QUALITY AND SAFETY IN RADIATION ONCOLOGY

IMPLEMENTING TOOLS AND BEST PRACTICES FOR PATIENTS, PROVIDERS, AND PAYERS

ADAM P. DICKER
TIM R. WILLIAMS
ERIC C. FORD
Quality and Safety in Radiation Oncology
Quality and Safety in Radiation Oncology

Implementing Tools and Best Practices for Patients, Providers, and Payers

Editors

Adam P. Dicker, MD, PhD
Chair and Professor of Radiation Oncology
Pharmacology and Experimental Therapeutics
Sidney Kimmel Medical College at Thomas Jefferson University
Sidney Kimmel Cancer Center
Philadelphia, Pennsylvania

Tim R. Williams, MD
Medical Director, Radiation Oncology
The Lynn Cancer Institute
Radiation Oncology Department
Boca Raton Regional Hospital
Boca Raton, Florida

Eric C. Ford, PhD
Professor
Department of Radiation Oncology
University of Washington Medical Center
Seattle, Washington

© Demos Medical Publishing
Contents

Contributors ix
Preface xiii
Acknowledgments xv

Part I. Basic Concepts in Quality, Value, and Safety
Tim R. Williams

1. Introduction to Concepts in Quality, Value, and Safety for Health Care Systems 3
   Tim R. Williams and Alex Bastian

2. Defining Value, American Style 11
   Alex Bastian, Yohan Cho, and Tim R. Williams

3. Defining Value in European Health Care Systems 25
   Alex Bastian, Yohan Cho, and Tim R. Williams

4. Systems Engineering and Process Control 35
   Ajay Kapur and Louis Potters

5. Error Prevention and Risk Management 47
   Suzanne B. Evans and Derek Brown

   Ronald C. Chen, James R. Broughman, and Jeff M. Michalski

Part II. Quality in Radiation Physics
Eric C. Ford

7. Equipment and Software: Commissioning From the Quality and Safety Perspective 73
   Indra J. Das and Aaron Andersen

   Baozhou Sun and Eric E. Klein

9. Quality Considerations in External Beam Radiotherapy Simulation and Treatment Planning 99
   Eric C. Ford

10. Quality Considerations in Brachytherapy 107
    R. Adam B. Bayliss and Bruce R. Thomadsen

11. Quality Considerations in Proton and Particle Therapy 125
    Brian Winey, Benjamin Clasie, and Yuting Lin

12. Quality Assurance and Verification of Treatment 135
    Bernard L. Jones and Moyed Miften

© Demos Medical Publishing
13. Plan Quality: The Good, the Bad, and the Ugly 143
   Kevin L. Moore

14. Special Procedures and Equipment 155
   H. Harold Li, Michael B. Altman, and H. Omar Wooten

15. Individual Patient-Based Peer Review for Photon and Proton Therapy 163
   Genevieve M. Maquilan, Henry K. Tsai, and Hiral P. Fontanilla

   Mary Coffey and Ola Holmberg

Part III. Quality in the Clinical Practice Setting
   Adam P. Dicker

17. Creating a Culture of Safety and Empowerment 179
   Dennis R. Delisle and Nicholas A. Thompson

18. Clinical Testing of Quality and Safety Measures 189
   Kevin L. Moore and Arno J. Mundt

19. Quality and Safety Training in Graduate Medical Education 199
   Erin F. Gillespie, Derek Brown, and Arno J. Mundt

20. Training: How to Teach Quality and Safety 205
   Laura A. Doyle, Amy S. Harrison, and Yan Yu

21. Making Metrics Matter 211
   Amy S. Harrison, Laura A. Doyle, and Yan Yu

22. The Patient Perspective on Quality and the Role of the Patient in Safety 221
   Suzanne B. Evans and Sharon Rogers

23. The Impact of Incidents: Reporting, Investigating, and Improving 229
   Jennifer L. Johnson and Suzanne B. Evans

24. Medical Error Disclosure 239
   John Banja and Suzanne B. Evans

25. Error Identification and Analysis 249
   Yan Yu, Laura A. Doyle, and Amy S. Harrison

26. Introduction and History of Patient Safety Organizations:
    The Formation of a National Safety Program 253
   Tom Piotrowski

27. Quality Assurance and Safety Requirements for Clinical Trials 261
    Ying Xiao, Stephen F. Kry, and David S. Followill

28. Incident Learning Systems 267
    Gary Ezzell, Stephanie Terezakis, and Courtney Buckey

29. Quality and the Role of the Dosimetrist and Therapist in Radiation Oncology 275
    Joumana Dekmak and Kathy Lash

30. Quality and the Electronic Medical Record in Radiation Oncology 287
    James Mechalakos and Sonja Dieterich

31. Quality and the Third-Party Payer 297
    Heather A. Curry
Appendix: Quality Assurance and Patient Safety Checklists  309
  Checklist for Peer Review Paper  309
   External Beam Plan Checklist  310
   Plan Checks and Pretreatment Physics Quality Assurance Checks  312
   Checklist for Stereotactic Body Radiation Therapy Treatment  313
   Time Out Checklist  314
   Checklist of a Gamma Knife Perfexion Daily Quality Assurance Example  316
   Checklist for Treatment Day Quality Assurance of Perfexion  317
   Checklist for Gamma Knife Perfexion Pretreatment  318
   Checklist for Treatment Delivery Quality Assurance of Perfexion Patients  319
   Checklist for Gammamed HDR Treatment  320
   Checklist for Commissioning  322

Index  323
Contributors

Michael B. Altman, PhD
Assistant Professor
Department of Radiation Oncology
Washington University School of Medicine
St. Louis, Missouri

Aaron Andersen, MS
Physicist
Department of Radiation Oncology
Indiana University School of Medicine
Indianapolis, Indiana

John Banja, PhD
Professor
Center for Ethics
Emory University
Atlanta, Georgia

Alex Bastian, MBA
Vice President
Health, Market Access
GfK Custom Research
San Francisco, California

R. Adam B. Bayliss, PhD
Assistant Professor
Department of Human Oncology
University of Wisconsin
Madison, Wisconsin

James R. Broughman, BS
Medical Student
University of North Carolina School of Medicine
Chapel Hill, North Carolina

Derek Brown, PhD
Associate Professor
Department of Radiation Medicine
and Applied Sciences
University of California, San Diego
La Jolla, California

Courtney Buckey, PhD
Assistant Professor
Department of Radiation Oncology
Mayo Clinic Arizona
Phoenix, Arizona

Ronald C. Chen, MD, MPH
Associate Professor
Department of Radiation Oncology
University of North Carolina School of Medicine
Chapel Hill, North Carolina

Yohan Cho
Senior Consultant
Market Access
GfK Custom Research
New York, New York

Benjamin Clasie, PhD
Assistant Professor
Department of Radiation Oncology
Francis H. Burr Proton Therapy Center
Massachusetts General Hospital
Harvard Medical School
Boston, Massachusetts

Mary Coffey
Adjunct Associate Professor
Department of Radiation Oncology
Trinity College Dublin School of Medicine
Dublin, Ireland

Heather A. Curry, MD
Director, Radiation Oncology
NantHealth
Philadelphia, Pennsylvania

Indra J. Das, PhD, FACR, FASTRO
Vice Chair, Professor and Director of Medical Physics
Department of Radiation Oncology
New York University Langone Medical Center
Laura and Isaac Perlmutter Cancer Center
New York, New York

© Demos Medical Publishing
CONTRIBUTORS

Joumana Dekmak, MS, BS, RT (T)
Radiation Therapist Supervisor
Department of Radiation Oncology
University of Michigan Health System
Ann Arbor, Michigan

Dennis R. Delisle, ScD, MHSA, FACHE
Executive Project Director
Thomas Jefferson University and Jefferson Health
Philadelphia, Pennsylvania

Sonja Dieterich, PhD, FAAPM
Associate Professor
Department of Radiation Oncology
University of California, Davis
Sacramento, California

Laura A. Doyle, MS
Clinical Instructor
Department of Radiation Oncology
Sidney Kimmel Cancer Center
Thomas Jefferson University
Philadelphia, Pennsylvania

Suzanne B. Evans, MD, MPH
Associate Professor
Department of Therapeutic Radiology
Yale University School of Medicine
New Haven, Connecticut

Gary Ezzell, PhD
Chair, Medical Physics Division
Associate Professor of Radiation Oncology
Department of Radiation Oncology
Mayo Clinic Arizona
Phoenix, Arizona

David S. Followill, PhD
Imaging and Radiation Oncology Core (IROC)
Houston/MD Anderson Cancer Center
Houston, Texas

Hiral P. Fontanilla, MD
Department of Radiation Oncology
Princeton Radiation Oncology
Monroe, New Jersey

Eric C. Ford, PhD
Professor
Department of Radiation Oncology
University of Washington Medical Center
Seattle, Washington

Erin F. Gillespie, MD
Resident Physician
Department of Radiation Medicine and Applied Sciences
University of California, San Diego
La Jolla, California

Amy S. Harrison, MS
Medical Physicist, Operations and Safety Director
Department of Radiation Oncology
Sidney Kimmel Cancer Center
Thomas Jefferson University
Philadelphia, Pennsylvania

Ola Holmberg, PhD
Unit Head, Radiation Protection of Patients
Radiation Safety and Monitoring Section
International Atomic Energy Agency
Vienna, Austria

Jennifer L. Johnson, MS, MBA
Senior Medical Physicist and Director for Safety
Department of Radiation Physics
The University of Texas MD Anderson Cancer Center
Houston, Texas

Bernard L. Jones, PhD
Assistant Professor
Department of Radiation Oncology
University of Colorado School of Medicine
Aurora, Colorado

Ajay Kapur, PhD
Director of Medical Physics Research and Education
Department of Radiation Medicine
Hofstra Northwell School of Medicine
Hempstead, New York

Eric E. Klein, PhD
Vice President / Director of Medical Physics
Professor of Radiation Medicine
Northwell Health System
Lake Success, New York

Stephen F. Kry, PhD
Imaging and Radiation Oncology Core (IROC)
Houston/MD Anderson Cancer Center
Houston, Texas

Kathy Lash, BS, RT (R) (T)
Director of Community Practices
Department of Radiation Oncology
University of Michigan Health System
Ann Arbor, Michigan

© Demos Medical Publishing
CONTRIBUTORS

H. Harold Li, PhD
Associate Professor
Department of Radiation Oncology
Washington University School of Medicine
St. Louis, Missouri

Yuting Lin, PhD
Fellow
Department of Radiation Oncology
Massachusetts General Hospital
Harvard Medical School
Boston, Massachusetts

Genevieve M. Maquilan, MD
Resident Physician
Department of Radiation Oncology
University of Texas Southwestern Medical Center
Dallas, Texas

James Mechalakos, PhD
Associate Attending Physicist
Department of Medical Physics
Memorial Sloan Kettering Cancer Center
New York, New York

Jeff M. Michalski, MD, MBA
Professor
Department of Radiation Oncology
Washington University School of Medicine
St. Louis, Missouri

Moyed Miften, PhD
Professor and Chief of Physics
Department of Radiation Oncology
University of Colorado School of Medicine
Aurora, Colorado

Kevin L. Moore, PhD, DABR
Associate Professor / Associate Division Director
Department of Radiation Medicine and Applied Sciences
University of California, San Diego
La Jolla, California

Arno J. Mundt, MD
Professor and Chair
Department of Radiation Medicine and Applied Sciences
University of California, San Diego
La Jolla, California

Tom Piotrowski, RN, MSN
Vice President of Clinical Informatics / Executive Director
Clarity PSO
Clarity Group, Inc.
Chicago, Illinois

Louis Potters, MD, FACR, FASTRO
Chairman and Professor
Department of Radiation Medicine
Hofstra Northwell School of Medicine
Hempstead, New York

Sharon Rogers, MHSc, BA
Hospital Ombudsman
University Health Network
Toronto, Ontario, Canada

Baozhou Sun, PhD
Instructor
Department of Radiation Oncology
Washington University School of Medicine
St. Louis, Missouri

Stephanie Terezakis, MD
Associate Professor
Department of Radiation Oncology and Molecular Radiation Sciences
Johns Hopkins University School of Medicine
Baltimore, Maryland

Bruce R. Thomadsen, PhD
Professor
Department of Medical Physics
University of Wisconsin School of Medicine and Public Health;
Director
Patient Safety Organization
The Center for the Assessment of Radiological Sciences
Madison, Wisconsin

Nicholas A. Thompson, MPH
Communication and Change Enablement Analyst
Thomas Jefferson University and Jefferson Health
Philadelphia, Pennsylvania

Henry K. Tsai, MD
Princeton Radiation Oncology
Monroe, New Jersey

Tim R. Williams, MD
Medical Director, Radiation Oncology
The Lynn Cancer Institute
Radiation Oncology Department
Boca Raton Regional Hospital
Boca Raton, Florida

© Demos Medical Publishing
Ying Xiao, PhD
Professor
Department of Radiation Oncology
Perelman School of Medicine
University of Pennsylvania
Philadelphia, Pennsylvania

Yan Yu, PhD, MBA, FAAPM, FASTRO
Vice Chair, Professor and Director of Medical Physics
Department of Radiation Oncology
Sidney Kimmel Cancer Center
Thomas Jefferson University
Philadelphia, Pennsylvania

Brian Winey, PhD
Assistant Professor
Department of Radiation Oncology
Francis H. Burr Proton Therapy Center
Massachusetts General Hospital
Harvard Medical School
Boston, Massachusetts

H. Omar Wooten, PhD
Assistant Professor
Department of Radiation Oncology
Washington University School of Medicine
St. Louis, Missouri
Preface

This book started as two independent conversations by Tim Williams and me with Rich Winters, formerly the Executive Editor at Demos over 3 years ago. To give context, in 2010, the New York Times highlighted medical errors that occurred in the field of radiation oncology. ASTRO created the Target Safely campaign at that time and in June 2011, ASTRO’s Board of Directors approved a proposal to establish a national radiation oncology-specific incident learning system. A standardized system provides an opportunity for shared learning across all radiation oncology institutions and may be an added value to institutions that track incidents independently. This incident learning system represents a key commitment of Target Safely; ASTRO’s patient protection plan, designed to improve the safety and quality of radiation oncology. ASTRO partnered with the American Association of Physicists in Medicine (AAPM) to develop RO-ILS: Radiation Oncology Incident Learning System®, the only medical specialty society-sponsored incident learning system for radiation oncology. In part, thanks to Dr. Lawrence Marks, University of North Carolina, I became a member of the Radiation Oncology Healthcare Advisory Council (RO-HAC) for the RO-ILS program. While immersing myself in details of RO-ILS I had the great pleasure working with Dr. Eric Ford, discussing quality and safety initiatives at the institutional and national level.

Given the increasing importance and high profile nature of Quality and Safety for the field of Radiation Oncology, at the Annual ASTRO meeting in 2013 I had discussed the idea of a book on the topic of Quality and Safety with Rich Winters. Rich mentioned that he had a similar conversation with Dr. Tim Williams of the Eugene M. and Christine E. Lynn Cancer Institute at Boca Raton Regional Hospital who was also interested in the subject. Tim was the ASTRO Board Chair in 2010 when the Target Safely initiative was launched. I met with Tim and we discussed our ideas and we both realized that his perspective was different from mine, yet together they would synergize quite well. I suggested we needed a medical physicist who has extensive experience in quality and safety and brought up Dr. Eric Ford, University of Washington, as a partner in the collaborative effort.

This project has been great for bringing together a diverse group of people whose training and backgrounds complement each other, all for the goal of helping patients. The book views Quality and Safety in the context of “Value” and the Institute of Medicine Quality Initiative.

Radiation Oncology is a technically demanding and complex field. It requires a team of individuals working together to deliver high-quality care. We believe by sharing information, as this book does, we can improve the lives of patients worldwide.

Adam P. Dicker, MD, PhD
Acknowledgments

Carolyn, Michal, Josh, Shimshon and Yehuda. Thank you, love you, you’re the best!!!!!!

I would like to thank the leadership at Thomas Jefferson University who helped us combine sophisticated patient-first, multidisciplinary care with high-impact science for cancer patients.

—Adam P. Dicker

I would like to extend my sincere gratitude to the Chapter authors, who devoted much time and effort to deliver an unbelievable work product, our editors at Demos for their guidance and support, and most of all to Adam and Eric, whose patience and understanding with me is the true unsung story of this book.

—Tim R. Williams

I would like to thank my many partners in delivering quality care and understanding what it means, Gabrielle Kane, Jing Zeng, Matt Nyflot and the many active team members at University of Washington including Lulu Jordan, Patty Sponseller, Josh Carlson, and many others. I also thank my mentors, colleagues, and friends at Johns Hopkins, Ted DeWeese, Stephanie Terezakis, and Peter Pronovost.

—Eric C. Ford
INTRODUCTION
This chapter discusses the nature of error, the classification of error, concepts in quality management, types of error mitigation, hierarchy of interventions, and the role of culture in error prevention.

PSYCHOLOGY OF ERROR: UNDERSTANDING OUR INHERENT BIASES
An exploration of the psychology of error is useful, as it provides context and insight that can help us understand the causes of error and develop functional error mitigation strategies.

The way we look at the world is often not rational. Our impressions are largely guided by heuristics, or “rules of thumb,” as to how the world usually works. Psychologists have proposed a dual system processing theory for thought (1,2). There are many different modifications and permutations of this theory, but they all center on a similar basic concept: the belief that there is one mode of thinking in which a person is carefully considering the information in front of him or her, and making a rational judgment based on the information at hand. In the other mode of thinking, one is not thinking deliberately, and mental shortcuts are used to make snap judgments about a situation. These modes have been called systems 1 and 2 (2), central and peripheral (1), or slow and fast thinking (3), respectively. Clearly, the process of “fast thinking” leaves the individual much more prone to error; however, it is argued that this is the predominant mode of how we process interactions in our daily lives.

While we cannot remove ourselves from the tendencies and biases that we are subject to, it is helpful for all members in the department of radiation therapy to understand that these sorts of cognitive processes are not aberrant, but rather a natural part of human interactions. Therefore, the strategies for error mitigation include awareness of biases we are subject to, as well as avoiding error mitigation strategies that include the elimination of these biases, as this is impossible. An appreciation for the ubiquity of bias in medicine will aid in the understanding of how errors can occur. Consider, if you will, the following biases, along with examples of how they might be evident in radiation oncology.

Confirmation bias
This is the tendency to believe, seek out, and interpret information that supports one’s own preconceptions (4). This is similar to myside bias (5). For instance, a clinical example of this is a patient who arrives with a diagnosis of metastatic breast cancer and known bony metastases. She reports early satiety, intermittent fever, weight loss, and night sweats. Imaging reveals a large amount of axillary, supraclavicular, inguinal, and also periportal adenopathy. Her endocrine therapy is adjusted, and for 2 months there is progression. A biopsy is performed, and later diffuse large B-cell lymphoma is diagnosed. Another example is a physicist trying to troubleshoot a pretreatment quality assurance (QA) of an intensity-modulated radiation therapy (IMRT) plan that is off by more than 5%. The physicist may attribute this discrepancy to poor QA equipment setup, incorrect normalization of the verification plan, or poor choice of measurement location, when, in fact, there is a mechanical problem with the linac and the plan is not being delivered correctly.

Availability bias
This is the tendency to overestimate the occurrence of events that are known to the individual. Events may be more “available” depending on how unusual or emotionally charged they are (6). An example of this in radiation oncology is a treatement team that traumatically lost a patient to aortic dissection on the treatment table after the patient complained of chest pain and was hypotensive. The next patient who comes in with chest pain and hypotension is sent to the emergency department for a workup rather than treated for his radiation esophagitis and dehydration. Additionally, one can witness admitting inpatient teams “running specials” on a given disease...
process, as they are admitting multiple people with the same “available” diagnosis.

Ambiguity effect

The tendency to avoid choices for which information is missing makes the probability of a given effect seem “unknown” (7). For instance, a radiation oncology team is faced with a choice of two boards for prone positioning. After a detailed physics evaluation showing the two to be equivalent, the team decides to go with the more expensive, better-known company rather than the cost-effective start-up, whose service record is “unknown.”

Sunk cost fallacy (also known as irrational escalation)

This occurs when individuals continue with a poor decision, based on a perceived prior amount of work already invested in that decision (8). A clinical example might be a radiation oncologist and dosimetrist who have spent many hours planning a treatment for IMRT to the head and neck, whose plan has already undergone quality assurance by the physicist. At chart rounds, peer review suggests that an additional lymph node basin should be included. The radiation oncologist decides not to include this nodal basin based on sunk cost fallacy.

Inattentional blindness

This is the phenomenon of items in plain sight remaining invisible to the viewers because they are not looking for these items or expecting to find them (9). An example of this is a very large thyroid mass being overlooked on a CT simulation for breast cancer, as the radiation oncologist is focused on definition of the lumpectomy cavity, the field design, and the heart and lung contours. Another example of this is a physicist checking a treatment plan, paying very close attention to doses to organs at risk and planning target volume coverage, but missing a 150% hotspot in unsegmented normal tissue.

Outcome bias

This is the tendency to judge a decision or action by its eventual outcome, rather than the quality of the decision at the time it was made (10). An example in radiation oncology might be regretting a decision to install an orthovoltage machine in a clinic seeing a tremendous amount of skin cancer, once reimbursement for orthovoltage treatments plummets.

Intervention bias

This describes the tendency of the treatment team (and sometimes the patient or family) to prefer intervention, whether it is with drugs, diagnostic tests, noninvasive procedures, or surgeries, when supportive care would be a reasonable alternative (11). A common example in radiation therapy is the preference of the medical team (and often the patient) to perform palliative radiation treatment on a patient with a massive burden of disease when it is unlikely to be of benefit, and the patient is clearly at the end of life. An additional example is an incident of minor severity and low occurrence frequency prompting institution of a lengthy therapy checklist to prevent its reoccurrence, out of an intrinsic preference for action.

Counterfactual thinking

This is the tendency of individuals to create alternative, imaginative outcomes after the actual outcome is known but has an undesirable result (12,13). An example in radiation oncology is a decision to treat an apical lung tumor at the high end or low end of acceptable doses. Out of concern for the brachial plexus and lung doses, the radiation oncologist elects to treat at the lower end of the acceptable dose. The patient subsequently suffers a local recurrence. The radiation oncologist spends significant time imagining a choice for a higher dose, imagining cure rather than serious radiation pneumonitis or brachial plexopathy.

Cognitive dissonance

This refers to the mental anxiety experienced by an individual who holds several contradictory beliefs or values at the same time, or is faced with new information that contradicts existing beliefs or values (14,15). An example in radiation oncology might be the treatment team who consider themselves exceptionally careful, yet make a catastrophic error in treatment. It is this sort of affront to one’s positive self-concept that makes errors so challenging to recognize and admit.

Situational awareness

This refers to the ability of individuals to process what is happening around them in order to inform decision making (16). Loss of situational awareness may also be colloquially termed “losing the forest for the trees.” Consider the tragic situation of Elaine Bromiley, a 37-year-old mother of two, who was admitted for an elective endoscopic sinus surgery and septoplasty (17). She suffered hypoxic brain damage, ultimately fatal, during the induction of general anesthesia, largely because of loss of situational awareness by the physician team. They were unable to intubate or ventilate her—a situation that calls for emergency tracheostomy—but they were caught up in the tasks of repeated attempts to intubate or ventilate her, ignoring the passage of time with profound hypoxia.

As evidenced by the preceding examples of bias, it is clear that the process of information processing and
cognitive reasoning is quite complex, and subject to many biases even under the best conditions. But, as the later sections explore, the best conditions are not always present, and error can result.

CLASSIFICATIONS OF ERROR: HUMAN FACTORS ANALYSIS AND CLASSIFICATION SYSTEM

The accurate and timely classification of error forms the backbone of retrospective quality and safety improvement efforts and is the basis for prospective error mitigation strategies. Error classification also enables aggregate trending over time and comparisons across departments and institutions—both valuable tools in the effort to improve quality and safety.

Historically, when an incident occurs in the health care environment, teams of experts use root cause analysis (RCA) to classify the incident and to determine the underlying, root causes. More recently, a classification system based on Reason’s error categorization, Human Factors Analysis and Classification System (HFACS), has been demonstrated to improve standardization of error classification, increase the specificity of error classification, and facilitate the development of actionable corrective plans (18).

Before we discuss HFACS in more detail, it is useful to describe several other error categorization schemes that can be used in addition to those presented in HFACS. Errors can be classified as sporadic, where the same incident is unlikely to occur again in the same process step or under the same conditions; or systematic, where, given the same conditions, the same error is likely to occur again. Errors can also be classified as “active” and “latent” (19). Active errors occur at the point of contact between a human and some aspect of a larger system. Latent errors are less apparent failures of organization or design that contributed to the occurrence of errors. As we examine HFACS, we will attempt to identify error categories as either systematic/sporadic or active/latent.

The HFACS methodology, originally designed for use within the U.S. Navy and Marine Corps to identify common root causes among aviation-related accidents, looks at an incident from the perspective of Reason’s four

![Diagram of HFACS framework](https://via.placeholder.com/150)


© Demos Medical Publishing
levels of error causation: (a) Unsafe Act, (b) Preconditions for Unsafe Acts, (c) supervision, and (d) organizational influences (20,21). Each level contains several causal categories that act as bins for “nanocodes,” which are similar to the root causes used in a RCA. Figure 5.1 shows the HFACS framework.

Because errors and violations are associated with human performance, it is worthwhile considering the performance categories that individuals are working in when the error occurs. Rasmussen defined three human performance categories: (a) skill-based, (b) rule-based, and (c) knowledge-based (22). Skill-based activities are typically straightforward routine tasks that have been performed for some time. Rule-based activities are more complex or critical tasks that may be only occasionally performed, and error results from improper application of a rule or a plan. Knowledge-based activities deal with unfamiliar and/or unprepared-for tasks, and the intended outcome may not be achieved because of a knowledge deficit. Unsafe acts are therefore categorized into errors (skill-based, perceptual, decision) and violations. Errors represent normal accepted behavior that fails to meet the desired outcome, whereas violations are the result of intentional departures from accepted practice. A decision error occurs when information, knowledge, or experience is lacking, whereas skill-based errors occur because of inattention or memory failure. A perceptual error occurs when sensory input is somehow unintelligible or unusual. Because unsafe acts occur while the individual is performing the task, these are typically seen as active errors, and depending on whether the error is skill-based, decision, or perception, unsafe acts can be categorized as either sporadic or systematic.

Preconditions for Unsafe Acts refer to causal factors, such as the work environment and conditions that existed at the time of the incident. These include environmental factors, personnel factors, and the conditions of the operator. Environmental factors include both the physical environment and the technological environment. Personnel factors relate to how people work together and communicate with one another. The conditions of the operator describe situations in which the person performing the task is either distracted or incapable of performing the task. Preconditions for Unsafe Acts may be either active or latent, depending on whether the causal factor related to the condition of the operator or environmental factors, and are more likely to be categorized as systematic when similar conditions would likely result in the same error.

The supervision category outlines factors that relate to leadership and administration and how these may have played a role in the incident. There are four subtypes in this category: leadership (provision of inadequate training, guidance, or oversight), operational planning (scheduling and the assignment of work), failure to correct known problems, and supervisory ethics (permitting individuals to perform work outside of their scope). Supervisory factors are latent errors because these factors are typically “once removed” from the incident, and are likely to be systematic because similar conditions would likely result in the same error.

Organizational influences describe factors such as resource management and organizational culture that can directly affect supervisors and also contribute to incidents. The three subtypes in this category are resource management (allocation of human resources and equipment budgets), organizational climate (culture), and operational processes (high-level policies and procedures). Organizational influences, like supervision factors, are considered latent, systematic errors.

For further consideration, in an alternative but similar methodology, the Health Care Performance Initiative has further classified failure modes into individual or system failures. It divides individual failure modes into five categories: those related to competency, consciousness, communication, critical thinking, or compliance (23). Table 5.1 lists these in further detail. It divides system failure modes into five categories as well: those related to structure, culture, process, policy and protocol, or technology and environment (23). Table 5.2 lists these in further detail.

CONCEPTS IN QUALITY MANAGEMENT

Much like Socrates, to whom is attributed the quote that “The unexamined life is not worth living,” one might say that failure to examine one’s department for quality measures decreases the possibility that a quality department might exist. Without understanding one’s event rates, one cannot understand whether small improvements or major overhauls are required. Several theories about the ideal strategy for this analysis, which will be explored now.

Dr. W. Edwards Deming (24) initially spoke about this process in regard to the Shewhart cycle (25): design the product, make the product, test the product through market research, and then redesign the product. From this came the Deming cycle, which espouses the principles “plan, do, study, act” (PDSA). The planning phase encompasses identification of areas for improvement, and identification of changes required. The “do” phase is for implementation. The study phase is for measurement of the success of the process or the outcome measure. The “act” phase refers to the need to make further changes based on those results. The intent with this system is that it is viewed as a cycle that is never completed, leading to continuous quality improvement, or continuous process improvement. This principle is applied to the optimization of medical processes for the purposes of increasing quality of care and increasing patient safety. It is important to note that the “check” or “study” piece of this process means that routine audits will have to be performed in order to measure progress or identify interventions that were less effective than hoped.
### Table 5.1 Taxonomy of Individual Failure Modes

<table>
<thead>
<tr>
<th>Competency (Knowledge and Skills)</th>
<th>Consciousness (Attention)</th>
<th>Communication (and Information Processing)</th>
<th>Critical Thinking (Cognition)</th>
<th>Compliance (Motivation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unformed skills/habits (inability to do something well; while possessing knowledge, lacks performance reliability gained through experience)</td>
<td>Inattention (preoccupied; rushing or hurrying; not paying attention)</td>
<td>Incorrect assumption (assuming a thing to be true or correct that was in fact wrong)</td>
<td>Situational awareness (unawareness or lacking knowledge of what is going on; failure to perceive that acts or conditions deviated from desired path)</td>
<td>Indifference (inadequate care or attention to people or things of responsibility; carelessness, informality, or casual attitude, yet with no deliberate intention to cause harm)</td>
</tr>
<tr>
<td>Normalized deviance (conforming to an individual’s standard, type, or custom, where behavior is sharply different from the generally accepted standard)</td>
<td>Distraction (divided or diverted attention)</td>
<td>Misinterpretation (forming an understanding that is not correct from something that is said or done)</td>
<td>Failure to validate/verify (failure to find or test the truth of something; failure in the cognitive process of establishing a valid proof)</td>
<td>Shortcut (deliberate, conscious act to take a quicker or more direct route—a route that deviates from the designated or optimal path)</td>
</tr>
<tr>
<td>Inadequate knowledge (lacks fundamental knowledge-in-the-head of operating procedures or principles, or knowledge of available protocols)</td>
<td>Habit intrusion or reflex (act performed without conscious thought; a settled or regular tendency or practice)</td>
<td>Information overload (overburdened with too much sensory or cognitive input or information)</td>
<td>Mindset (primed or biased by pattern or preconceived notion; a fixed mental attitude or disposition that pre-determines a person’s response to interpretations or situations)</td>
<td>Overconfident (excessively confident or presumptuous; failure to stop when questions arise; proceeding in the face of uncertainty)</td>
</tr>
<tr>
<td></td>
<td>Spatial disorientation (feeling lost or confused, especially with respect to direction or position; confused because of misleading information)</td>
<td></td>
<td>Tunnel vision (the tendency to focus exclusively on a single or limited objective or view; overly focused on details of task; failure to see the big picture)</td>
<td>Reckless (acting without thought or care for the consequences of one’s acts; acting overtly with full knowledge that an act could cause harm)</td>
</tr>
<tr>
<td></td>
<td>Bored, fatigued, or unfit for duty (feeling weary; extreme tiredness because one is unoccupied; has no interest because of physical or mental activity or external influence)</td>
<td>Lapse (a momentary fault or failure in behavior; inadequate mental tracking; inadvertently forgot to complete something)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As part of this attitude toward continuous improvement, the department is encouraged to adopt the practice of systems thinking (26). This refers to an analytic focus on the systems and complex interplays involved in the delivery of health care, rather than on individual error. This thinking paradigm seeks the systems approach solution to what is nearly always a multifactorial problem. This can be quite challenging, as there has always been an emphasis in health care on individual training and individual responsibility.

However, the practice of blame and shame is counter to patient safety, as it inhibits future incident reporting and damages safety culture.

When looking at a quality management system, it is also useful to consider the International Organization for Standardization (ISO) 9000 and Six Sigma approaches. The concept of sigma is a metric that indicates how well a given process is performing (27,28). If the value of sigma is high, then the performance quality of the process is also high. Essentially, sigma measures the ability of a process to result in defect-free work, with patient dissatisfaction considered a defect. “Six Sigma” refers to the concept that product defects (or error) would occur at a rate of six standard deviations away from being a matter of chance (29). Part of this process is the cycle “design, measure, analyze, improve, control” (DMAIC), which is similar to PDSA; however, it contains the additional step of control. Control refers to sustaining gains made through quality improvement. Part of the Six Sigma process is the generation of fishbone diagrams, in which the causes of a certain event are identified for later remediation. ISO 9000 is another quality management approach that is based on the principles of customer focus, leadership, involvement of people, process approach, systems approach to management, continual improvement, factual

---

**Table 5.2 Taxonomy of System Failure Modes**

<table>
<thead>
<tr>
<th>Structure</th>
<th>Culture</th>
<th>Process</th>
<th>Policy and Protocol</th>
<th>Technology and Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure model (wrong model, incompatible missions)</td>
<td>Inadequate vision or mission (lacking or poorly executed mission)</td>
<td>Omitted actions (key activity is missing or incomplete)</td>
<td>Lacking or informal (no policy or protocol)</td>
<td>Input/output (visual display, alarms, control configuration)</td>
</tr>
<tr>
<td>Inadequate structure (span of control, levels of leadership, leveraging positions and experience)</td>
<td>Noncollaboration (disruptive competition, defensiveness, poor teamwork, low morale)</td>
<td>Excessive actions (contains low-value activities)</td>
<td>Usability (poor presentation or information depiction, low credibility, poor access)</td>
<td>Human capability (symbols, codes, anthropometry, devices, human control, physical work)</td>
</tr>
<tr>
<td>Inadequate job function (overlap or gaps in roles, responsibilities, or expectations)</td>
<td>Operational leadership (lacking or inadequate command, prioritization, or assignment in response to emergent or emerging issues)</td>
<td>Poorly sequenced (poor flow, excessive branching of work process activities)</td>
<td>Understandability (difficult to comprehend because guidance detail is lacking or inadequate for the knowledge and skill level of the user)</td>
<td>Arrangement (physical arrangement of work space, department, facility, or campus negatively impacting performance)</td>
</tr>
<tr>
<td>Resource allocation (insufficient infrastructure, people, budget, equipment, or other resources)</td>
<td>High-reliability environment (setting does not incorporate error prevention expectations and focus including competency, consciousness, critical thinking, communication, compliance)</td>
<td>Inadequate interface (lack of or poorly designed handoffs of information, resources, or products)</td>
<td>Knowledge in environment (inadequate or underutilized job aids, forcing functions)</td>
<td>Environment (lighting, noise, climate, motion negatively impacting performance)</td>
</tr>
<tr>
<td>Collaboration mechanisms (wrong or inadequate collaboration mechanisms)</td>
<td></td>
<td>Inadequate checks (lack of or poorly designed checks, inspections, or reviews)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

approach to decision making, and mutually beneficial supplier relationships (30). Criticisms of both ISO 9000 and Six Sigma (31) include the large time and financial investments required for their implementation.

The final approach mentioned here for quality management is the practice of peer review. There is no question that the practice within radiation oncology of reviewing cases on a routine basis with one’s peers is somewhat exceptional in medicine. This is partially because of the tangible work product in radiation oncology—the treatment plan, which lends itself to review, unlike the surgical technique or diagnostic acumen, which is much more challenging to review for each and every patient treated, and certainly is less easy to adjust midstream. Peer review is an integral part of QA (32,33), and has been shown to result in changes in management in a nontrivial percentage of patients (34,35). Readers are directed to the guidance from the American Society for Radiation Oncology (ASTRO) on this subject for a detailed treatise on the components of peer review (36).

TYPES OF ERROR MITIGATION—PROACTIVE MITIGATION STRATEGIES AND RETROSPECTIVE CORRECTIVE ACTIONS

Whether in response to an incident or as part of a quality improvement project, the development of effective error mitigation strategies is unquestionably important. This is where the “rubber meets the road.” You have done your prospective analysis, your RCA, or, better yet, your HFACS. You have figured out where the problems are and all the different pathways and factors that contributed, or could have contributed, to an incident. Now it is time to design, or redesign, a process, modifying or including new steps that will reduce the likelihood that an incident will occur.

Where do you turn for ideas? How would you know whether the newly implemented strategy will be effective? What about communicating the plan to everybody in your department? These are important questions, and if they are not adequately addressed, your error mitigation strategy is more likely to fail. In what follows we offer a brief, prescriptive approach to the development and implementation of error mitigation strategies.

Once you know where the problems are, the first step is to brainstorm potentially useful error mitigation strategies. The Hierarchy of Interventions can be useful in providing guidance on what broad type of error mitigation strategy might be appropriate (37). Another option is to see what others have done and then adapt this to your specific clinical situation. Sutlief (38) has performed an extensive review of 11 sources that describe more than 1,000 incidents and corrective actions. The relative distribution of corrective actions is shown in Figure 5.2. The safety barrier category includes corrective actions such as additional QA steps and checklists/time-outs.

---

**FIGURE 5.2 Relative frequency of corrective actions used—analysis of more than 1,000 reported incidents and corrective actions. (Source: Adapted from Ref. (38). Sutlief S, Brown D. Taxonomy of corrective actions in radiation therapy. Med Phys. 2014;41(6):433.) With permission.**
Once you have decided on your error mitigation strategy, the second step is to implement it safely and effectively. Safe implementation involves: (a) making sure that the error mitigation strategy will actually address the problems you have identified and (b) ensuring that you do not make things worse by creating negative downstream effects. One way to do this is to redo the analysis that found the problems with the new error mitigation strategy in place. Another is to test the error mitigation strategy in a mock setting. This is a great way to see whether it actually works and will help significantly in elucidating any potential negative downstream effects.

Effective implementation involves: (a) generating buy-in from colleagues who are directly impacted and (b) understanding corrective actions as behavior changes. Generating buy-in is incredibly important, and is largely dependent on good communication. Some strategies for generating buy-in include involving as many stakeholders in the process as possible, communicating to individuals and groups through conversations and presentations, and valuing and acting on feedback from your colleagues. Some strategies for promoting behavior change are practice rehearsal (discussed earlier), the use of clinical guidelines and care maps, and the inclusion of reminders—checklists or time-outs—that improve compliance with new processes.

The third step is to realize that the error mitigation strategy that you have developed will very likely have to be modified once it is implemented system-wide. You should continually seek feedback and be open to adjusting the error mitigation strategy as necessary.

HIERARCHY OF INTERVENTIONS

The Hierarchy of Interventions provides a structured evaluation of the effectiveness of different broad types of error mitigation strategies. Vaida, working at the Institute for Safe Medication Practices, first developed the hierarchy in 1999, and Woods et al presented a modified version in 2008. In the order of most effective to least effective, the Hierarchy of Interventions is presented as follows: (a) forcing functions; (b) automation, computerization, and technology; (c) standardization and protocols; (d) staffing organization; (e) policies, rules, and expectations; (f) checklists and double-checks; (g) risk assessment and communication error; (h) education and information; and (j) personal initiative—vigilance. This hierarchy is often represented as a pyramid, as in Figure 5.3.

It is potentially illustrative to categorize the most commonly cited corrective actions, presented in section IV based on the Hierarchy of Interventions. Table 5.3 presents this categorization.

Interestingly, none of the most commonly cited corrective actions fall into either of the two most effective intervention categories: “Forcing functions” and “Automation,
At a more practical level, an unfortunate reality is that team members experience burnout, with subsequent erosion of their professionalism (43,44). The term in medicine is the disruptive individual: the team member who exhibits derogatory, hypercritical, aggressive, or angry responses in the workplace. This disruptive behavior impacts patient safety in a negative fashion (45). While some individuals display inappropriate behavior as a response to fatigue, stress, or burnout, it is nonetheless essential that the institution or clinic does not tolerate this behavior. There is an emerging realization that there has to be a dedicated effort to promote professionalism by the institution (46). The Brigham model (46) approaches this from the frame that support is needed for the clinician, and a tiered approach to interventions is followed to help individuals correct their negative behavior pattern. This starts with an awareness of how actions were perceived, with an emphasis on helping these persons understand the impact of their intentions, while recognizing that their intent was sometimes radically different. Although many of the efforts are best developed at the physician level, these principles should be followed for the whole team. Physicians are called upon to lead by example (47), and while a formalized professionalism initiative may be outside the reach of many clinics, at the very least, leadership should exhibit exemplary professionalism.

**SUMMARY**

An effective quality management program depends on a robust understanding of error, appropriate classification of error, the application of a systems approach to problem solving, the careful choice and implementation of error mitigation strategies, and a commitment to continuous improvement. With a lot of hard work and ingenuity, clinics can be made safer.
REFERENCES


32. Fogarty GB, Hornby C, Ferguson HM, et al. Quality assurance in a radiation oncology unit:


Plan Quality: The Good, the Bad, and the Ugly

Kevin L. Moore

THE UGLY TRUTH ABOUT PLAN QUALITY

Any characterization of treatment plans as ugly might sound a bit hyperbolic, yet most clinicians have encountered such plans and, unfortunately, patients have in fact been treated with such plans. Another characterization of this phenomenon could be a blunder, defined in Merriam-Webster as “a gross error or mistake resulting usually from... ignorance or carelessness.” The human element as the source of a blunder is unavoidable in this definition, and can thus be distinguished from several well-documented errors (1) in radiotherapy that, at least initially, could be sourced to other mechanisms, such as software glitches, data transfer errors, or hardware malfunctions. Figure 13.1 shows a real-world instance of an ugly treatment plan, wherein two organs-at-risk (OARs) were not appropriately spared when the human planner was performing the intensity-modulated radiation therapy (IMRT) optimization (see dotted lines on the dose volume histogram [DVH]). It was, thankfully, caught before the patient was treated, and an improved final treatment plan was created (solid lines on the DVH). In most instances, treatment planning blunders are not of this magnitude, although assessing the frequency and severity of treatment plan quality deficiencies is not at all easy; this is discussed in more detail later in this chapter.

Classifying treatment plan quality blunders is instructive, as the spectrum of preventative actions depends strongly on the source of the errors. For this we consider three categories of planning blunders: errors of commission (performing the wrong action), errors of omission (failing to perform a required action), and errors of ignorance (lack of understanding of proper actions). Generally, errors of omission and ignorance are more easily prevented through automated measures.

FIGURE 13.1 An example of a treatment planning blunder in head-and-neck IMRT (dotted lines), where the optimization process failed to avoid the larynx and upper esophagus. This ugly plan was reviewed and approved for treatment, but luckily was identified as suboptimal before the patient was treated. When a replan (solid lines) was developed, the delivered dose to these two organs-at-risk was reduced substantially, with no cost to the coverage of the PTVs.
and/or standardized quality checks, whereas errors of commission are more difficult to automate. For example, manual or automated checks can quite easily detect whether an OAR is either empty or not present (error of omission), but it is significantly harder to detect whether a clinical target volume (CTV) has been incorrectly contoured by the physician (error of commission). In Table 13.1, some examples of plan quality blunders are presented, including both manual preventative measures and potentially automated preventative measures.

In an attempt to classify the ugly plan presented in Figure 13.1, it would seem that this cannot be identified as an error of either omission or ignorance, as the OARs were correctly contoured and included in the plan optimization. Thus, this planning blunder must be sourced back to an error of commission, wherein the erroneous action was the use of inappropriate optimization objectives that allowed far more dose than necessary to be delivered to these OARs in the initial plan. Standard quality assurance (QA) for treatment planning workflows include both a physician and a physicist review of candidate treatment plans before treatment (11), but, of course, this falls into the manual preventative measures column of Table 13.1, and, unfortunately in this case, the on-screen physician review failed to detect the suboptimal nature of the initial treatment plan. When IMRT began in the mid-1990s, such a blunder could perhaps have been identified as an error of ignorance, but after 20 years of collective experience with inverse-optimized planning, we cannot claim complete ignorance of what is possible with IMRT.

Such an extreme case is almost certainly the exception, and yet the reader is encouraged to pause and consider just how often such a case could have been delivered to patients in your clinic. Although nearly all clinical scenarios have multiple plan quality criteria and cannot be distilled down to a single number, as a thought experiment one can posit a theoretical axis of plan quality whereby plans can be scored from catastrophic and dangerous (leftmost side of Figure 13.2) to fully optimal with respect to limits of the radiation delivery method used (rightmost side of Figure 13.2). Now consider that every plan ever delivered at your clinic will reside somewhere on this axis, and looking at all of them in aggregate will form a frequency histogram. If this distribution could be known, clinicians could quantify the probability that any given plan might be fatally flawed and institute quality control measures accordingly. With the exception of a handful of studies that will be discussed later in this chapter, clinicians typically do not know what the actual distribution of plan quality is in their own clinic.

Radiotherapy is, quite regrettably, in a state where no one knows how widespread plan quality deficiencies are in clinical practice, nor the degree to which these quality deficits negatively impact patient care.

The quantitative tools that could query large pluralities of patient treatments to generate a true picture of the plan quality distribution in Figure 13.2 are not yet available to most clinicians. However, there are methods that, were they put into the hands of clinicians, could not only illuminate the distribution of plan quality from retrospective samples (without QC curve in Figure 13.2), but also fundamentally alter the shape of the distribution itself through true quantitative quality control tools (with QC curve in Figure 13.2).

The remainder of this chapter focuses on such methods of knowledge-based treatment plan quality control—that is, combating the problem of errors in plan optimization that lead to bad and ugly plans—but the reader is encouraged to consult the references in the last column of Table 13.1 for literature on tools to address the other types of planning errors.

### Table 13.1 Examples and Categorization of Treatment Planning Blunders

<table>
<thead>
<tr>
<th>Type of Radiotherapy Planning Blunder</th>
<th>Examples</th>
<th>Manual Preventative Measures</th>
<th>Automated Preventative Measures</th>
</tr>
</thead>
</table>
| Errors of commission                 | • Contouring errors  
• Fusion errors  
• Prescription errors  
• Inappropriate beam energy  
• Poorly optimized plan | • Pretreatment physics review  
• Peer review (chart rounds) | • Autocontouring checks (2–5)  
• Library search  
• Knowledge-based plan quality control (6,7) |
| Errors of omission                   | • OARs not contoured  
• OARs not included in optimization | • Checklists  
• Peer review (chart rounds) | • Templates  
• Autocontouring (2,3)  
• Reporting tools (8–10) |
| Errors of ignorance                  | • Wrong assumptions (e.g., integrity of CT scan)  
• Dose calculation errors  
• Previous treatment not accounted for | • Pretreatment physics review | • Reporting tools (8–10) |
FROM QUALITATIVE TO QUANTITATIVE PLAN QUALITY ASSESSMENTS

The treatment plan quality problem can be attributed to a number of factors: the anatomic variations between patients, the complexity of clinical goals, the paucity of quantitative metrics to judge the optimality of a given plan, and, of course, the subjectivity and relative experience of the humans involved in the planning process. As explored in many of the other chapters in this book, human error is the vexing source of many critical failure modes in clinical radiotherapy, and treatment plan quality deficiencies are no exception. A fundamental tenet of treatment plan quality control is to remove, to as large a degree as possible, the reliance of plan quality assessments on subjective human experience. To accomplish this, one must be able to make accurate quantitative predictions of plan quality metrics on a patient-specific basis, and so the question now turns to how such quantitative predictions are made.

Some plan quality metrics are universal across patients and are well suited as immutable constraints on some aspects of plan quality. Examples of such instances are the specification of the minimum allowable coverage criteria on a target volume (e.g., \( V_{95\%} \geq 95\% \) (12)) or a hard constraint on a serial organ (e.g., Cord \( D_{\text{max}} \leq 45 \) Gy). Such specifications form the core of radiotherapy protocol guidelines such as those found at the National Cancer Institute (NCI) cooperative groups (13). However, it must be clearly stated that not all dosimetric goals should be set at the same value for all patients. There are numerous clinical examples, but one storied and instructive case study is the parotid gland in head-and-neck cancer. There is a factor of three difference between the region of highest dose response for the parotid gland (20–30 Gy (14–17)) and the therapeutic dose response curve of head-and-neck tumors (60–70 Gy (18)), and yet the advent of IMRT has allowed in-field dose gradients to be tuned to such a degree that salivary function can be preserved in cases that would be impossible with 3D-CRT (19). The planner is thus presented with the ability to effect a broad range of final mean doses to the parotid gland, and the “optimal” value is not constant between patients. On one end of the spectrum, sparing the parotid gland too much at the expense of clinical target coverage is highly undesirable, as evidenced by the increased risk of local failures observed under such a planning strategy (20). On the other end of the spectrum, it is trivially possible to insufficiently spare the parotid gland with IMRT planning by failing to create any negative dose gradient across the parotid gland, that is, in the 3D-CRT limit of opposed lateral fields. Given that salivary flow is a smooth inverse function of dose (15,16), a reasonable planning specification for a parotid gland could be written “as low as possible without losing target coverage,” and yet this does not advance one’s knowledge of what the mean dose to the parotid gland should be for any particular patient.

This ignorance of what is possible for any given patient clearly has a cost. In Moore et al (6), a simple mathematical model was developed from a plurality of previous cases that allowed a patient-specific prediction of mean dose based on the geometric overlap of a parotid gland with the planning target volume, with approximately \( \pm 10\% \) accuracy. The effect of incorporating this prediction into the IMRT planning process was dramatic (Figure 13.3), resulting in significantly lowered and much less variable mean doses delivered patient-to-patient.

![Figure 13.2 Cartoon view of clinical plan quality distributions with and without quality control measures.](image)
This reduction in variability is not merely an academic matter; the patients planned before the model-driven feedback were at much higher risk of incurring salivary complications based on the aforementioned parotid gland dose response curve (6).

The means by which patient-specific predictions are obtained follows from the root cause of the dosimetric variation itself. In most cases, patients are planned according to standardized clinical goals determined entirely by the staging of their disease, meaning that the only predominant difference between one patient and another is their unique anatomy. The understanding of the degree to which a radiotherapy plan can or cannot meet a set of clinical goals for a given patient thus follows directly from the geometric properties of that patient’s underlying anatomy. Although this is simply stated and quite intuitive, this confounding element is the primary reason behind the plan quality variations observed in the clinical practice of IMRT.

The remedy for this lack of clarity is to construct some predictive apparatus that quantifiably connects patient anatomy to expected dosimetry, and like all other scientific endeavors, the best guide for all numerical predictions is careful observation of controlled experimental data and a theoretical framework that allows for accurate future predictions. For radiotherapy clinicians, the obvious source of this data would be previously treated patients who shared the same clinical goals. If there were anatomic quantities that were strongly predictive of output dosimetry, we would need only plot several patients on these axes, as in the diagrammatic representation in Figure 13.4. As the geometric variable(s) for new patients can be computed irrespective of whether a plan has yet been generated, plotting the data on these axes can facilitate a knowledge-based prediction for the new case via two possible strategies. A similar patient (or patients) could be identified who shares similar geometric variables, an approach that could be characterized as a knowledge-based library search. Another approach to utilizing the data cast on these axes would be to develop some curve fit to the data, which could be characterized as a knowledge-based model that converts some geometric variable(s) $X$ to some output dosimetric variable(s) $Y$ via some functional relationship $F$, that is, $Y_{\text{predicted}} = F(X)$.

Returning to the primary goal of this chapter, which is to describe how to identify and eliminate treatment planning blunders from clinical practice, either the knowledge-based library approach or the knowledge-based modeling does much to advance this goal. Should a candidate plan differ greatly from either the matched plan in the library search or the model-predicted value, this could trigger a closer evaluation of whether the difference is “legitimate” (plan is truly optimal and the discrepancy is due to other factors) or is truly a blunder (plan is suboptimal and can be further improved). Both strategies have advantages and disadvantages associated with their use, which are summarized in Table 13.2.

Ultimately, the use of prior experience to inform future cases depends most strongly on the correlative power of the chosen geometric variable(s) to the dosimetric quantity of interest in the absence of plan...
quality variations; that is, how accurately the geometric quantities explain the observed dosimetric variations if all suboptimal plans have been eliminated. The reader would be readily forgiven for balking at this statement as a logical Catch-22, whereby plan quality deficiencies must be eliminated before one could develop the means to construct a quantitative plan quality control system to eliminate plan quality deficiencies. This is a core problem facing knowledge-based methods that purport to identify suboptimal treatment plans, and the quantification of aggregate plan quality deficiencies is the problem to which we must now turn our attention.

**HOW BAD AND HOW OFTEN? QUANTIFYING THE FREQUENCY AND CLINICAL SEVERITY OF SUBOPTIMAL PLANS WITH KNOWLEDGE-BASED PREDICTIONS**

Returning now to the theoretical plan quality distribution of Figure 13.2, we consider how one might answer the key question of ascertaining the nature of the distribution from a sample of $N$ patients’ treatment plans, that is, quantifying the frequency and clinical severity of suboptimal plans in clinical practice. One instructive example from the literature provides a hint at the question of the frequency of suboptimal plans, alternatively interpreted as quantifying the variation in plan quality due to less-than-ideal treatment plan development. In Nelms et al (21), the authors provided a single prostate patient’s data set (simulation CT and contouring structure set) to the community as part of the ROR plan challenge, an annual contest whereby entrants attempt to achieve the highest possible plan quality metric (PQM) score according to a specified combination of DVH-based target and organ metrics. Upon analysis of the more than 125 submitted cases, the authors ultimately reported the following: “There is a large inter-planning variation in plan quality defined by a quantitative PQM score that measures the ability of the planner to meet very specific plan objectives. Plan quality was not statistically different between different treatment planning systems or delivery techniques and was not correlated to metrics of plan complexity. Certification and education demographics, experience, and confidence level of the planner were not good predictors of plan quality.” Figure 13.5 shows the observed distribution of PQM scores, exhibiting an extremely wide variance across the submitted cases. That the treatment planning system and delivery technique were not correlated with PQM score provides evidence that the limitation in achieving high-quality plans is not the fault, at least in this case, of the tools at the disposal of the planners. Although this study was limited to a single patient’s data set and was unable to ascertain any specific factors that correlated with suboptimal plan submission, the breadth of plan quality variations in normal clinical practice could be reasonably inferred to be quite large given the very reasonable assumption that the individuals submitting contest submissions were a self-selected group who presumably thought their plans were good enough to warrant submission to a plan quality contest.

Bookending this study are several investigations, drawn from multi-institutional clinical trials, correlating the failure to meet dosimetric protocol constraints with worsened outcome (22–25). The outcomes tracking and centralized quality assurance of these trials provide strong evidence that buttresses the intuition that large deviations from clinical quality parameters really do subvert patient outcome, and yet these studies were unable to provide an answer to the question of whether any of the protocol-noncompliant cases could actually have been made to be compliant through better planning.

A knowledge-based methodology of collecting previous patients and developing sufficiently accurate dosimetric
predictions does have the ability to address both the frequency and the clinical severity aspect of the plan quality problem. Focusing on one specific knowledge-based modeling strategy, we can examine how making accurate patient-specific DVH predictions could accomplish this. In Appenzoller et al (26), a knowledge-based method is described whereby parametric models are developed on the basis of the correlation of expected dose to the minimum distance from a voxel to the planning target volume (PTV) surface. A three-parameter probability distribution function (PDF) was used to model iso-distance OAR subvolume dose distributions. The knowledge-based DVH models were obtained by fitting the evolution of the PDF with distance, yielding accurate organ-specific models that make DVH predictions based on the synthesized experience of the training set. Figure 13.6 depicts the comparison of a predicted DVH to a candidate clinical DVH, putting into sharp relief the potential of accurate DVH predictions for plan quality control. Depending on the component (e.g., $V_3$ or derivative element (e.g., equivalent uniform dose [EUD] (27)) of the DVH that matters clinically, a comparison of the predicted dosimetric value against the candidate plan’s value gives immediate quantitative feedback (within the accuracy of the DVH prediction) as to whether and by how much a plan could be improved.

With this model-building and plan quality scoring apparatus, the elucidation of the plan quality distribution becomes a matter primarily of finding a properly representative data set. In Moore et al (28), parametric knowledge-based model DVH predictions were brought to bear on a multi-institutional clinical trial to directly assess the frequency and clinical severity of suboptimal treatment planning on a large scale: The Radiation Therapy Oncology Group (RTOG) 0126 protocol, A Phase III Randomized Study of High Dose 3DCRT/IMRT Versus Standard Dose 3DCRT/IMRT in Patients Treated for Localized Prostate Cancer. Briefly, this work examined high-dose IMRT patients (79.2 Gy prescription dose in 44 fractions), and after using knowledge-based DVH prediction models (26) to eliminate other plan quality indices as problematic, the study ultimately focused on grade 2+ late rectal toxicities as the primary plan quality marker. Using an outcomes-validated Lyman–Kuthcher–Burman (LKB) model for the normal tissue complication probability (NTCP) for late rectal toxicities (29), comparisons between clinical and model-predicted DVHs yielded the absolute excess risk (actual NTCP – predicted NTCP) from suboptimal planning: $94/219 (42.9\%)$ had $\geq 5\%$ excess risk, $20/219 (9.1\%)$ had $\geq 10\%$ excess risk, and $2/219 (0.9\%)$ had $\geq 15\%$ excess risk. The results of this study definitively demonstrated that poor-quality IMRT planning frequently put protocol patients at substantial and unnecessary risk of normal tissue toxicities. The excess risks were directly found to be unnecessary, as the validated LKB model quantified the potential risk reductions and replanning demonstrated the achievability of this reduction without compromise of target structures or other OARs. The characterization of $+4.7\%$ average excess risk as “substantial” requires context, to which two key points of comparison are available from Michalski et al (30). On the high-dose arm of RTOG 0126, at median follow-up of 3 years, patients treated with 3D-CRT had a $22.0\%$ cumulative incidence of grade 2+ GI toxicity, whereas IMRT patients had a $15.1\%$ cumulative incidence ($\chi^2 = .039$). A $4.7\%$
risk reduction in the IMRT group rate might have cut the incidence by nearly a third. Further, the predicted 4.7% risk reduction is on a par with the toxicity rate difference between 3D-CRT and IMRT of 22.0% to 15.1% = 7.1%, implying that quality-controlled IMRT planning could have yielded nearly as much clinical benefit as uncontrolled IMRT planning offered over 3D-CRT (Figure 13.7).

It is highly likely that, at least for the time period from which the plans were drawn (2003–2008), this study is representative of the plan quality variations present in the general population. Given the types of institutions that participate in national trials and the efforts expended to meet protocol requirements, it is plausible that the plan quality variations in this work actually underestimate the variability in wider clinical practice. The observed quality variations in this study hint, troublingly, at potentially larger variations in sites where the spread between the prescription dose and organ tolerances is even wider, for example, the aforementioned case of parotid glands in head-and-neck cancer. Further investigations into plan quality variation distributions in different disease sites and larger clinical data sets will likely expand as knowledge-based tools become more widely available.

AUTOMATED PLANNING WITH KNOWLEDGE-BASED DOSE PREDICTIONS

Of course, the promise of knowledge-based plan quality control is not just a quantification of clinical plan quality

FIGURE 13.7 Secondary study on RTOG 0126 quantified excess risk of late rectal complication due to suboptimal IMRT planning. (A) Scatter plot shows knowledge-based prediction of NTCP versus the actual treated plans’ NTCP. (B) Frequency histogram of the same data shows a mean excess risk of 4.7% ± 3.9% and represents the largest and most institutionally diverse measurement of plan quality variations to date. (Source: From Ref. (28). Moore KL, Schmidt R, Moiseenko V, et al. Quantifying unnecessary normal tissue complication risks due to suboptimal planning: a secondary study of RTOG 0126. Int J Radiat Oncol Biol Phys. 2015;92:228–235.)
variations, but an entire reshaping of the distribution itself (Figure 13.2). With objective dosimetric predictions for individual patients, it is an easily logical step to move from knowledge-based quality control (comparing a candidate plan with the knowledge-based prediction) to fully automated knowledge-based planning (KBP), whereby the dosimetric predictions are directly incorporated into the plan optimization process. The utility of this can be readily seen by a reconsideration of the ugly plan in Figure 13.1, where the only plausible explanation for these dangerous OAR DVHs is that inappropriate DVH objectives and/or weights were entered into the IMRT optimization engine. Had knowledge-based DVH predictions resembled closely the final plan’s solid DVH curves available, and had the IMRT optimization objectives been based on these, the inferior plan (dotted DVHs) would not have resulted. There are several means by which knowledge-based DVH predictions are used to automate the generation of the discrete DVH-based IMRT objectives (8,31–33), including one commercial offering (34) at the time of publication. Knowledge-based DVH prediction models are necessary but not sufficient for such automated KBP, as the estimated curves have to be converted to DVH objectives and priority weights. Here we see a two-fold layering to any automated KBP schema, whereby the DVH prediction algorithm (the “model”) must be sufficiently accurate and the relative weighting of the generated DVH-based objectives have to be properly calibrated against each other (the autoplanning “routine”). The independent validation of both the models and the autoplanning routines is a critical step, and several approaches have been presented in the literature (26,28,33–36), by setting aside a sizeable number of retrospectively collected patients from the model training set and then applying the DVH estimation models and autoplanning routines to these “new” patients to confirm that the output of the autoplanning system is in line with previous experience. The process of applying a trained dose prediction model to a set of unseen validation patients is the primary protection against overfitting and, assuming the validation set contains predominantly high-quality plans, quantifies the uncertainty in the model’s dose metric prediction. If the plans in the validation set have unknown quality, the incorporation of some replanning endeavor will be required to separate whether prediction inaccuracies are true model uncertainty or should be ascribed to plan quality variations.

**QUIS CUSTODIET IPSO CUSTODES? (WHO WILL GUARD THE GUARDS THEMSELVES?)**

If knowledge-based automated planning is supposed to protect clinicians and patients from suboptimal plans, how can we be certain that it is performing this task flawlessly in all cases? To explore the implications of this question, as a thought experiment we can envision a knowledge-based automated planning system for an arbitrary disease site, an arbitrary delivery technique, and a set of several dosimetric plan quality indices of interest. For every plan quality index, we can consider three quantities:

- $D_{\text{optimal}}$—The optimum dose value for the plan quality index, taking into account the delivery technique and the necessary trade-offs of this metric against the other plan quality indices
- $D_{\text{model}} \pm \delta D_{\text{model}}$—The knowledge-based model prediction for the plan quality index, including some error bounds quantifying the precision of the prediction
- $D_{\text{autoplan}}$—The resultant value of the plan quality index upon use of the knowledge-based autoplanning routine to generate a deliverable plan with the given delivery technique

It must be noted that $D_{\text{optimal}}$ is never actually known for all metrics. The closest we might get to discovering $D_{\text{optimal}}$ for all the quality metrics in an individual case would be to go through something like the multiuser planning contest that generated the distribution in Figure 13.5. The highest scoring plan could be reasonably assumed to approach $D_{\text{optimal}}$ for each of the multiple planning criteria, but of course, such a process is not possible in normal clinical practice. Table 13.3 examines a theoretical instance of three OAR quality indices with comparisons for $D_{\text{optimal}}$, $D_{\text{model}} \pm \delta D_{\text{model}}$, and $D_{\text{autoplan}}$ in each dose metric.

With $D_{\text{optimal}}$ unknown, it is not obvious how to determine whether any of a particular plan’s quality metrics might be falling victim to one of the error types, nor how this will be limited. Of the two error types, “false negatives” may be the most concerning, as these would represent precisely the inadequate dose sparing that knowledge-based planning is designed to mitigate. This consideration makes the quantitative assessment of the model uncertainty $\delta D_{\text{model}}$ all the more critical, and although Table 13.3 utilized a standard deviation to describe the model error, other quantifications (such as box-and-whisker plots examining the full range of observed model prediction variation) can also be valuable quantitative feedback to the clinician as to how far the model prediction might deviate.

In addition to normal and expected model prediction variation, there is the matter of geometric outliers, that is, patients whose anatomy lies outside the observed range of geometric variables in whatever space has been used to train the knowledge-based system. This can be envisioned by a new patient taking a geometric variable value to the left or to the right of all of the training data in the cartoon picture of Figure 13.4. One must then consider whether the knowledge-based model can be legitimately extrapolated to this as-yet unobserved value or, if a knowledge-based library approach is used, how to
match this patient to the existing database. To complicate things further, all modern knowledge-based systems (31,33,37) utilize a set of geometric inputs comprising a multivariable space that is not readily visualized, so a mechanism to identify outliers in a multidimensional space (38) will be required to understand the degree to which an individual observation is within the range of or outside the geometric variation in the training data. The user’s response to identification of a geometric outlier should be one of caution and heightened skepticism, not only with regard to the values of the model’s dosimetric prediction but to the model uncertainty estimates as well. In these instances, extra scrutiny of the autoplaning routine’s output is a necessity, up to and including manual adjustment of the optimization parameters in a replanning effort to ensure that the resultant plan is not readily improved in any of the plan quality dimensions. It would be advisable to both retain the KBP result to compare against the replan(s) and to incorporate some reporting system (10,36) that can quickly analyze all quality metrics together in case one was improved at the expense of another.

The imperative to continually expand the breadth and accuracy of KBP routines finally brings us to the notion of knowledge-based model maintenance. As geometric outliers are discovered in the course of normal clinical practice, it would be highly advantageous to incorporate new information into updated models and improved autoplanning routines. Unlike the treatment planning system (TPS) beam models used for dose calculation that are rarely (if ever) changed after commissioning, the merging of new patient plans into the knowledge-based models is readily accomplished by retraining the system with an expanded patient pool. Besides the inclusion of geometric outlier patients to broaden the dynamic range of the knowledge-based predictions, it is entirely possible to continually add high-quality plans, perhaps even filtering out training plans that appear inferior to the quality distribution within the training sample itself. The difficulties of this value proposition lie in the bookkeeping aspects of keeping track of several models and in the concern that the model output may drift to a new state, thus inadvertently driving the clinical trade-offs in an undesirable direction. The latter worry can be substantially alleviated by maintaining a large and constant validation pool of patients to which separate models can be applied. The performance of differing models can be assessed on the aggregation of the many PQMs across a constant plurality of patients, so any unintended consequences of adjustment of the training pool on the model’s performance should be transparent and easily discovered.

### CONCLUSIONS

Although this chapter has focused on knowledge-based dose predictions as a potential solution to the problem of plan quality variations, it is notable that knowledge-based techniques would still remain susceptible to several of the radiotherapy planning blunders listed in Table 13.1. Without the addition of further safety measures, for example, it is easy to see how prescription errors, contouring errors, and dose calculation errors could still propagate through the planning process even if KBP

---

**Table 13.3 Example of a Set of Dosimetric Plan Quality Metrics and How Knowledge-Based Model Predictions and Autoplans Might Compare With The Fully Optimal Values**

<table>
<thead>
<tr>
<th>D_{optimal} (Gy)</th>
<th>D_{model} ± δD_{model} (Gy)</th>
<th>D_{autoplan} (Gy)</th>
<th>Error Type</th>
<th>Consequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>32.4</td>
<td>32.6 ± 0.7</td>
<td>32.8</td>
<td>False-negative</td>
<td>OAR underspared, plan is suboptimal with respect to this metric</td>
</tr>
<tr>
<td>23.8</td>
<td>26.1 ± 2.3</td>
<td>26.0</td>
<td>False-negative</td>
<td>OAR overspared, targets or other OARs are compromised, plan is suboptimal owing to incorrect trade-off</td>
</tr>
<tr>
<td>11.0</td>
<td>9.3 ± 0.8</td>
<td>9.4</td>
<td>False-positive</td>
<td>OAR overspared, targets or other OARs are compromised, plan is suboptimal owing to incorrect trade-off</td>
</tr>
</tbody>
</table>

δD_{model} is expressed here as a standard deviation, but could be represented by any statistics from the application of the model to an established validation set. The first dose metric is an example of normative behavior where the model prediction closely approximates the ideal value, and the autoplan result closely follows this. The second dose metric is an example of a “false-negative” error, whereby the model prediction overestimates by a significant margin the truly optimal value, resulting in a final plan value that is not as low as possible for the associated organ. The third dose metric is an example of a “false-positive” error, whereby the model prediction is significantly lower than the truly optimal value, and the resultant plan value forces an inappropriate trade-off to achieve this erroneously aggressive value.
reduces or eliminates the chance of highly suboptimal plans being generated. As should be clear from a survey of the other chapters in this book, a modern quality and safety apparatus must include a portfolio of checks and redundant quality systems to ensure consistently optimal behavior. We should go into the future with strong optimism that the rate of “ugly” and “bad” plans will decrease, although it will take quite some time to automate all treatment sites and for these systems to become ubiquitous.

REFERENCES


INTRODUCTION

Incidents in radiation therapy can have dramatic consequences. Fortunately, most incidents are minor and the majority are preventable. Mitigation strategies can be developed after understanding the nature, frequency, and potential severity of such events through the process of incident learning (1). Prospective utilization of methods such as failure modes and effects analysis (FMEA) and analysis of event root causes can identify areas for systematic improvement that can lead to enhanced patient safety (2). The incorporation of advanced technologies such as intensity-modulated radiation therapy (IMRT) and stereotactic radiation hold the promise of therapeutic ratio improvement. However, it is known that the implementation of such technologies can be associated with an increased risk for error, particularly if not properly evaluated and quality assured (3).

As our field advances, it becomes increasingly important that appropriate systems to report and learn from incidents and errors be implemented both intradepartmentally and intranstitutionally, as well as nationally. Intradepartmental and hospital-wide reporting systems can aid in uncovering specific process errors that may be unique to a departmental culture or workflow. Incident learning systems are in place in many institutions throughout the country (4–6). It is feasible to use a hospital patient safety reporting system for reporting, but the complexity of the radiation oncology process often cannot be captured in a generic hospital system. A radiation oncology–specific system has the advantage of capturing relevant field-specific information that can be used for unique quality improvement (7).

Patients are at risk for events regardless of the type of center in which they are being treated. National and international reporting systems facilitate the collection of data across a range of sites, including both community and academic settings, to determine the common elements that may put patients in harm’s way. The Radiation Oncology Incident Learning System (RO-ILS), a centralized national reporting system in the United States, allows input of data from radiation treatment centers across the United States and was established to facilitate the reporting of events nationally (8). In this way, safety standards can be derived on the basis of those incidents occurring on the ground. An international reporting system (Radiation Oncology Safety Information System [ROSIS]) has been online for years, collecting information through a voluntary reporting structure (9). The International Atomic Energy Agency (IAEA) is also spearheading the development of an incident learning system named SAFRON (Safety in Radiation Oncology). Consensus recommendations for incident learning database structures in radiation oncology have been published by the American Association of Physicists in Medicine (AAPM) Work Group on the Prevention of Errors (WGPE) (10). Incident learning has the potential to improve patient safety and quality by facilitating knowledge sharing that can lead to incorporation and prioritization of process improvement initiatives and quality assurance program evaluation.

Incident reporting systems can also potentially be used to identify treatment-related factors that may increase the risk of radiation safety events. Patient- and disease-specific risk factors could be identified and used for error prevention prospectively at the time of treatment plan review and during the course of treatment delivery. The treatment process may then be modified for those patients thought to be at higher risk for errors. However, voluntary incident reporting systems are not without their pitfalls. The potential for large variability in the nature and frequency of reports received when incident reporting is voluntary is well described. Ultimately, reports that are submitted are subject to the biases of the observers reporting; thus, reports could represent a limited sample of the events actually occurring in the care setting (11). Therefore, incident reports may not include all areas that pose risk for patients. Despite this issue, incident reporting systems still provide an objective means for evaluation of actual events occurring during the workflow or to patients themselves. Ultimately, these incidents provide us with a way in which we can prioritize and develop mitigation strategies to reduce the incidence of events.

INCIDENT LEARNING SYSTEMS

Incident learning systems have been used in individual radiation oncology departments for many years. Here we distinguish between incident learning systems and event...
Reporting systems, which have been part of radiation oncology practice for decades. Every department reacts, formally or informally, when it is recognized that a patient has been treated contrary to plan. The staff deals with the issues associated with that particular patient, which may involve making adjustments to the patient’s subsequent treatments (or not), reporting to risk managers or regulators (or not), and making changes to processes (or not). A true incident learning system (ILS) formalizes the process by which the organization recognizes and responds to events in order to make treatments safer. Several steps characterize that process:

1. Identifying that something has happened that should be addressed
2. Reporting the issue
3. Investigating the circumstances
4. Analyzing the issue for proximate and underlying causes
5. Designing corrective actions
6. Implementing corrective actions
7. Evaluating the effectiveness of the corrective actions

An effective ILS must accomplish each of these steps. A number of authors have published their experience in designing, implementing, learning from, and evaluating the effectiveness of an institutional ILS, and we will refer to those as we consider each of these steps in turn.

Identifying that something has happened

What should be reported to the ILS? It is generally understood that any delivery of a therapeutic radiation dose contrary to the intended plan constitutes an event that should be reported. Similarly, an intervention that discovered and prevented a misapplied dose, that is, a “near miss,” should be reported. Several authors comment that it is also important to recognize and respond to the “various hassles that practitioners experience in their everyday clinical work” (12). Ford et al (7), chose to use the term incident

 mainly to highlight the fact that all deviations are potentially of interest even those that do not necessarily impact the patient. Other terms could be used equally well such as “variance,” “event,” or “condition.” Incidents also include deviations or variations to expected workflow, conditions that would impede the smooth completion of a task without a workaround of the standard process.

Reporting and responding to all such process variances, not just those clearly related to safety, has a number of benefits. One is that it promotes a culture of reporting and expecting that things can be made better, so that staff are more likely to report issues that are clearly related to safety. Another is that eliminating hassles and workarounds and improving efficiency also improves safety. This is Reason’s well-known Swiss cheese model of how errors come about (13). “In this way, information is gathered from frontline staff about contributory factors that might at some point lead to patient harm if left unaddressed” (12).

Authors have published the frequency of incidents reported in their ILS, and the numbers have varied widely. Those that encourage staff to report all process variances have seen 1 (14) to 0.6 (4) to 0.14 (15) events per patient, whereas lower numbers are more common, e.g., 0.1 to 4.7 per 100 patients treated (5,16–22). Mutic et al (4), who reported 0.6 events per patient, classified 7.8% of them as of high or critical importance (see following text), corresponding to 4.7 per 100 patients treated. Others have seen ratios on the order of 1:10:30:600 for critical:serious:minor categories of incident (23). These data support an expectation that a robust ILS could see on the order of 0.1 to 1 report per patient, some 60% to 90% of which will be of low or minor immediate significance. This has ramifications for the system required to analyze and respond to the reports.

Reporting the issue

The reporting mechanism should be readily accessible and easy to navigate. Several authors have reported significant increases in the number of reports submitted after shifting from a paper-based system to an electronic system, ranging from factors of 1.8 (20) to 4 (4) to 9 (14), although that effect has not been universally seen (17). In fact, introduction of an unwieldy electronic system was found to decrease reporting (24).

If the desire is to have frontline staff members routinely report all manner of incidents, then it is important that the link to the reporting tool be ubiquitous within the department and that the initial report require a minimal amount of information. The frontline reporter must feel few barriers to reporting. Submitting that initial report should then flag others who have the responsibility to dig into the details and contributing causes, flesh out the narrative description, and classify the incident according to the chosen taxonomy. That two-step workflow has been used by institutionally designed systems (4,25,26), recommended by an AAPM Working Group (7), and used by the national ILS for Radiation Oncology (RO-ILS) sponsored by the American Society for Radiation Oncology (ASTRO) and AAPM (8).

In order to promote reporting, making the system easy to use is necessary but not sufficient. Other key strategies include ensuring that the response to reports is supportive and nonpunitive (27); rewarding reporters, as with a “Good Catch” award (28, 24); and providing reliable feedback to the individual reporters and to all the staff about the changes
made in response to the reports (27–29). In addition, it is necessary to provide training to staff on when and how to use the system. Using standardized, hypothetical events in the training can promote consistent and effective use (10). Even institutions with sophisticated internal reporting systems can experience reluctance to report, especially among physicians, who have expressed more concerns than other staff about getting colleagues in trouble, liability, effect on departmental reputation, and embarrassment (30). In the long term, it will be helpful to incorporate into medical residency training the expectation to report concerns in the ILS.

Investigating the circumstances; analyzing the issue for proximate and underlying causes

The initial report will rarely provide sufficient information to permit a useful response to be determined. Those charged with reviewing the report will need to augment the initial narrative with a fuller description of the incident and surrounding circumstances. The narrative is key; effective corrective action can be designed only after a deep understanding of the incident, relevant circumstances, and contributing causal factors has been achieved. It can be a challenge to condense that description into a succinct narrative, but necessary if the report is to be the basis for later review. Furthermore, if the report is intended to be useful and comprehensible to people outside the department (e.g., if it will be reported to a national system), then the authors of the narrative need to keep that target audience in mind and avoid department-specific jargon.

The reviewers will also need to classify aspects of the incident in order to enable statistical analysis of the population of reports. Such analysis is useful in at least two ways. One is to help identify patterns of failure so that attention can be effectively focused, and the other is to help measure the effectiveness of the ILS. As an example of the former, it can be useful to see what types of errors in treatment planning are caught by the physics plan check (often the first available safety barrier) and what types are not caught until the first treatment (after having gotten through some checks). Having data elements that keep track of where the error occurred and where it was found can help sort that out. As an example of the latter, authors have reported a decrease in the average severity of reported incidents over time, even as the number of reports has increased (29) or in the frequency of actual treatment deviations as compared with near misses (23). Designing a classification scheme that is comprehensive, robust, and manageable is a difficult task. The consensus document published by the AAPM Working Group (7) offers one approach.

A particularly tricky issue is estimating the severity of the incident. Because of the latency of radiation effects, both beneficial and detrimental, assigning severity nearly always entails a degree of speculation and subjectivity. There is no standard accepted scale. Most systems use a 4 or 5 scale; as an example, that proposed by Mutic et al (4) is summarized here:

1. Low: Not a dose delivery–related event but an inconvenience to the department or patient. These are events that result in inefficiencies and repeat or unnecessary work.
2. Medium: Limited toxicity (may not require medical attention) or minor underdose to the planning target volume. These events have the potential to result in minor deviations of delivered doses. Minor deviations are considered to be less than 10% or 10 mm in magnitude.
3. High: Potentially serious toxicity or injury (may require medical attention) or major underdose to the target. These events have the potential for significant deviations from the intended treatment and are all greater than 10% or 10 mm in magnitude.
4. Critical: Selecting this option causes the software to immediately send alpha pages to the involved area supervisors; it is intended for use only in urgent situations.

The AAPM working group consensus document (10) uses a 10-point scale for dosimetric severity and medical severity. Carefully designing the severity scale, educating the staff who do the scoring as to its use, and testing the consistency of its application are very important, especially when the number of reports requires some prioritization as to which receive the most attention, such as a root cause analysis.

Further complicating the issue of severity is the problem of near misses. If the error was avoided, or was caught very early in the treatment, then does one score what happened or what might have happened? The proper answer may be to score both (15). Only scoring what actually happened would discount a potentially catastrophic near miss, but imagining the worst possible scenario for every potential error can dilute the focus on the most concerning incidents.

There is value in having an intradepartmental team review the reports to look for underlying causes and to design corrective action (25). Very often, the underlying issues will involve workflows and communication channels affecting multiple sections of the department, and insights will come from multiple areas. The relevant supervisor should do an initial investigation to ascertain the details of the event so that the committee will have good information to work with. This implies a multiple-step workflow: the frontline person puts in a report, the supervisor looks into the situation, and then the intradepartmental committee discusses those reports requiring broader review.
NATIONAL AND INTERNATIONAL REPORTING SYSTEMS

Expanding from using a local incident learning or practice improvement system to a national or international one makes intuitive sense. Learning about incidents, near misses, and unsafe conditions from other institutions offers substantially greater opportunities than a single internal system can provide. One caveat, however, is that the heterogeneous nature of the training, technology, resources, infrastructure, staffing, and workflows at other facilities, and in other countries, can make applying the lessons learned to your own department more challenging. Another issue to confront is the potential lack of uniformity, or interreviewer reliability, of reports from different organizations within the same reporting system. Analyses of this variability have been done, and multiple groups have shown that results from reviews are divergent—with suggestions that longer narratives and improved taxonomies will lead to more consistent scoring results (31–33).

Many national and international groups have waded into this arena, and in this section, we report on some of the more well-known efforts—some which have been in existence for years and others that are in the more nascent stages of development. Unfortunately, at the time of this writing, although the need for national and international systems has been well recognized, experiences with the systems have not yet been well published.

Radiation Oncology Safety Information System

The ROSIS was established in 2001 by the European Society of Therapeutic Radiology and Oncology (ESTRO) as a patient safety tool with exclusive emphasis on radiation oncology. It is a voluntary reporting system for both incidents and near misses, with an international focus. Reports related to linear accelerators, cobalt-60 units, and brachytherapy devices are accepted. As of early 2014, more than 150 departments had registered with ROSIS; initially registered departments were located within Europe, but there is now a more diverse global distribution of departments in ROSIS (34).

ROSIS classifies incidents into two broad areas: process-related (where the occurrence is related to a failure in the process), and non–process-related (including hardware and software failures, falls, etc.). The process-related activities are further divided into subgroups related to the point at which they occurred.

At the time of initial registration, a departmental profile is established related to equipment, staff, and environment. However, clinic details are confidential, are stored separately from the incident information, and are not searchable. Incident reports can be made through a variety of methods: transfer via local ROSIS patient safety module, direct input via the website, transfer and translocation via web service (cross-organizational system), and transfer and translocation via web service (local system) (34).

The incident reports are anonymized and stored in a searchable database, and made available to participants on the ROSIS website in their original text. As a further means of communication and engagement, ROSIS has published six themed newsletters to date, some with spotlight reports. Annual workshops are also held, which focus on generalized quality and safety topics, as well as ROSIS updates. The most up-to-date information about ROSIS can be found at http://rosis.info.

Safety in Radiation Oncology

The SAFRON system was developed by the Radiation Protection of Patients (RPOP) unit at the IAEA and piloted in 2012 by an international group of health care professionals from Argentina, Australia, Canada, Ireland, Malaysia, Nigeria, Sweden, United Kingdom, and the United States. It was designed to be both a facility-specific system and a support for an international error reporting system. As of June 2013, there were more than 1,100 reports, including both incidents and near misses—although not all of those were entered prospectively; some were from past reports to the IAEA and to ROSIS (35). The integration of prospective risk analysis together with retrospective reporting enables the system to be proactive, which is of value when considering the rapid development of new medical technology (36).

Facilities wishing to participate in SAFRON must first register with NUCLEUS, the agency’s information resource catalog, and then separately register with SAFRON through a web portal. Currently, only external beam reports are accepted. SAFRON is searchable, and sorts incidents by such details as where in the process the incident was discovered, who discovered the incident, how it was discovered, and any word in the free-text fields (e.g., summary and description of incidents). It is possible to use these criteria in any combination as search filters. Another search feature permits searching on the basis of incidents from the home institution, or the database at large. The identifying information for facilities that are not your own remains anonymous. Participants can review causality information and what corrective actions have been taken by the reporting institution in order to build a resilient system and reduce the likelihood of having a similar incident at their institution. SAFRON publishes themed newsletters that share tips for using SAFRON, and spotlight cases. Aggregate data about types of reports are also offered. The newsletters are e-mailed, but only the most recent edition is currently available. The most up-to-date information about SAFRON can be found at https://rpop.iaea.org/SAFRON.
National System for Incident Reporting in Radiation Treatment

The Canadian Partnership for Quality in Radiotherapy (CPQR) has partnered with the Canadian Institute of Health Information (CIHI) on the National System for Incident Reporting in Radiation Treatment (NSIR-RT) project. The NSIR-RT is still under development, but it aims to be a dynamic tool that will include a searchable database. Events are expected to relate to “learning about, and reporting events arising along the RT delivery continuum, from the decision to treat to treatment completion” (37). The system looks at data for events within Canada, in three areas: actual incidents that reach the patient, near misses, and reportable circumstances (hazards not involving a patient). It is designed for use as a national event reporting system in conjunction with a local system, or to serve as the local system in facilities that do not have one of their own.

During the development of the system, existing incident reporting systems, including WHO, ASTRO, AAPM, and SAFRON, were reviewed to identify relevant elements and highlight their similarities, differences, and existing gaps in order to elaborate the key elements for the NSIR-RT. Efforts were made to maintain compatibility with those systems for future international collaboration. The taxonomy was launched in 2015, and a consensus process is under way. A published report about this consensus process indicated that 13 data elements were thought to warrant inclusion, and that near misses would be included in the implicit definition of an “incident” and not listed separately (38). The most up-to-date information about CPQR can be found at www.cpqr.ca/.

Radiation Oncology Incident Learning System

In June 2011, ASTRO approved a proposal to create a RO-ILS, and was later joined in the endeavor by AAPM. In 2013, a contract was signed with Clarity PSO, a patient safety organization, to afford special legal protections to participants. The patient safety work product (PSWP) entered into the system is privileged and confidential, allowing for analysis in a safe and protected environment. A beta test was executed between September 2013 and November 2014 among 14 sites from a variety of practice types.

As of February 24, 2015, 46 institutions had signed contracts with Clarity PSO; of these, 27 sites had entered 739 patient safety events into local database space, with 358 (48%) events pushed to the national database (39). The database accepts input related to both events and near misses, and functions as both a local reporting system (if desired) and a national database. Currently, the database is searchable only for reports from your own institution; because reports are not anonymized, they cannot be shared with all participants. The main method for dissemination of information is through quarterly newsletters. The most up-to-date information about RO-ILS can be found at www.astro.org/Clinical-Practice/Patient-Safety/ROILS/Index.aspx.

Other systems of note

Other systems do exist, and the motivated reader can find more information about them. These include, but are not limited to, STARSweb interface from the Health Service Executive (HSE) in Ireland (www.hse.ie); the National Health Service system in England and Wales, which is a collaboration between the National Patient Safety Agency, the Royal College of Radiologists, the Health Protection Agency, the Society and College of Radiographers, and other stakeholders, and based on Toward Safer Radiotherapy (40); RIRAS (Radiotherapy Incident Reporting and Analysis System) from the Center for Assessment of Radiological Sciences (CARS) with certification from the Agency for Healthcare Research and Quality (AHRQ) in the United States (www.cars-psi.org/); ROIRLS (Radiation Oncology Incident Reporting and Learning System) in the United States for Veterans Administration patients; two systems sponsored by the Conference of Radiation Control Program Directors (CRCPD) in the United States (www.crcpd.org/): a pilot tracking system for machine-based radiation medical events, and the Committee on Radiation Medical Events (H38) system to collect events reported to state agencies.

METHODS FOR CREATING AN EFFECTIVE FEEDBACK/LEARNING SYSTEM; MEASURES OF EFFECTIVENESS

Consensus recommendations on the development of incident reporting systems were recently published by the AAPM. This report includes process maps that list the key processes in the radiation oncology workflow to aid in the codification of incidents for both external beam therapy and brachytherapy (10). Severity scales are also discussed and categorized into a clinical/medical severity scale, as well as a dose deviations scale for dosimetric consequences that may have ultimately small or virtually no clinical impact. It is recommended that incidents be coded with both scales in the reporting system. The guidelines also include a taxonomy list of root causes and contributing factors to classify incidents, with the goal of identifying possible areas for improvement.

Additional factors must be considered in developing and maintaining an incident reporting system. The system itself may be relatively straightforward to develop from an information technology (IT) perspective, particularly as there are several models to use (4,6,23). However, the time, training, and resources required to evaluate the incidents submitted must be considered, particularly if the system is to be used for process improvement intradepartmentally.
Additionally, it is critical to provide feedback to the staff to incentivize departmental members to spend the time using the system, particularly if the system is voluntary (30). Thus, it is not enough to simply have an incident reporting system. To have a positive impact in the clinic, considerable resources must be directed toward maintaining the system, and conducting the analyses necessary to develop the corrective actions that will ultimately make patients safer.

In order for incident reporting systems to flourish, a safety culture must be encouraged. In a multi-institutional survey conducted to assess barriers to safety reporting, physicians and other staff felt it was their responsibility to report near misses and incidents (97% overall). Physicians were significantly less likely than other groups to report minor near misses ($P = .001$) and minor errors ($P = .024$). Physicians were significantly more concerned about getting colleagues in trouble ($P = .015$), liability ($P = .009$), effect on departmental reputation ($P = .006$), and embarrassment ($P < .001$) than their colleagues. Thus, even though all members of the radiation oncology team observe errors and near misses, physicians are significantly less likely to report events than other colleagues. Therefore, specific barriers to physician reporting should be investigated and analyzed to improve the systems in place today (30).

Ultimately, an effective incident reporting system should lead to mitigation strategies to prevent the same or a similar event from occurring, and thereby enhance patient safety. Several case examples have been previously published that describe effective changes that have led to improvement in processes (26). Although a mitigation strategy may be implemented, it is not always clear that the approach actually had its intended effect on improving safety or quality. The Plan–Do–Study–Act (PDSA) approach is one such method to qualitatively, and perhaps quantitatively, evaluate a new process to determine whether it succeeded in its intent. Although the PDSA approach has not been explicitly studied in radiation oncology, it is commonly used in health care for quality improvement. The PDSA cycle begins with the “plan,” which involves the identification of the issue to be addressed. After the goal is set, the “do” step involves the actual implementation of the new process. Subsequently, the “study” step involves monitoring and collecting data on the outcomes after the new process has been implemented. New areas for improvement may be identified during the “study” step. Lastly, the “act” step completes the cycle and integrates the information learned from studying the outcomes to inform the next steps in meeting the goal of quality improvement. This may result in a change in the goal or methods that may be minor or major. The PDSA cycle is repeated with each of the four steps, in an iterative fashion, in order to provide effective feedback and continual improvement (41).

Ultimately, the value of an incident learning system depends on how effectively it contributes to making positive change happen. The staff must feel that the effort they put into reporting makes a difference in their lives and in the lives of patients; otherwise, that effort will not be sustained.

REFERENCES


