 5, 2017; accepted March 2, 2017. **Address correspondence to:** E.P.H. e-mail: erich.huang@nih.gov

Published by Elsevier Inc. on behalf of The Association of University Radiologists.
http://dx.doi.org/10.1016/j.acra.2017.03.002

PET
positron emission tomography

PFS
progression-free survival

PPV
positive predictive value

RECIST
Response Evaluation Criteria in Solid Tumors

SUV
standardized uptake value

INTRODUCTION

Advanced imaging, namely, novel imaging methods or standard of care imaging used in novel clinical contexts, has proven valuable in guiding oncology treatment, clinical trial design, and drug development. Endpoints of many Phase II trials such as overall response rate (ORR) via Response Evaluation Criteria in Solid Tumors (RECIST) (1) were based on anatomic imaging. Changes in tumor metabolism in gastric cancer patients following the initial cycles of chemotherapy as measured by positron emission tomography with [18F]-fluorodeoxyglucose (FDG-PET) can serve as an early measure of response and thus help guide treatment adjustments in these patients after the first few cycles of chemotherapy (2–4). Possible roles of imaging are summarized in Table 1; definitions were borrowed from literature on clinical trial design and in vitro and genomic biomarkers to maintain a common vocabulary between these research communities.

Although the imaging literature currently contains many promising results, the incorporation of advanced imaging into patient care and clinical trial design has not been ubiquitous (5). In addition to the high cost of the investigation and development of imaging agents (6) and relatively high regulatory barriers for using imaging in a clinical study (eg, Investigational New Drug application requirements (7)), most of the results so far, although important, are insufficient evidence of the utility of advanced imaging in guiding disease management or as an integral part of the design of a treatment clinical trial. Many clinical imaging studies indicate statistically significant correlations and high sensitivities and specificities, which by themselves neither necessarily translate into adequate performance in a particular role nor sufficiently justify the extra effort and resources to adopt the imaging.

This paper serves as a roadmap for research to advance an imaging procedure to the point where it can justifiably be used in guiding disease management or as an integral part of the design of an oncology treatment clinical trial. In general, this entails evaluating the performance of the imaging through a sequence of progressively larger and more definitive clinical imaging studies. Gatsonis and Hillman and Gatsonis define a framework of such a sequence that draws upon parallels with Phase I, Phase II, and Phase III clinical trials of oncology treatments (8,9). The first studies, which Gatsonis and Hillman and Gatsonis call "Phase I", involve *discovery* and include those focusing on standardization of the image acquisition and processing protocol and evaluation of metrological aspects such as test-retest repeatability. Next come *introductory studies* evaluating the association between imaging measurements and clinical outcomes (eg, association between baseline FDG standardized uptake value [SUV] and overall survival [OS] in gastric cancer patients undergoing chemotherapy) or of the ability of the imaging to facilitate detection of clinically relevant characteristics such as metastases ("Phase II" according to Gatsonis and Hillman and Gatsonis). Then come larger *mature studies* directly and definitively evaluating the performance of the imaging in its intended role, typically in a multi-institutional setting ("Phase III"); for example, such a study of FDG-PET in the assessment of early response to chemotherapy in gastric cancer patients may involve a randomized study comparing the survival of early nonresponders (ie, those not showing appreciable decreases in FDG SUV after the initial cycles of chemotherapy) who switch treatments to that of early nonresponders who do not (10). So far, these mature studies have not received much attention in the imaging literature. Not only have comparatively few of them been performed to date, but also, the imaging literature contains substantial gaps in methodology regarding how to design and execute such studies. However, related concepts have been addressed more extensively in the genomic and in vitro biomarker literature and can be adapted to imaging to bridge these gaps.

The sequences of clinical studies needed to evaluate imaging in each of its possible roles are presented, with relevant designs and statistical analyses. The emphasis will be more on guiding the reader through this process rather than the statistical details of the analysis methodology or designs of such studies. Presenting novel study designs or statistical methodology is not the purpose of this paper either. Study designs and analysis techniques from the imaging literature are reviewed whenever such methodology is available. Otherwise, analogous