

# *Draft Comparative Effectiveness Review*

---

Number XX

## **Radiotherapy Treatments for Head and Neck Cancer Update**

**Prepared for:**

Agency for Healthcare Research and Quality  
U.S. Department of Health and Human Services  
540 Gaither Road  
Rockville, MD 20850  
www.ahrq.gov

**Contract No. 290-2007-10058**

**Prepared by:**

Blue Cross and Blue Shield Association  
Evidence-based Practice Center  
Chicago, IL

**Investigators:**

Thomas A. Ratko, Ph.D.  
GW Douglas, M.S., M.P.H.  
Jonas DeSouza, M.D.  
Suzanne E. Belinson, M.P.H., Ph.D.  
Naomi A. Aronson, Ph.D.

**AHRQ Publication No.**  
**< Month Year >**

This report is based on research conducted by the Blue Cross and Blue Shield Association Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2007-10058). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, that is, in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the specific permission of copyright holders.

Persons using assistive technology may not be able to fully access information in this report. For assistance, contact [EffectiveHealthCare@ahrq.hhs.gov](mailto:EffectiveHealthCare@ahrq.hhs.gov).

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this report.

**Suggested citation:** Ratko TA, Douglas GW, DeSouza, J, Belinson SE, Aronson, NA. Radiotherapy Treatments for Head and Neck Cancer Update. Comparative Effectiveness Review No. xx. (Prepared by Blue Cross and Blue Shield Association Evidence-based Practice Center under Contract No. 290-2007-10058.) AHRQ Publication No. xx-EHCxxx-EF. Rockville, MD: Agency for Healthcare Research and Quality; <Month Year>. [www.effectivehealthcare.ahrq.gov](http://www.effectivehealthcare.ahrq.gov).

## Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. These reviews provide comprehensive, science-based information on common, costly medical conditions, and new health care technologies and strategies.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see [www.effectivehealthcare.ahrq.gov/reference/purpose.cfm](http://www.effectivehealthcare.ahrq.gov/reference/purpose.cfm).

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the Web site ([www.effectivehealthcare.ahrq.gov](http://www.effectivehealthcare.ahrq.gov)) to see draft research questions and reports or to join an email list to learn about new program products and opportunities for input.

We welcome comments on this systematic review. They may be sent by mail to the Task Order Officer (TOO) named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to [epc@ahrq.hhs.gov](mailto:epc@ahrq.hhs.gov).

Richard Kronick, Ph.D.  
Director  
Agency for Healthcare Research and Quality

Yen-Pin Chiang, Ph.D.  
Acting Director, Center for Outcomes and Evidence  
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.  
Director, EPC Program  
Center for Outcomes and Evidence  
Agency for Healthcare Research and Quality

William Lawrence, M.D.  
Task Order Officer  
Center for Outcomes and Evidence  
Agency for Healthcare Research and Quality

## **Acknowledgments**

### **Key Informants**

In preparing this update, the EPC did not consult Key Informants who represent the end-users of research. For CER No. 20, the EPC sought Key Informant input on the priority areas for research and synthesis. Key Informants were not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

## **Technical Expert Panel**

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who participated in developing this report follows:

## **Peer Reviewers**

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential nonfinancial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential nonfinancial conflicts of interest identified.

The list of Peer Reviewers follows:



# Radiotherapy Treatments for Head and Neck Cancer Update

## Structured Abstract

**Objectives.** This report is an update of a comparative effectiveness review (CER) published in final form in May 2010 on the benefits and harms of radiotherapy (RT) to treat patients with head and neck cancer (CER No. 20). RT modalities included conventional two-dimensional RT (2DRT), three-dimensional conformal RT (3DRT), intensity-modulated RT (IMRT) and proton-beam RT (PBRT)

In this CER update we included 3DRT, IMRT, PBRT, and stereotactic body RT (SBRT). Conventional 2DRT was not considered because it is considered obsolete for this setting. Brachytherapy was not included because its use is highly limited and specialized in head and neck cancer patients.

This review uses the same Key Questions as CER No. 20. The Key Questions asked whether any of these modalities is more effective than the others: (1) in reducing normal tissue toxicity and adverse events, and improving quality of life (QOL); (2) in improving local tumor control, time to disease progression, and survival; and (3) when used in certain anatomic locations or patient subpopulations; and, finally, whether (4) there is more variation in patient outcomes with any modality secondary to user experience, treatment planning, or target volumes.

The main finding of CER No. 20 was that late xerostomia was reduced and QOL domains related to xerostomia were improved in patients treated with IMRT compared with those who received either 3DRT or 2DRT. Evidence was insufficient to draw conclusions on survival or tumor control; adverse events other than late xerostomia (e.g., mucositis, dysphagia, skin toxicities, or osteoradionecrosis of the jaw); whether patient and tumor characteristics affected relative outcomes; or whether physician experience and treatment characteristics affected relative clinical outcomes such as survival or treatment-associated adverse events.

**Data sources.** A medical librarian searched MEDLINE<sup>®</sup>, EMBASE<sup>®</sup>, and the Cochrane Controlled Trials Registry for English-language articles. The overall search was performed for a period dating 12 months before the final literature search in CER No. 20 (September 28, 2009) through April 2013. For SBRT, the literature was searched for the period January 1, 1990, through April 2013. The search was updated at the time the draft CER was posted for review by the Agency for Healthcare Research and Quality (AHRQ). A search of the gray literature included clinical trial registries and information from manufacturers if available.

**Review methods.** We sought only comparative studies that reported clinical outcomes and QOL among our populations of interest. We found noncomparative studies to be uninformative in CER No. 20, so we excluded them from the update. Data were abstracted for each Key Question by one reviewer, with independent data verification. The study limitations of randomized controlled trials (RCTs) and other comparative studies were assessed using the United States

Preventive Services Task Force (USPSTF) criteria. The strength of the body of evidence for specific outcomes was assessed according to the latest AHRQ *Methods Guide*.

**Results.** We identified 6,661 unique titles and screened 262 in full text. Of the latter, nine (N=1,072) met the inclusion criteria, including one RCT (N=60). According to USPSTF criteria, the RCT was deemed fair quality, whereas the other studies were of poor quality. 3DRT and IMRT were compared in eight studies, including the RCT. One study compared 3DRT and SBRT; none compared IMRT and SBRT. As in CER No. 20, no evidence was identified on PBRT. Outcomes of therapy included overall survival, local control, adverse effects, and QOL.

**Conclusions.** New evidence on the comparative effectiveness of RT modalities for head and neck cancer is limited and heterogeneous for each comparison of 3DRT versus IMRT or SBRT. We did not identify any evidence for PBRT. New moderate strength evidence strengthens the CER No. 20 finding of reduced late xerostomia with IMRT compared with 3DRT, with no relative change in other conclusions on adverse events or QOL. New evidence was insufficient to draw conclusions about the relative effects of IMRT and 3DRT on overall survival or locoregional tumor control. New evidence is insufficient to draw conclusions on the comparative effectiveness of 3DRT or IMRT versus SBRT or PBRT.

# Contents

<b>Executive Summary</b> .....	<b>ES-1</b>
<b>Introduction</b> .....	<b>1</b>
Objectives .....	1
Epidemiology and Burden of Head and Neck Cancer .....	2
Overview of Multimodal Clinical Management of Head and Neck Cancer .....	2
RT in Head and Neck Cancer .....	3
Overview .....	3
Summary .....	5
Key Questions .....	5
PICOTS .....	6
Population(s).....	6
Interventions .....	7
Comparators .....	7
Outcomes .....	7
Timing .....	8
Settings .....	8
Analytic Framework .....	9
Organization of the Report .....	10
<b>Methods</b> .....	<b>11</b>
Overview .....	11
Study Inclusion Criteria .....	11
Literature Search .....	11
Search strategies .....	11
Data Abstraction and Data Management.....	12
Review of titles and abstracts .....	12
Full-text review .....	12
Data abstraction .....	12
Evidence tables .....	13
Assessment of Methodological Risk of Bias (Quality or Limitations) of Individual Studies .....	13
Data Synthesis .....	14
Grading the SOE for Individual Comparisons and Outcomes .....	14
Assessing Applicability .....	18
<b>Results</b> .....	<b>19</b>
Overview .....	19
Results of Literature Searches .....	19
Electronic Search .....	19

Grey Literature (Publication Bias) .....	20
Description of All Included Studies .....	21
Study Limitations .....	21
Study Design and Patient Characteristics .....	21
Key Question 1: Comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding adverse events and QOL .....	25
Overview .....	25
Key Points .....	30
Qualitative Synthesis .....	30
Strength of Evidence for Key Question 1 .....	34
Key Question 2: Comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding tumor control and patient survival .....	38
Overview .....	38
Key Points .....	41
Qualitative Synthesis .....	41
Strength of Evidence for Key Question 2 .....	43
Key Question 3: Comparative effectiveness of 3DRT, IMRT, SBRT, or PBRT for specific patient and tumor characteristics .....	47
Key Points .....	47
Key Question 4: Comparative effectiveness of 3DRT, IMRT, SBRT, or PBRT because of differences in user experience, treatment planning, treatment delivery, and target volume delineation .....	47
Key Points .....	47
<b>Discussion .....</b>	<b>48</b>
CER Update Strength of Evidence Relative to CER No. 20 Findings .....	48
Applicability of the Findings .....	51
Key Questions 1 and 2 .....	51
Key Questions 3 and 4 .....	52
Findings in Relationship to What Is Already Known .....	52
Limitations of Current Review and Evidence Base .....	53
Research Gaps .....	53
Lack of Clinical Trial Evidence on RT Interventions for Head and Neck Cancer .....	53
Potential Impediments to Comparative Studies of RT Interventions for Head and Neck Cancer .....	54
Summary and Conclusions .....	54
<b>References .....</b>	<b>56</b>
<b>Definition of Terms .....</b>	<b>59</b>
<b>Summary of Protocol Amendments .....</b>	<b>60</b>

## Tables

- Table 1. SOE rating domains: required and optional  
Table 2. Overall SOE categories and criteria for assignment  
Table 3. Design and characteristics of studies included in the CER No. 20 Update and CER No. 20  
Table 4. Summary of key reported acute (<90 days posttreatment) comparative toxicity outcomes  
Table 5. Summary of key reported late (>90 days posttreatment) comparative toxicity outcomes  
Table 6. Key Question 1: Evidence synthesis for key reported comparative grade 2 or higher toxicity outcomes  
Table 7. SOE for Key Question 1: Adverse effects and QOL  
Table 8. Summary of key reported comparative oncologic outcomes  
Table 9. Key Question 2: Evidence synthesis for key reported comparative oncologic outcomes  
Table 10. SOE for Key Question 2: Oncologic outcomes  
Table 11. Comparison of relevant CER No. 20 and update conclusions  
Table 12. Summary of applicability of evidence for Key Questions 1 and 2

### **Figures**

- Figure 1. Analytic framework for comparative effectiveness of RT for head and neck cancer  
Figure 2. PRISMA diagram for disposition of literature search results

### **Appendices**

- Appendix A. Search strategy  
Appendix B. USPSTF Study Quality Ratings  
    Table B-1. USPSTF study limitations ratings  
    Table B-2. Comparative study design and patient characteristics  
    Table B-3. Summary of radiotherapy techniques  
    Table B-4. Key acute (<90 days posttreatment) toxicities  
    Table B-5. Key late toxicities  
    Table B-6. Key oncologic outcomes  
Appendix C. Summary of reasons for study exclusion at the level of full-text review

# Executive Summary

## Introduction

### Objectives

In May 2010, the Agency for Healthcare Research and Quality (AHRQ) published the results of Comparative Effectiveness Review (CER) No. 20, *Comparative Effectiveness and Safety of Radiotherapy Treatments for Head and Neck Cancer*, prepared by the Blue Cross and Blue Shield Association Evidence-based Practice Center (EPC). In 2011, AHRQ published a surveillance report that used methods developed by the RAND and Ottawa EPCs to prioritize an update of AHRQ CER No. 20 in 2013. In preparing this update, we considered whether the interventions that we included in CER No. 20 remained applicable to current radiation oncology practice. In particular, we examined the applicability of conventional opposed beam 2DRT and brachytherapy in modern radiation oncology practice. Our conclusion, based on the current literature and input from our Technical Expert Panel (TEP) members, was that 2DRT is no longer in routine use in the U.S. for definitive treatment of head and neck cancer, thus we excluded it from the update.

Brachytherapy is an invasive technique that was the first form of radiotherapy (RT) in clinical use, dating back to 1901. Historically, it has been used extensively in many tumor types, including head and neck cancer. The primary advantage of brachytherapy over traditional opposed external beam two-dimensional radiotherapy (2DRT) has been its capability to conform a high, localized radiation dose to the implanted tumor, limiting exposure to noninvolved tissues. However, as conformal external beam RT methods (e.g., three-dimensional conformal RT [3DRT], intensity-modulated RT [IMRT]) have become more prevalent in the past 2 decades, this advantage of brachytherapy has been mitigated. Brachytherapy can be used in select head and neck cancer cases as a means of dose escalation in conjunction with external beam irradiation.<sup>1,2</sup> However, this practice has become uncommon because sufficient dose escalation can usually be achieved in these cases with a noninvasive approach (conformal RT). Brachytherapy alone is very rarely employed, except with small (T1) tumors of the nasal vestibule, lip, or oral cavity.<sup>3-7</sup> These presentations of head and neck cancers are relatively uncommon (1 percent to perhaps 5 percent of all cases), and RT is typically not first-line treatment in many cases. Therefore, because use of brachytherapy alone for primary management of head and neck malignancies has limited applicability in modern radiation oncology practice, we did not seek evidence of it for this current CER; we focused instead on RT modalities that are used as the sole RT intervention for a given presentation of head and neck cancer.

For this update, we reviewed and assessed new evidence on the comparative effectiveness of 3DRT, IMRT, and proton-beam RT (PBRT). We also systematically reviewed and assessed evidence on stereotactic body RT (SBRT), a newer RT modality that was not widely available when we prepared CER No. 20. This update used the same key questions as in CER No. 20 and, for the most part, the same methods and search strategies, modified to address the changes in the list of interventions. We organized clinical evidence according to treatment setting, abstracted only from comparative studies

(randomized or nonrandomized) of the conformal RT methods used in treatment for any head and neck cancer.

## **Epidemiology and Burden of Head and Neck Cancer**

Head and neck cancer is a heterogeneous disease characterized by complex clinical and pathologic presentations. Squamous cell carcinoma of the head and neck (SCCHN) constitutes approximately 90 percent of all head and neck cancers, and accounted for approximately 3 percent (about 50,000) of all new cancer cases and 2 percent (approximately 12,000) of all cancer deaths in 2010 in the U.S.<sup>8</sup> More than 600,000 people were diagnosed with SCCHN worldwide in 2008.<sup>8</sup>

## **Overview of Multimodal Clinical Management of Head and Neck Cancer**

Aggressive multimodality treatments with curative intent may include surgery, RT, and chemotherapy. RT is the mainstay of treatment, offered to nearly 75 percent of all head and neck cancer patients with either curative or palliative intent. RT may be used alone or as a part of multimodality approach, often with significant long-term side effects.

## **RT in Head and Neck Cancer**

### **Overview**

Conformal RT refers to modalities in which cytotoxic radiation beams are “shaped” to cover the tumor volume plus a surrounding tissue margin to treat microscopic disease that may reside there.

### **Conformal RT Modalities**

Here we briefly review important characteristics of each conformal RT modality considered in this CER update.

### **Three-Dimensional Conformal Radiotherapy**

3DRT allows for more accurate and precise dose calculations, with very rapid dose fall-off in surrounding tissues than with 2DRT because 3DRT takes into account axial anatomy and complex tissue contours.<sup>9</sup> 3D anatomic information from diagnostic computed tomography (CT) scans is used to deliver multiple highly focused beams of radiation that converge at the tumor site.

### **Intensity-Modulated Radiotherapy**

Compared with 3DRT, IMRT is a newer, more complex, and resource-intensive form of RT that delivers a high dose of ionizing radiation conformally to the target volume while sparing uninvolved, healthy tissues.<sup>9, 10</sup> By varying the beam intensity across shaped radiation fields, IMRT theoretically reduces radiation dose to organs at risk more than conventional RT.

## **Stereotactic Body Radiotherapy**

Stereotactic body RT (SBRT) delivers relatively large ablative doses of radiation in fewer treatment sessions than other conformal modalities.<sup>11</sup> Regimens generally comprise a total dose by definition in five or fewer fractions. The tumor location can be tracked in four dimensions (including time) using several CT imaging techniques that depend on the platform, tracking on bony structures or implanted fiducials.

## **Proton-Beam Radiotherapy**

Proton-beam RT (PBRT) has become increasingly available in the last few years. It has theoretical advantages over photon therapy because PBRT lacks an “exit dose,” potentially enabling physicians to deliver high-energy conformal doses to the tumor volume while almost completely sparing normal healthy tissue.

## **Summary**

The optimal means of delivering external beam ionizing radiation in sufficient doses to cure a patient with SCCHN requires a fine balance between treatment effectiveness and associated toxicity. A surveillance study prepared in 2011 by the Ottawa and RAND EPCs suggested rationale to update CER No. 20, based on signals of new evidence that would change several conclusions of that report. Taken together, the emergence of new technology and evidence suggesting potential differences between interventions in some outcomes prompted AHRQ to prioritize this update of CER No. 20.

## **Key Questions**

The following 4 key questions were addressed:

### **Key Question 1**

What is the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding adverse events and QOL [quality of life]?

### **Key Question 2**

What is the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding tumor control and patient survival?

### **Key Question 3**

Are there differences in the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT for specific patient and tumor characteristics?

### **Key Question 4**

Is there variation in the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT because of differences in user experience, treatment planning, treatment delivery, and target volume delineation?

## PICOTS

### Population(s)

#### Key Questions 1–4

Populations of interest included patients with head and neck cancer. To define what constitutes head and neck cancer, we consulted clinical resources such as the National Cancer Institute's (NCI) Physician Data Query Cancer Information Summary.<sup>12</sup> The definitions include:

- pharynx (hypopharynx, oropharynx, and nasopharynx)
- larynx
- lip and oral cavity
- paranasal sinus and nasal cavity
- salivary gland
- occult primary of the head and neck

All therapeutic strategies were included. RT can be delivered as a primary (curative) intent therapy or as an adjunct to surgery. We sought direct evidence for one intervention compared with another, with or without chemotherapy or surgery.

### Interventions

#### Key Questions 1–4

- 3DRT
- IMRT
- SBRT
- PBRT

Interventions may occur as part of a multimodal treatment strategy if the comparisons only differ with respect to the RT given.

### Comparators

#### Key Questions 1–4

All therapies were compared with each other as part of a continuum of treatment for patients with head and neck cancer.

### Outcomes

#### Key Questions 1, 3, and 4

**Final outcomes:** QOL and adverse events including: radiation-induced xerostomia and dysphagia.

**Intermediate outcomes:** Salivary flow and probability of completing treatment according to protocol.

We sought evidence related to user experience, treatment planning, and target volume delineation within the context of Key Question 4.

### **Key Questions 2–4**

**Final outcomes:** Overall survival and cancer-specific survival.

**Intermediate outcomes:** Local control and time to recurrence.

### **Timing**

All durations of followup were considered.

### **Settings**

Inpatient and outpatient.

## Analytic Framework

Figure A provides an analytic framework illustrating the population, interventions, outcomes, and adverse effects that guided the literature search and synthesis. It links the interventions of interest directly with final health outcomes (e.g., overall survival) and adverse events (e.g., xerostomia) as well as indirectly with final outcomes via intermediate outcomes (e.g., local control, disease-free survival).

**Figure A. Analytic framework for comparative effectiveness of RT for head and neck cancer**

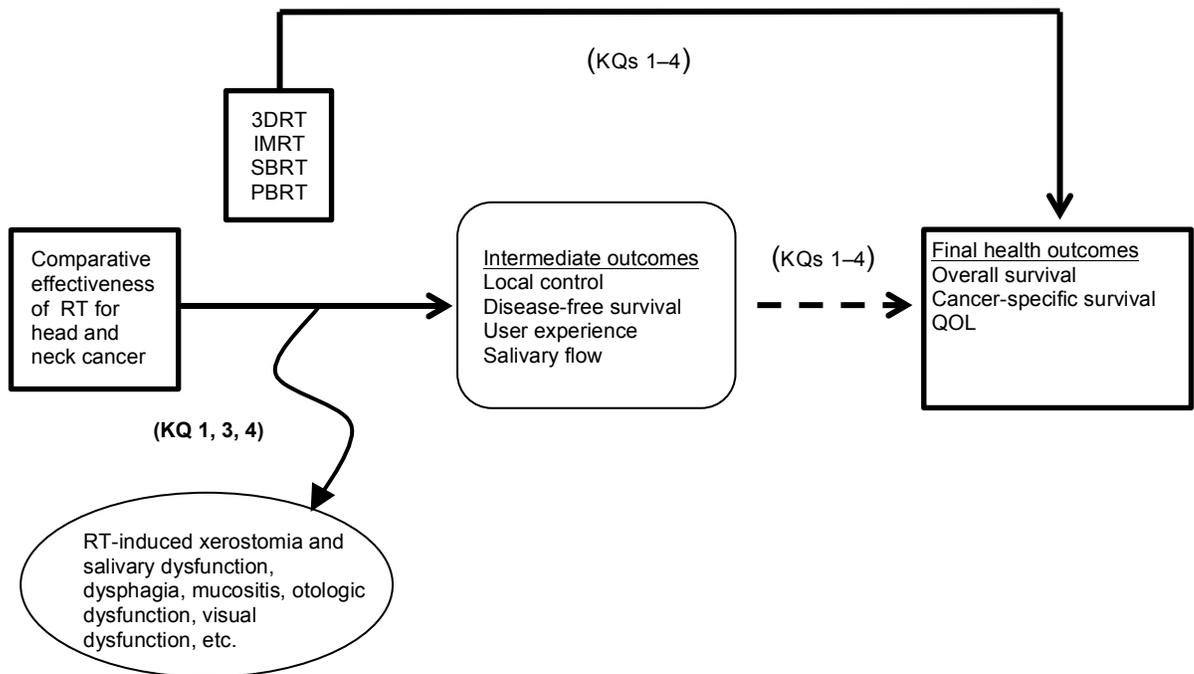


Figure A depicts the key questions (KQs) within the context of the PICOTS described in the previous section. In general, the figure illustrates how the interventions 3DRT, IMRT, SBRT, and PBRT may result in intermediate outcomes (e.g., local tumor control, disease-free survival) and long-term outcomes (e.g., overall survival, cancer-specific survival, QOL). Also, adverse events (e.g., radiation-associated xerostomia and salivary dysfunction, dysphagia, mucositis, otologic dysfunction, visual dysfunction) may occur at any point after the treatment is received.

# Methods

## Overview

This section describes the methods used to produce this CER update. The methodological practices we followed derived from the *Methods Guide* and its subsequent updates.<sup>13</sup> We also consulted the article published by Tsertsvadze et al. on methods to update CERs.<sup>14</sup>

## Study Inclusion Criteria

We included only full-length reports that describe the final results of randomized controlled trials (RCTs) and nonrandomized comparative studies (observational, case-control, and cohort studies) that meet the PICOTS criteria (see above).

## Literature Search

### Search strategies

An experienced medical librarian designed and performed all searches for this CER update. The literature search for the update was backdated to 12 months before the final literature search for CER No. 20 (dated September 28, 2009). For SBRT (and any other new interventions, we subsequently determined merited inclusion), the literature was searched electronically for citations from January 1, 1990, through April 2013. The search will be updated at the time the draft is posted for peer review by AHRQ. We searched the following databases:

- MEDLINE<sup>®</sup>
- EMBASE<sup>®</sup>
- Cochrane Controlled Trials Register

## **Data Abstraction and Data Management**

Literature search results were transferred to EndNote<sup>®</sup> and subsequently into Distiller for study screening.

### **Review of titles and abstracts**

We developed data collection forms for abstract review, full-text review, and data extraction. Two CER team members performed the initial title and abstract screen. To be excluded, a study must have been independently excluded by both team members.

### **Full-text review**

Full-text articles were reviewed against the PICOTS to determine their inclusion in the systematic review. The reason for excluding each article retrieved in full-text was recorded in the Distiller database.

### **Data abstraction**

We abstracted data into tables created in the Systematic Review Data Repository. Each article included was abstracted by a single reviewer. A second reviewer assessed the data extraction against the original articles for quality control.

The data elements abstracted included the following:

- Patient characteristics
- Treatment characteristics
- Outcome assessment (see PICOTS and Analytical Framework sections)

### **Evidence tables**

The same abstraction tables were used for all studies. The dimensions of each evidence table may vary by key question, but the tables contain common elements such as author, year of publication, sample size, study type, intervention(s), and comparator(s).

## **Assessment of Methodological Risk of Bias (Quality or Limitations) of Individual Studies**

In adherence to the *Methods Guide*,<sup>13</sup> the general approach to grading the quality or limitations of individual comparative studies was performed by using a United States Preventive Services Task Force (USPSTF) method.<sup>15</sup> Individual study quality assessment accounted for the following study elements:

- Number of participants and flow of participants through steps of study
- Treatment-allocation methods (including concealment)
- Use of blinding
- Study design (prospective vs. retrospective)
- Use of an independent outcome assessor

## Data Synthesis

The qualitative synthesis emphasized comparative studies sorted by specific head-to-head comparisons of interventions, specific treatment settings, patient characteristics, specific outcomes, and status relative the evidence hierarchy and study quality assessment.

## Grading the Strength of Evidence for Individual Comparisons and Outcomes

Studies were assessed for relevance against target populations, interventions of interest, and outcomes of interest. The system used for rating the strength of the overall body of evidence is outlined in the recently updated (2013) chapter from the *Methods Guide*<sup>13</sup> and is based on a system developed by the GRADE Working Group.<sup>16</sup> This system explicitly addresses the following domains: study limitations, directness, consistency, precision, and reporting bias.

The overall strength of evidence (SOE) grade is classified into four categories as shown in Table A

**Table A. Overall SOE categories and criteria for assignment**

Grade	Definition	Criteria for assignment
High	We are very confident that the estimate of effect lies close to the true effect for this outcome.	The body of evidence has few or no deficiencies. We believe that the findings are stable.
Moderate	We are moderately confident that the estimate of effect lies close to the true effect for this outcome.	The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	We have limited confidence that the estimate of effect lies close to the true effect for this outcome.	The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Insufficient	We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome.	No evidence is available or the body of evidence has unacceptable deficiencies, precluding judgment.

## Assessing Applicability

We assessed applicability of findings with the AHRQ Comparative Effectiveness *Methods Guide* using the PICOTS (Population, Intervention, Comparator, Outcome, Timing, Setting) framework.<sup>13, 17</sup> Included studies were assessed for relevance against target populations, interventions and comparators of interest, and outcomes of interest. We anticipated that results would be applicable only to the specialized populations of interest by key question.

# Results

## Overview

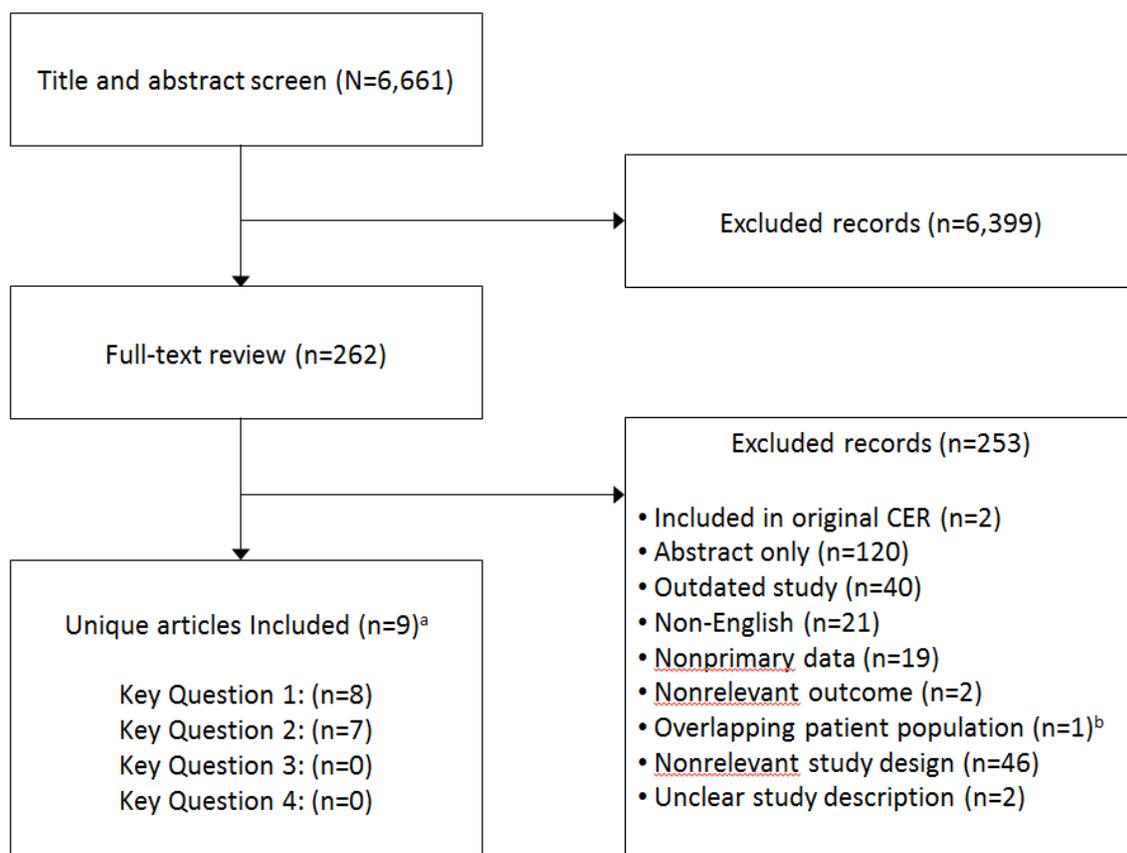
In this section, we first report our literature search results and PRISMA diagram, which depicts the flow of articles through the review according to our screening and inclusion criteria. We then provide an overview of the design, patients, and quality (risk of bias) of all included studies. We then lay out a qualitative synthesis of the evidence focusing on key outcomes related to CER No. 20.

## Results of Literature Searches

### Electronic Search

In our initial literature search for this update, we identified 6,661 unique titles and screened 262 in full text. Of the latter, nine (total N=1,072) met the CER inclusion criteria.<sup>18-26</sup> The flow of articles through the screening and study selection process is shown in the PRISMA diagram (Figure B).

**Figure B. PRISMA diagram for disposition of literature search results**



<sup>a</sup> Six studies addressed both Key Questions 1 and 2.

<sup>b</sup> Overlapping patient population refers to the studies in which the same patients were included in more than one study.

In this case, only one study was included to avoid oversampling. Decision to include a study was based on the clarity in reporting relevant patients and/or outcomes.

## **Grey Literature (Publication Bias)**

We did not include any information based on comprehensive searches of meeting abstracts. We examined the bibliographies of all papers screened in full text to identify peer-reviewed articles the electronic search may have missed.

We accessed the Web site ClinicalTrials.gov to identify ongoing phase 3 RCTs that would meet the criteria for inclusion based on our protocol. After a MEDLINE search of the NCT number(s) and title(s), we did not find any published results; it is unknown whether any data have been reported. At submission of this draft, we received Scientific Information Packets from one manufacturer of RT equipment. Information contained therein had no effect on our analysis.

## **Description of All Included Studies**

Nine studies met the inclusion criteria for this CER update. All are generally described in this section; other details and results specific to a particular key question are considered in the relevant subsections to follow.

## **Study Limitations**

We assigned a “fair” USPSTF rating to the RCT of Gupta primarily because the study was not double-blinded, particularly its outcomes assessments. The investigators did not report an intention-to-treat analysis, but this is moot because they reported a 97 percent followup rate in each of two study arms. Gupta reported aggregated survival results in patients with tumors in different sites. However, the distribution of tumor sites and characteristics between arms was similar. Overall, the two study arms were statistically similar and comparable.

The eight nonrandomized studies were retrospective database analyses, one of which used a historical comparator group. Overall, these eight studies were poorly designed, executed, and reported.

## **Study Design and Patient Characteristics**

In total, 3DRT and IMRT were compared in eight new studies, including one small (N=60) RCT.<sup>24</sup> One study compared 3DRT and SBRT<sup>26</sup>; none compared IMRT and SBRT. As in CER No. 20, no evidence was identified on PBRT.

Overall, similar to what we identified for CER No. 20, the body of studies in the update is heterogeneous in terms of tumor site and stage, treatment setting, and treatment intent (e.g., curative vs. palliative or recurrent). Patients were generally in their mid-fifties, with males predominating across studies. Tumor sites included the hypopharynx, larynx, nasal sinus, nasopharynx, oral cavity, and oropharynx. Four studies involved patients with single tumor sites. The majority of patients across studies had locally advanced (stage III and IV) cancer, although small proportions of patients had stage I or II disease.

The treatment settings included concurrent chemoradiotherapy (CCRT); RT with or without concurrent chemotherapy (CCT); CCRT with or without surgery; and adjuvant

postoperative RT. Where it appears in all tables throughout this report, the term RT ± CCT refers to treatment regimens in which all patients received RT, but not all received concurrent CT. This is distinct from the setting of CCRT, in which all patients were reported to have received RT and CT concurrently.

# **Key Question 1: Comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding adverse events and QOL**

## **Overview**

Acute and late toxicity outcomes were not collected consistently across studies. Four studies reported no acute toxicities,<sup>19, 21, 23, 26</sup> and three reported no late toxicities.<sup>21, 23, 26</sup> Only the study by Chen 2012 reported QOL evidence according to RT modality.<sup>21</sup>

Investigators did not adjust results to account for chemotherapy-associated toxicities independently of RT-associated toxicities, which complicates interpretation of toxicity evidence for many adverse events (e.g., mucositis). This is somewhat ameliorated by our focus on studies in which chemotherapy regimens are similar between study arms, thus potentially isolating the effect of the RT modality on such outcomes. However, we focused this update, as we did CER No. 20, on those grade 2 or higher toxicities associated with RT in the head and neck: dysphagia, salivary gland function, and xerostomia.

## **Key Points**

- New comparative evidence assessed in this update strengthens the conclusion from CER No. 20 that the risk of grade 2 or higher late xerostomia is significantly lower in patients treated with IMRT than 3DRT.
- Evidence remains insufficient to draw relative conclusions on adverse events other than late xerostomia (e.g., mucositis, dysphagia, skin toxicities, osteoradionecrosis of the jaw) that alter conclusions of CER No. 20.
- Posttreatment toxicities were reported inconsistently across studies, precluding comparisons within the body of evidence. We are uncertain whether the limited evidence on RT-associated toxicities overall reflects their absence or that the investigators did not systematically collect or report them.

## **Qualitative Synthesis**

In Table B and below, we summarize new evidence and the SOE related to Key Question 1 on toxicities actually reported in multiple studies according to the intervention comparison, treatment setting, and timeframe (acute vs. long-term).

### **RT-Associated Toxicities**

Three studies of IMRT compared with 3DRT in the setting of concurrent CRT showed statistically significant reduction in late xerostomia.<sup>20, 24, 25</sup> The rate of late xerostomia also was significantly lower with IMRT than 3DRT in single studies in the setting of RT with or without concurrent CT,<sup>18</sup> or postoperative RT,<sup>19</sup> respectively. The same set of studies reported evidence on acute and late dysphagia.

### **RT-Associated QOL**

One nonrandomized study reported QOL evidence on IMRT versus 3DRT in the setting of RT with or without concurrent CT. Chen et al. reported on mean QOL scores

using the University of Washington Quality of Life validated, self-administered tool.<sup>21</sup> In this study, the salivary gland domain was the only specific component of this score wherein significant differences were observed between the IMRT and 3DRT groups at both 1 and 2 years ( $p < 0.001$  at both points). Other domains (pain, appearance, activity, recreation, swallowing, chewing, speech, shoulder, taste, saliva, mood, anxiety) showed no differences according to RT modality. At 1 year after completion of RT, the global QOL was rated as “very good” or “outstanding” among 51 percent of patients treated with IMRT compared with 41 percent of those treated with 3DRT ( $p = 0.11$ ). However, at 2 years, the corresponding percentages were 73 percent and 49 percent, respectively ( $p < 0.001$ ), showing a benefit of IMRT. Multivariate analysis showed no effect on QOL scores of age, sex, radiation intent, radiation dose, T stage, primary site, or use of concurrent CT and neck dissection. The use of IMRT was the only variable associated with improved QOL ( $p < 0.01$ ).

**Table B. Key Question 1: Evidence synthesis for key reported comparative grade 2 or higher toxicity outcomes**

Comparison	Outcome	Timeframe	Number of studies (number of patients)	Individual study statistically significant results (p-value)	Study limitations (risk of bias)	Directness	Consistency	Precision	Overall SOE
3DRT vs. IMRT	Xerostomia	Late	Three studies <sup>20, 24, 25</sup> (N=509)	All three studies showed statistically significant benefit of IMRT vs. 3DRT.	Moderate  One “fair” quality small RCT (n=60, Gupta 2012) plus two “poor” quality non-randomized studies result in a “moderate” study limitations rating	Direct  All three studies directly compared IMRT and 3DRT.	Consistent  All three studies showed a statistically significant reduction of late grade >2 xerostomia with IMRT compared with 3DRT.	Precise	Moderate  The body of evidence comprises one “fair” quality RCT, for a provisional SOE of “high”. We downgraded the SOE one level based on the “moderate” risk of bias of the body of evidence. Although the Gupta trial was relatively small, its statistically significant result coupled with similar findings of the much larger non-randomized evidence merits an overall rating of precise.
	Dysphagia	Acute	Three studies <sup>20, 24, 25</sup> (N=509)	Only 1 study showed a statistically significant benefit of IMRT vs.	Moderate  One “fair” quality small RCT (n=60, Gupta 2012) plus two	Direct	Inconsistent  One study non-randomized study showed a statistically	Imprecise  The Gupta RCT only included 60 cases,	Insufficient  A “high” provisional SOE based on the Gupta RCT was reduced three

				3DRT <sup>20</sup>	“poor” quality non-randomized studies result in a “moderate” study limitations rating.		significant reduction with IMRT compared with 3DRT. <sup>20</sup> The other non-RCT showed a directionally same but nonsignificant effect that favored IMRT over 3DRT. Gupta 2012 showed a lower but also nonsignificant rate difference of acute dysphagia with 3DRT compared with IMRT.	compared with 449 for the other 2 studies. It was likely not sufficiently powered to detect slight changes in rates of adverse effects, particularly in the face of much larger, albeit “poor” quality non-RCT evidence.	levels for three reasons: (1) inconsistent rating; (2) imprecise rating based on the small size of the Gupta RCT and its nonsignificant result; and (3) the two nonrandomized studies were of “poor” quality, heterogeneous, and subject to a high risk of bias, thus increasing the risk of bias to “moderate” for the body of evidence.
	Late	Two studies <sup>20, 25</sup> (N=707)	Only 1 study showed a statistically significant benefit of IMRT vs. 3DRT (grade $\geq 2$ ) <sup>20</sup>	High	Two “poor” quality, nonrandomized studies comprise the body of evidence.	Direct	Inconsistent	Precise	Insufficient
							One study showed a statistically significant effect of IMRT compared with 3DRT, with the second study showing a reduction, albeit nonsignificant reduction.		The two nonrandomized studies were “poor” quality and heterogeneous, with high risk of bias that compromises the value of their results.

CCRT = concurrent chemoradiotherapy; CCT = concurrent chemotherapy; IMRT = intensity-modulated radiotherapy; RT = radiotherapy; 3DRT = three-dimensional conformal radiotherapy.

## **Key Question 2: Comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding tumor control and patient survival**

### **Overview**

Key oncologic outcomes were not reported consistently across studies, Not all outcomes were collected in each study. Three studies, including the Gupta RCT, reported data on overall survival, local control, or locoregional control among patients treated with IMRT compared with 3DRT in the setting of concurrent CRT.<sup>20, 24, 25</sup>

### **Key Points**

- As we found in CER No. 20, comparative evidence assessed in this update was insufficient to draw relative conclusions on any oncologic outcomes.
- The key oncologic outcomes were not reported universally across studies, so we could not make comparisons across a larger body of evidence.

### **Qualitative Synthesis**

In Table C, we summarize new evidence and the SOE related to Key Question 2 on oncologic outcomes actually reported in multiple studies.

In general, evidence on tumor control and survival outcomes is sparse. No statistically significant differences were reported for overall survival, local control, or locoregional control among studies of 3DRT versus IMRT in any setting compiled there. The only statistically significant oncologic result we found was in disease-free survival with IMRT compared with 3DRT in the postoperative adjuvant setting for paranasal sinus cancer (72 percent vs. 60 percent,  $p=0.02$ ).

**Table C. Key Question 2: Evidence synthesis for key reported comparative oncologic outcomes**

Comproison	Outcome	Number of studies (number of patients)	Individual study statistically significant results (p-value)	Risk of bias	Directness	Consistency	Precision	Overall SOE
3DRT vs. IMRT	Overall survival	Three studies <sup>18, 24, 25</sup> (N=509)	No statistically significant difference in overall survival was reported in any study.	Moderate  One “fair” quality small RCT (n = 60, Gupta 2012) plus two “poor” quality non-randomized studies result in a “moderate” study limitations rating.	Direct  All three studies directly compared IMRT and 3DRT.	Consistent  All three studies showed no statistically significant difference between 3DRT and IMRT in rate of overall survival at 2 or 5 years.	Imprecise  The Gupta 2012 RCT was likely not sufficiently powered to detect slight changes in rates of overall survival with IMRT compared with 3DRT, particularly in the face of much larger, albeit “poor” quality non-RCT evidence.	Insufficient  A “high” provisional SOE based on the Gupta RCT was reduced three levels for three reasons: (1) imprecise rating based on the small size of the Gupta RCT and its nonsignificant result; (2) the two nonrandomized studies were of “poor” quality, heterogeneous, and subject to a high risk of bias, yielding an overall “moderate” risk of bias; and (3) the relative larger size of these 2 studies compared to Gupta, accounting for 88% of all patients in the body of evidence, obscure the findings of the latter, resulting in an overall SOE rating of “insufficient”.
	Locoregional control	Two studies <sup>24, 25</sup> (N=305)	No statistically significant difference in locoregional control was reported in	Moderate  One “fair” quality RCT (Gupta 2012) and a much	Direct  Both studies directly compared IMRT and	Consistent  Both studies showed no statistically significant	Imprecise  The Gupta 2012 RCT is was likely not sufficiently powered to detect	Insufficient  A “high” provisional SOE based on the Gupta RCT was reduced three SOE levels basically as outlined

			either study.	larger “poor” quality non-randomized study result in a “moderate” study limitations rating.	3DRT.	difference between 3DRT and IMRT in rate of overall survival at 2 or 5 years.	slight changes in rates of locoregional control with IMRT compared with 3DRT, particularly in the face of much larger, albeit “poor” quality non-RCT evidence.	above for overall survival. Note the patients in the nonrandomized study comprised more than 80% of the evidence base, obscuring Gupta’s results.
--	--	--	---------------	---	-------	---	--	---

CCRT = concurrent chemoradiotherapy; IMRT: intensity-modulated radiotherapy; RT: radiotherapy; 3DRT: three-dimensional conformal radiotherapy.

### **Key Question 3: Comparative effectiveness of 3DRT, IMRT, SBRT, or PBRT for specific patient and tumor characteristics**

#### **Key Points**

- In CER No. 20, no comparative studies addressed these issues. In this update, we did not identify any new evidence that specifically addressed Key Question 3.

### **Key Question 4: Comparative effectiveness of 3DRT, IMRT, SBRT, or PBRT because of differences in user experience, treatment planning, treatment delivery, and target volume delineation**

#### **Key Points**

- In CER No. 20, no comparative studies addressed these issues. In this update, we did not identify any new evidence that specifically addressed Key Question 4.

## Discussion

### Strength of Evidence Relative to CER No. 20

Table D provides a summary of the conclusions we drew for the relevant interventional comparisons for each key question in CER No. 20 and in this update. Because 2DRT and SBRT are not commonly addressed in CER No. 20 and the update, they are not included in Table D. Moderate strength evidence from the update shows a reduction of the incidence of late grade 2 or higher xerostomia with IMRT compared with 3DRT, which strengthens the conclusion on this toxicity and comparison from CER No. 20. Evidence in the update is insufficient to show a difference between IMRT and 3DRT in overall survival or locoregional tumor control rates. We found no new evidence to alter any conclusions of CER No. 20 for any other toxicity or oncologic outcomes or comparisons.

**Table D. Comparison of relevant CER No. 20 and update conclusions**

Key question	Comparison	Clinical outcome	CER No. 20 total evidence base	CER No. 20 conclusions	CER No. 20 update total evidence base	CER No. 20 update conclusions	Cumulative update conclusions (action needed)
Key Question 1: What is the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding adverse events and QOL?	3DRT vs. IMRT	Grade $\geq 2$ late xerostomia	One good quality RCT and six poor quality non-RCTs	Moderate SOE shows significant reduction in incidence	One fair quality RCT, two poor quality non-RCTs	Moderate SOE shows significant reduction in incidence	Raises SOE to “high” (no further study required)
		Other RT-associated grade $>2$ toxicities (e.g., acute or late dysphagia, salivary gland dysfunction, swallowing function)	Variously, one good quality RCT, 13 poor quality non-RCTs	Insufficient evidence to draw conclusions	Variously, one good quality RCT, eight poor quality non-RCTs	Insufficient evidence to draw conclusions	Insufficient evidence to draw conclusions (further study required)
		QOL	Three poor quality non-RCTs				
	3DRT or IMRT vs. PBRT	Any	No evidence identified	No evidence identified, insufficient	No evidence identified	No evidence identified, insufficient	
Key Question 2: What is the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding tumor control and patient survival?	3DRT vs. IMRT	Overall survival, locoregional control	Variously, one good quality RCT, six poor quality non-RCTs	Insufficient evidence to draw conclusions	One fair quality RCT, three poor quality non-RCTs	Insufficient evidence to draw conclusions	
	3DRT or IMRT vs. PBRT	Any	No evidence identified	No evidence identified, insufficient	No evidence identified	No evidence identified, insufficient	
Key Question 3: Are there differences in comparative	3DRT or IMRT vs. PBRT	Any	No evidence identified	No evidence identified,	No evidence identified	No evidence identified,	

effectiveness of 3DRT, IMRT, SBRT, and PBRT for specific patient and tumor characteristics?				insufficient		insufficient	
Key Question 4: Is there variation in comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT because of differences in user experience, treatment planning, treatment delivery, and target volume delineation?	3DRT or IMRT vs. PBRT	Any	No evidence identified	No evidence identified, insufficient	No evidence identified	No evidence identified, insufficient	

CCT = concurrent chemotherapy; CCRT = concurrent chemoradiotherapy; CER = Comparative Effectiveness Review; CT = computed tomography; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; QOL = quality of life; RCT = randomized controlled trial; RT = radiotherapy; SBRT = stereotactic body radiotherapy; SOE = strength of evidence; 3DRT = three-dimensional conformal radiotherapy.

## **Applicability of the Findings**

In general, applicability assessment would depend on a body of evidence sufficient to form new conclusions about the comparative outcomes of 3DRT, IMRT, SBRT, and PBRT in treatment of head and neck cancer. However, comparative evidence that meets study selection criteria for this CER update is sparse for 3DRT, IMRT, and SBRT, and nonexistent for PBRT. In the absence of sufficient evidence, additional factors may be considered in making a treatment decision. Those could include relative convenience and cost, issues outside the scope of this CER.

In preparing this update, we considered the interventions that we included in CER No. 20 and whether all remained applicable to current radiation oncology practice. In particular, we examined the role of conventional opposed beam 2DRT and brachytherapy in modern radiation oncology practice. Our conclusion, based on the current literature and input from our TEP members, was that 2DRT is no longer in use in the U.S. for definitive treatment of head and neck cancer, thus we excluded it from the update. Further, although brachytherapy can be used in select cases as a means of dose-escalation in conjunction with external beam irradiation for head and neck cancer<sup>1,2</sup> this practice has become uncommon because sufficient dose escalation can often be achieved in these cases with a noninvasive approach (e.g., conformal RT). Brachytherapy alone is very rarely employed, except in small (T1) tumors of the nasal vestibule, lip, or oral cavity, which are relatively uncommon (1 percent to perhaps 5 percent of all cases).<sup>3-7</sup> Therefore, because use of brachytherapy alone for primary management of head and neck malignancies has limited applicability in modern head and neck radiation oncology practice, we did not seek evidence of it for this CER; we focused instead on RT modalities that are used as the sole RT intervention for a given presentation of head and neck cancer.

## **Key Questions 1 and 2**

The degree to which the evidence presented in this report is applicable to clinical practice is a function of the similarity between populations in the included studies and the patient population that receives clinical care in diverse settings. It also is related to the relative availability of the interventions. Because of the overall weakness of evidence for Key Questions 1 and 2, we have primarily limited comments to the relevance of the PICOTS elements, a practical and useful structure to review the applicability in a systematic manner (Table E).

**Table E. Summary of applicability of evidence for Key Questions 1 and 2**

PICOTS Domain	Applicability of evidence
Populations	<ul style="list-style-type: none"> <li>Overall patients included in the evidence base of this CER update are typical of the head and neck cancer population treated with RT based on age, sex, and tumor characteristics.</li> </ul>
Interventions	<ul style="list-style-type: none"> <li>3DRT, IMRT, and SBRT represent different technological approaches to the delivery of conformal photon RT. The major advantage of these interventions compared with traditional wide-field 2DRT is the ability to deliver tightly focused cytotoxic radiation by delineating the shape and size of the tumor using a CT-based or other imaging planning system.</li> <li>3DRT represents a minimum technical standard for delivery of conformal RT. It involves static fields with a fixed shape, modified by compensators (wedges and segments). 3DRT is widely available.</li> <li>IMRT offers beam strength attenuation through a multileaf collimator (tungsten), with dynamic field shapes for each beam angle. IMRT is not as widely available as 3DRT and requires a higher level of inverse planning effort and quality assurance.</li> <li>SBRT is a hypofractionated technique administered in five or fewer fractions; 3DRT and IMRT typically deliver radiation in many more fractions than SBRT.</li> <li>SBRT is not as widely available as 3DRT or IMRT, but its use is growing in other settings such as non-small-cell cancer. The institutional programmatic requirements for SBRT are similar to those for IMRT.</li> <li>Comparative evidence for PBRT is unavailable.</li> </ul>
Comparators	<ul style="list-style-type: none"> <li>See above for Interventions.</li> </ul>
Outcomes	<ul style="list-style-type: none"> <li>The major beneficial health outcomes in this CER are overall survival and late toxicities, particularly xerostomia.</li> <li>Overall survival is the primary outcome of interest for any cancer intervention study.</li> <li>Local control is of interest to patients because it measures the effectiveness of an intervention in disease control. On local failure, patients enter into a new category centered on systemic chemotherapy. This is a perilous position for typically medically frail patients.</li> </ul>
Timing	<ul style="list-style-type: none"> <li>The relevant periods occur from the time of treatment through followup over months (palliation) or years (overall survival).</li> </ul>
Setting	<ul style="list-style-type: none"> <li>The evidence for Key Questions 1 and 2 is mostly international, primarily obtained in tertiary institutional settings. More sophisticated interventions such as IMRT and SBRT require an institutional commitment to quality assurance and ongoing training that may be difficult to achieve in smaller community-based centers.</li> <li>We did not collect or analyze information to examine these issues.</li> </ul>

CCT = concurrent chemotherapy; CCRT = concurrent chemoradiotherapy; CER = Comparative Effectiveness Review; CT = computed tomography; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; RCT = randomized controlled trial; RT = radiotherapy; SBRT = stereotactic body radiotherapy; 3DRT = three-dimensional conformal radiotherapy.

## Key Questions 3 and 4

The current evidence base for Key Questions 3 and 4 is nonexistent based on our literature review. Therefore we cannot assess the applicability to clinical practice.

## Findings in Relationship to What Is Already Known

Our updated systematic literature search and review revealed no relevant evidence-based guidelines we could compare with our findings for any of the key questions.

## **Limitations of Current Review and Evidence Base**

Although the body of evidence we identified was more substantial for 3DRT and IMRT than SBRT, and nonexistent for PBRT, we have significant concerns about interstudy heterogeneity, with variability in RT dose, schedule of treatment, concurrent treatments, patient selection criteria, tumor size and location, and so forth.

## **Research Gaps**

The primary research gap we identified is a continuing lack of evidence from well-executed comparative studies (randomized or otherwise) to draw conclusions on the relative clinical benefits and harms of the RT interventions used in patients with head and neck cancer. We also identified some potential impediments to the type of rigorous comparative studies we suggest are necessary to determine their comparative effectiveness. We urge that rigorous methods be used for the conduct of RCTs, particularly intention-to-treat analysis and adjustment of survival data to account for all patients based on their treatment plans.

Primary outcomes would include overall survival, cancer-specific survival, and local control. Pre-specified systematic collection of adverse events using validated criteria (e.g., CTCAE) is necessary to permit accurate assessment of relative benefits and risks of the interventions.

As we allude to in the Introduction of this report, the potential impact of tumor tissue human papillomavirus positivity on oncologic outcomes and management of such patients has been increasing in importance. Studies are needed to identify reduced intensity therapies that still yield satisfactory oncologic outcomes.

## **Potential Impediments to Comparative Studies of RT Interventions for Head and Neck Cancer**

The general dissemination of conformal RT technologies into community clinical practice is a potential impediment to comparative study of those technologies. We acknowledge that randomized studies of 3DRT versus IMRT or PBRT may be very difficult to recruit and conduct, based on technical and potential ethical issues related to perceptions of unequal clinical benefit among the interventions.

## **Summary and Conclusions**

The main finding of CER No. 20 was that late xerostomia was reduced and QOL domains related to xerostomia were improved in patients treated with IMRT compared with those who received either 3DRT or 2DRT. Evidence was insufficient to draw relative conclusions on survival or tumor control; adverse events other than late xerostomia (e.g., mucositis, dysphagia, skin toxicities, osteoradionecrosis of the jaw); whether patient and tumor characteristics affected relative outcomes; or whether physician experience and treatment characteristics affected relative clinical outcomes such as survival or treatment-associated adverse events.

Moderate strength evidence from the update shows a reduction of the incidence of late grade 2 or higher xerostomia with IMRT compared with 3DRT, which strengthens the conclusion on this toxicity and comparison from CER No. 20. Evidence in the update is insufficient to show a difference between IMRT and 3DRT in overall survival or locoregional tumor control rates. We found no new evidence to alter any conclusions of CER No. 20 for any other toxicity or oncologic outcomes or comparisons.

## Cited References

1. Erickson BA, Wilson JF. Nasopharyngeal brachytherapy. *Am J Clin Oncol*. 1993 Oct;16(5):424-43. PMID: 8213626.
2. Esche BA, Haie CM, Gerbaulet AP, et al. Interstitial and external radiotherapy in carcinoma of the soft palate and uvula. *Int J Radiat Oncol Biol Phys*. 1988 Sep;15(3):619-25. PMID: 3138213.
3. Jorgensen K, Elbrond O, Andersen AP. Carcinoma of the lip. A series of 869 cases. *Acta Radiol Ther Phys Biol*. 1973 Jun;12(3):177-90. PMID: 4749173.
4. Marchese MJ, Nori D, Anderson LL, et al. A versatile permanent planar implant technique utilizing iodine-125 seeds imbedded in gelfoam. *Int J Radiat Oncol Biol Phys*. 1984 May;10(5):747-51. PMID: 6735761.
5. Marcus RB, Jr., Million RR, Mitchell TP. A preloaded, custom-designed implantation device for stage T1-T2 carcinoma of the floor of mouth. *Int J Radiat Oncol Biol Phys*. 1980 Jan;6(1):111-3. PMID: 7364600.
6. Mendenhall NP, Parsons JT, Cassisi NJ, et al. Carcinoma of the nasal vestibule. *Int J Radiat Oncol Biol Phys*. 1984 May;10(5):627-37. PMID: 6429094.
7. Mendenhall NP, Parsons JT, Cassisi NJ, et al. Carcinoma of the nasal vestibule treated with radiation therapy. *Laryngoscope*. 1987 May;97(5):626-32. PMID: 3573913.
8. Jemal A, Bray F, Center MM, et al. Global cancer statistics. *CA: A Cancer Journal for Clinicians*. 2011;61(2):69-90.
9. Burri M, Bevan A, Roach III M. Advances in radiation therapy: conventional to 3D, to IMRT, to 4D, and beyond. *CA Cancer J Clin*. 2005;55:117-34.
10. Lee N, Puri DR, Blanco AI, et al. Intensity-modulated radiation therapy in head and neck cancers: An update. *Head & Neck*. 2007;29(4):387-400.
11. Siddiqui F, Raben D, Lu JJ, et al. Emerging applications of stereotactic body radiation therapy for head and neck cancer. *Expert Rev Anticancer Ther*. 2011 Sep;11(9):1429-36. PMID: 21929316.
12. Pfister DG, Ang K-K, Brizel DM, et al. Head and Neck Cancers. *Journal of the National Comprehensive Cancer Network*. 2011 June 1, 2011;9(6):596-650.
13. Helfand M, Balshem H. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. August 2009. Rockville, MD. <http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayProduct&productID=318>.
14. Tsertsvadze A, Maglione M, Chou R, et al. Updating comparative effectiveness reviews: current efforts in AHRQ's Effective Health Care Program. *J Clin Epidemiol*. 2011 Nov;64(11):1208-15. PMID: 21684114.
15. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med*. 2001;20(3 Suppl):21-35. PMID: 11306229
16. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—Agency for Healthcare Research and Quality and the Effective Health-Care Program. *J Clin Epidemiol*. 2010;63:513-23. PMID: 19595577
17. Brown P, Brunnhuber K, Chalkidou K, et al. How to formulate research recommendations. *BMJ*. 2006 Oct 14;333(7572):804-6. PMID: 17038740.
18. Al-Mamgani A, Mehilal R, van Rooij PH, et al. Toxicity, quality of life, and functional outcomes of 176 hypopharyngeal cancer patients treated by (chemo)radiation: the impact of treatment modality and radiation technique. *Laryngoscope*. 2012 Aug;122(8):1789-95. PMID: 22833307.

19. Al-Mamgani A, Monserez D, Rooij P, et al. Highly-conformal intensity-modulated radiotherapy reduced toxicity without jeopardizing outcome in patients with paranasal sinus cancer treated by surgery and radiotherapy or (chemo)radiation. *Oral Oncol.* 2012 Sep;48(9):905-11. PMID: 22584070.
20. Al-Mamgani A, van Rooij P, Verduijn GM, et al. The impact of treatment modality and radiation technique on outcomes and toxicity of patients with locally advanced oropharyngeal cancer. *Laryngoscope.* 2013 Feb;123(2):386-93. PMID: 23404489.
21. Chen AM, Farwell DG, Luu Q, et al. Intensity-modulated radiotherapy is associated with improved global quality of life among long-term survivors of head-and-neck cancer. *Int J Radiat Oncol Biol Phys.* 2012 Sep 1;84(1):170-5. PMID: 22300572.
22. Dirix P, Vanstraelen B, Jorissen M, et al. Intensity-modulated radiotherapy for sinonasal cancer: Improved outcome compared to conventional radiotherapy. *International Journal of Radiation Oncology Biology Physics.* 2010;78(4):998-1004.
23. Guan X, Wang X, Liu Y, et al. Lymph node metastasis in sinonasal squamous cell carcinoma treated with IMRT/3D-CRT. *Oral Oncology.* 2013;49(1):60-5.
24. Gupta T, Agarwal J, Jain S, et al. Three-dimensional conformal radiotherapy (3D-CRT) versus intensity modulated radiation therapy (IMRT) in squamous cell carcinoma of the head and neck: a randomized controlled trial. *Radiother Oncol.* 2012 Sep;104(3):343-8. PMID: 22853852.
25. Lambrecht M, Nevens D, Nuyts S. Intensity-modulated radiotherapy vs. parotid-sparing 3D conformal radiotherapy. Effect on outcome and toxicity in locally advanced head and neck cancer. *Strahlenther Onkol.* 2013 Mar;189(3):223-9. PMID: 23319256.
26. Ozyigit G, Cengiz M, Yazici G, et al. A retrospective comparison of robotic stereotactic body radiotherapy and three-dimensional conformal radiotherapy for the reirradiation of locally recurrent nasopharyngeal carcinoma. *International Journal of Radiation Oncology Biology Physics.* 2011;81(4):e263-e8.

# Introduction

## Objectives

In May 2010, the Agency for Healthcare Research and Quality (AHRQ) published the results of Comparative Effectiveness Review (CER) No. 20, *Comparative Effectiveness and Safety of Radiotherapy Treatments for Head and Neck Cancer*, prepared by the Blue Cross and Blue Shield Association Evidence-based Practice Center (EPC). CER No. 20 examined evidence on clinical outcomes achieved with conventional or two-dimensional radiotherapy (2DRT), three-dimensional conformal RT (3DRT), intensity-modulated RT (IMRT), and proton-beam RT (PBRT). The main finding of CER No. 20 was that late xerostomia was reduced and quality of life (QOL) domains related to xerostomia were improved in patients treated with IMRT compared with those who underwent 3DRT or 2DRT. Evidence was insufficient to draw conclusions on overall survival or tumor control; adverse events other than late xerostomia (e.g., mucositis, dysphagia, skin toxicities, or osteoradionecrosis of the jaw); whether patient and tumor characteristics affected relative outcomes; or whether physician experience and treatment characteristics affected relative clinical outcomes such as survival or treatment-associated adverse events.

In 2011, AHRQ published a surveillance report that used methods developed by the RAND and Ottawa EPCs to prioritize an update of AHRQ CER No. 20 in 2013. In preparing this update, we considered whether the interventions included in CER No. 20 remained applicable to current radiation oncology practice. In particular, we examined the applicability of conventional opposed beam 2DRT and brachytherapy in modern radiation oncology practice. Our conclusion, based on the current literature and input from our Technical Expert Panel (TEP) members, was that 2DRT is no longer in routine use in the U.S. for definitive treatment of head and neck cancer, thus we excluded it from the update.

Brachytherapy is an invasive technique that was the first form of RT in clinical use, dating back to 1901. Historically, it has been used extensively in many tumor types, including head and neck cancer. The primary advantage of brachytherapy over traditional opposed external beam 2DRT has been its capability to conform a high, localized radiation dose to the implanted tumor, limiting exposure to noninvolved tissues. However, as conformal external beam RT methods (e.g., 3DRT and IMRT) have become more prevalent in the past 2 decades, this advantage of brachytherapy has been mitigated. Brachytherapy can be used in select head and neck cancer cases as a means of dose escalation in conjunction with external beam irradiation.<sup>1,2</sup> However, this practice has become uncommon because sufficient dose escalation can usually be achieved in these cases with a noninvasive approach (i.e., conformal RT). Brachytherapy alone is very rarely employed, except with small (T1) tumors of the nasal vestibule, lip, or oral cavity.<sup>3-7</sup> These presentations of head and neck cancers are relatively uncommon (1 percent to perhaps 5 percent of all cases), and RT is typically not first-line treatment in many cases. Therefore, because use of brachytherapy alone for primary management of head and neck malignancies has limited applicability in modern radiation oncology practice, we did not seek evidence of it for this current CER; we focused instead on RT

modalities that are used as the sole RT intervention for a given presentation of head and neck cancer.

For this update, we reviewed and assessed new evidence on the comparative effectiveness of 3DRT, IMRT, and PBRT. We also systematically reviewed and assessed evidence on stereotactic body RT (SBRT), a newer RT modality that was not widely available when we prepared CER No. 20. This update used the same key questions as in CER No. 20 and, for the most part, the same methods and search strategies, modified to address the changes in the list of interventions. We organized clinical evidence according to treatment setting, abstracted only from comparative studies (randomized or nonrandomized) of the conformal RT methods used in treatment for any head and neck cancer.

## **Epidemiology and Burden of Head and Neck Cancer**

Head and neck cancer is a heterogeneous disease characterized by complex clinical and pathologic presentations. Squamous cell carcinoma of the head and neck (SCCHN) specifically arises in the squamous epithelium of the upper aerodigestive tract (oral cavity, larynx, hypopharynx, oropharynx, nasopharynx, paranasal sinuses/nasal cavity). SCCHN constitutes approximately 90 percent of all head and neck cancers, and accounted for approximately 3 percent (about 50,000) of all new cancer cases and 2 percent (approximately 12,000) of all cancer deaths in 2010 in the U.S.<sup>8</sup> While these cancers in total comprise a relatively small percentage of all cancers, cumulatively they are the sixth most common cancer worldwide, with notable exceptions of high nasopharyngeal cancer incidence in South Eastern China and South Eastern Asia and high oral cavity cancer incidence in Melanesia and South Central Asia. More than 600,000 people were diagnosed with SCCHN worldwide in 2008.<sup>8</sup>

Major risk factors for the development of head and neck cancer include tobacco and alcohol abuse, with other less common risk factors including occupational exposures, nutritional deficiencies, and poor oral health.<sup>9</sup> Viral etiologies have also been established, with human papillomavirus (HPV) infection appearing to be a risk factor, particularly within the oropharynx, in younger people without a history of tobacco or alcohol abuse. The reported proportion of oropharyngeal cancers attributable to HPV in the U.S. has increased from 16.3 percent during the 1980s to 72.7 percent during the 2000s.<sup>10, 11</sup> Careful anatomic site stratification has shown that the age-adjusted incidence of oropharyngeal cancer is rising dramatically (estimated to be a 5 percent annual increase). In addition to HPV, an association has been made between Epstein-Barr virus and nasopharyngeal cancer.

## **Overview of Multimodal Clinical Management of Head and Neck Cancer**

Most patients with SCCHN present with locally advanced but curable disease; only a small percentage of these patients have demonstrable distant metastases. Treatment decisions are primarily determined by the size, location, and grade of the primary tumor; the extent of nodal involvement; and the estimated functional impact of therapy. Patient characteristics may include substantial comorbidities and poor performance status that must also be considered in devising a comprehensive treatment plan.<sup>9</sup>

Aggressive multimodality treatments with curative intent may include surgery, RT, and chemotherapy. RT is the mainstay of treatment, offered to nearly 75 percent of all head and neck cancer patients with either curative or palliative intent. RT may be used alone or as a part of multimodality approach, often with significant long-term side effects. In planning this CER, we sought to account for multimodal treatment strategies by organizing evidence according to treatment settings used in comparative studies of the RT approaches. RT-associated toxicities represent important clinical outcomes that can substantially reduce QOL and the ability of cancer patients to tolerate and complete the entire planned course of treatment.

The main challenge in RT for any type of cancer is to attain the highest probability of tumor control or cure with the least amount of morbidity and toxicity. However, improved outcomes with aggressive RT regimes come at the cost of increased treatment toxicity, mainly due to the close proximity of critical organs and the often large irradiation fields necessary to effect local tumor control in head and neck cancer patients. For example, xerostomia is the most prevalent toxicity of RT to the head and neck and a major cause of reduced QOL. In addition to patient perception of mouth dryness, it leads to impaired speech and swallow function, all of which also contribute to decreased QOL. Other prominent, related RT-associated toxicities include salivary gland dysfunction, accelerated dental caries and osteoradionecrosis.

Although RT-associated toxicities are highly problematic in any patient with head and neck cancer, such adverse events are considered to assume greater importance in patients identified with HPV compared with those with HPV-negative disease.<sup>11</sup> Patients with HPV-positive oropharynx cancer not only appear to have a different clinical phenotype from HPV-negative cancers, but they also have had better outcomes in multiple large studies, even when correcting for other known prognostic factors.<sup>12</sup> This trend has led investigators to research deintensification of treatment for patients with HPV-related head and neck cancers in order to limit toxicities, and alternatively intensification of treatment to improve tumor control in those with a significant HPV-negative cancer with a smoking history.<sup>9, 11</sup> In preparing this report, we sought to identify, where possible, HPV-positive patients as separate entities from HPV-negative patients.

## **RT in Head and Neck Cancer**

### **Overview**

RT designs have evolved over the past 30 years from being based on 2D to 3D images, incorporating increasingly complex computer algorithms.<sup>13</sup> 2DRT consists of a single beam from one to four directions with the radiation fields designed on 2D fluoroscopic simulation images. A quest to improve on survival rates and the adverse effect profile of conventional 2DRT has led to widespread adoption and application of conformal RT methods for definitive (curative) treatment of patients with SCCHN, with general abandonment of 2DRT in this role in the U.S.

Conformal RT refers to modalities in which cytotoxic radiation beams are “shaped” to cover the tumor volume plus a surrounding tissue margin to treat microscopic disease that may reside there. To standardize image-based tumor volume definitions for 3D radiation planning, the International Commission of Radiation Units and Measurements created terminology for use across institutions. Definitions include gross tumor volume

(GTV), clinical target volume (CTV), and planning target volume (PTV).<sup>14</sup> The GTV pertains to gross disease identified by clinical workup (e.g., physical exam and imaging), CTV includes the GTV and any areas at risk for microscopic disease, and PTV is an expansion of the PTV by a margin (usually 3–5 mm in the head and neck patient) to account for patient or organ motion and day-to-day setup variation.

## **Conformal RT Modalities**

Conformal external-beam photon-based RT modalities used to treat SCCHN include 3DRT, IMRT, and SBRT, which is also known as stereotactic ablative RT.<sup>13</sup> For purposes of this report, we use the term *SBRT*. Charged particle-based conformal external-beam therapy such as PBRT is also available. Here we briefly review important characteristics of each conformal RT modality considered in this CER update.

### **Three-Dimensional Conformal Radiotherapy**

3DRT allows for more accurate and precise dose calculations than achieved with 2DRT because 3DRT takes into account axial anatomy and complex tissue contours.<sup>13</sup> 3D anatomic information from diagnostic computed tomography (CT) scans is used to deliver multiple highly focused beams of radiation that converge at the tumor site. This allows accurate and precise conformity of the radiation to the tumor volume, with very rapid dose fall-off in surrounding tissues. A 3DRT treatment protocol typically comprises 60–70 Gray (Gy) delivered in 25–40 fractions (usually 1.8–2 Gy) over a period of 5–10 weeks.

### **Intensity-Modulated Radiotherapy**

In the 1990s, technological and computer treatment planning advances led to the development of IMRT.<sup>13, 15</sup> Compared with 3DRT, IMRT is a newer, more complex, and resource-intensive form of RT that delivers a high dose of ionizing radiation conformally to the target volume while sparing uninvolved, healthy tissues. A typical total dose of 60–70 Gy is usually delivered in 25–40 fractions over a period of 5–10 weeks. By varying the beam intensity across shaped radiation fields, IMRT theoretically reduces radiation dose to organs at risk (e.g., the parotid glands), potentially resulting in reduced xerostomia and improved QOL compared with conventional RT. A number of technological advances within the general category of IMRT are available or under investigation, such as segmental, dynamic, combined dynamic, and segmental in the same field, as well as conformal arc; each was noted in this CER as IMRT.

### **Stereotactic Body Radiotherapy**

SBRT delivers relatively large ablative doses of radiation in fewer treatment sessions than other conformal modalities.<sup>16</sup> Regimens generally comprise a total dose of 60 Gy at greater than 10 Gy per fraction, by definition in five or fewer fractions. The tumor location can be tracked in four dimensions (including time) using several CT imaging techniques that depend on the platform, tracking on bony structures or implanted

fiducials. SBRT can deliver very high biologically effective doses above 100 Gy equivalent that are needed to ablate a tumor and sterilize the tumor margins, minimizing damage to adjacent normal tissues. Conventionally fractionated schemes, delivering a similar total dose in 25–40 fractions, typically do not reach a similar biologically effective dose range.

## **Proton-Beam Radiotherapy**

Proton therapy has become increasingly available in the last few years. PBRT has theoretical advantages over photon therapy because PBRT lacks an “exit dose,” potentially enabling physicians to deliver high-energy conformal doses to the tumor volume while almost completely sparing normal healthy tissue.

## **Summary**

The optimal means of delivering external beam ionizing radiation in sufficient doses to cure a patient with SCCHN requires a fine balance between treatment effectiveness and associated toxicity. In CER No. 20, the compiled evidence demonstrated an advantage for IMRT over 3DRT and 2DRT in reducing late xerostomia and improving measures of xerostomia-related QOL. Evidence was insufficient to demonstrate any relative difference between interventions in measures such as overall survival or tumor control. Since CER No. 20 was published, a newer conformal technology—SBRT—has come into practice, whereas 2DRT has fallen out of use in the U.S. A surveillance study prepared in 2011 by the Ottawa and RAND EPCs suggested rationale to update CER No. 20, based on signals of new evidence that would change several conclusions of that report. Taken together, the emergence of new technology and evidence suggesting potential differences between interventions in some outcomes prompted AHRQ to prioritize this update of CER No. 20.

## **Key Questions**

The proposed key questions for CER No. 20, entitled *Comparative Effectiveness and Safety of Radiotherapy Treatments for Head and Neck Cancer*, were posted for public comment for 4 weeks during its development. At that time, changes to the key questions and the PICOTS were made based on comments received and discussion with the TEP for the report. In the surveillance assessment used to determine the priority to update the 2010 report, the language of the key questions was slightly modified, but unchanged in meaning.

The key questions we used for this update follow below. In addition to 3DRT, IMRT, and PBRT, we included SBRT, which was not part of CER No. 20. Based on input from TEP discussions and a review of the literature, we excluded 2DRT from further consideration and did not include brachytherapy. In response to TEP input, we also revised the language of Key Question 4 to expand the list of potential variables to consider.

## Key Question 1

What is the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding adverse events and QOL?

## Key Question 2

What is the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding tumor control and patient survival?

## Key Question 3

Are there differences in the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT for specific patient and tumor characteristics?

## Key Question 4

Is there variation in the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT because of differences in user experience, treatment planning, treatment delivery, and target volume delineation?

## PICOTS

### Population(s)

#### Key Questions 1–4

Populations of interest included patients with head and neck cancer. To define what constitutes head and neck cancer, we consulted clinical resources such as the National Cancer Institute's (NCI) Physician Data Query Cancer Information Summary and the National Comprehensive Cancer Network (NCCN).<sup>9</sup> The consensus definition of head and neck cancer includes tumors of:

- larynx
- pharynx (hypopharynx, oropharynx, and nasopharynx)
- lip and oral cavity
- paranasal sinus and nasal cavity
- salivary gland
- occult primary of the head and neck

The following tumors were excluded:

- brain tumors
- skull base tumors
- uveal/choroidal melanoma, other ocular and eyelid tumors
- otologic tumors
- cutaneous tumors of the head and neck (including melanoma)
- thyroid cancer
- parathyroid cancer
- esophageal cancer
- trachea tumors

All therapeutic strategies were included. RT can be delivered as a primary (curative)

intent therapy or as an adjunct to surgery. Chemotherapy can also be given as an adjunct to RT, particularly in patients with more advanced cancer (i.e., stages III or IV). We sought direct evidence for one intervention compared with another, with or without chemotherapy or surgery.

## Interventions

### Key Questions 1–4

- 3DRT: defined as any treatment plan where CT-based forward treatment planning is used to delineate radiation beams and target volumes in three dimensions.
- IMRT: defined as any treatment plan using intensity-modulated radiation beams and computerized inverse treatment planning.
- SBRT: defined as conformal RT (forward- or reverse-planned) delivered in 3–5 relatively larger doses of ionizing radiation than typically delivered in a standard conformal schedule of 25–35 doses.
- PBRT: defined as any treatment plan using proton-beam radiation.

Interventions may occur as part of a multimodal treatment strategy if the comparisons only differ with respect to the RT given.

## Comparators

### Key Questions 1–4

All therapies were compared with each other as part of a continuum of treatment for patients with head and neck cancer. Thus, we included studies in which an RT method was compared with a different method (e.g., with or without chemotherapy or surgery). We included all studies in which we could be reasonably certain additional treatments were contemporary and similar, leaving the major comparison that between RT modalities; those that we could not ascertain from the publication would be excluded.

To ensure chemotherapy or other treatments were similar and contemporary, we consulted accepted guidelines such as those from NCCN or NCI. We did not extract details on chemotherapy dosages or schedules, but rather ascertained their degree of general similarity and the proportions of patients who receive and complete such regimens. We categorized and synthesized evidence according to overall treatment (e.g., concurrent chemoradiotherapy or adjuvant RT), not mixing these settings in the strength of evidence (SOE) synthesis.

## Outcomes

### Key Questions 1, 3, and 4

**Final outcomes:** QOL and adverse events including; radiation-induced toxicities, xerostomia, mucositis, taste changes, dental problems, and dysphagia.

**Intermediate outcomes:** Salivary flow and probability of completing treatment according to protocol.

We sought evidence related to user experience, treatment planning, and target volume

delineation within the context of Key Question 4. Based on input received from the TEP, any outcomes not adequately addressed in the literature were stated as evidence gaps for primary research in the Discussion section of the report.

### **Key Questions 2–4**

**Final outcomes:** Overall survival and cancer-specific survival.

**Intermediate outcomes:** Local control and time to recurrence.

### **Timing**

All durations of followup were considered.

### **Settings**

Inpatient and outpatient.

## Analytic Framework

Figure 1 provides an analytic framework illustrating the population, interventions, outcomes, and adverse effects that guided our literature search and synthesis. It links the interventions of interest directly with final health outcomes (e.g., overall survival) and adverse events (e.g., xerostomia) as well as indirectly with final outcomes via intermediate outcomes (e.g., local control, disease-free survival).

**Figure 1. Analytic framework for comparative effectiveness of RT for head and neck cancer**

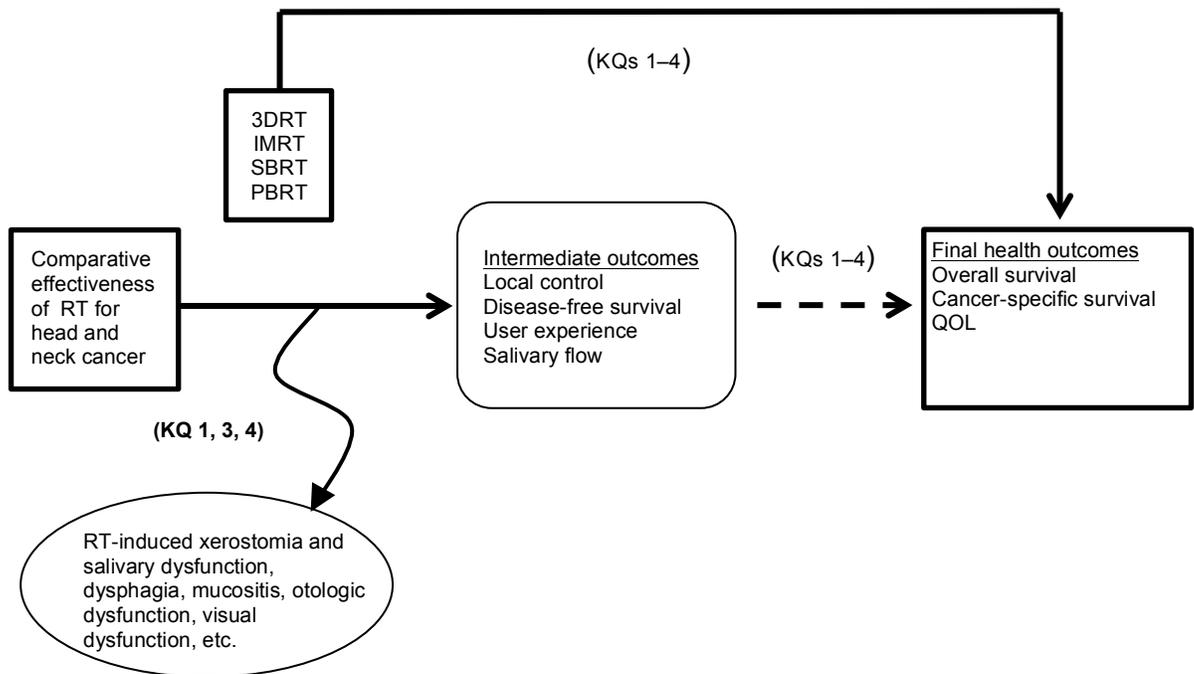


Figure 1 depicts the key questions (KQs) within the context of the PICOTS described in the previous section. In general, the figure illustrates how the interventions 3DRT, IMRT, SBRT, and PBRT may result in intermediate outcomes (e.g., local tumor control, disease-free survival) and long-term outcomes (e.g., overall survival, cancer-specific survival, QOL). Also, adverse events (e.g., radiation-associated xerostomia and salivary dysfunction, dysphagia, mucositis, otologic dysfunction, visual dysfunction) may occur at any point after the treatment is received.

## **Organization of the Report**

In the following sections of this CER update, we outline the Methods used in its preparation, including literature search strategies, methods used to select studies for inclusion, data elements and their abstraction, tabulation of results, assessment of study quality and risk of bias, and how we evaluated the SOE. In the Results section, we provide an overview of the literature search results and study inclusion and exclusion. We then present evidence for each key question, using bulleted key points and a summary of the results and tabulation of such. The Discussion section contains our assessment of the SOE as related to the conclusions of CER No. 20. Finally, we discuss the applicability of the evidence to clinical decisionmaking and gaps in the evidence base in the Discussion section. The report concludes with an overall summary that ties it together to the CER No. 20 findings.

# Methods

## Overview

This section describes the methods used to produce this CER update. Methodological practices followed were derived from the *Methods Guide* and its subsequent updates.<sup>17</sup> We also consulted the article by Tsertsvadze et al. on methods to update CERs.<sup>18</sup> The main parts in this section reflect the elements of the protocol established for the CER; certain methods map to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.

## Study Inclusion Criteria

We included only full-length reports that describe the final results of randomized controlled trials (RCTs) and nonrandomized comparative studies (observational, case-control, and cohort studies) of populations, interventions, comparisons, outcomes, treatment intervals, and settings that are part of the PICOTS (see above).

We excluded conference abstracts and other non-peer-reviewed sources, and noncomparative (single-arm) studies from this CER update. In preparation of CER No. 20, we collected a substantial body of evidence from single-arm studies. In our analysis, we found that the studies were very heterogeneous, with differences in patient populations, RT methods, treatment era, and adjunct treatments used, particularly cytotoxic chemotherapy regimens. As a consequence, we determined that the evidence was uninformative and not adequate for making valid comparisons or hypothesis generation.

## Literature Search

### Search strategies

An experienced medical librarian designed and performed all searches for this CER update. The literature search for the update was backdated to 12 months before the final literature search for CER No. 20 (dated September 28, 2009). See Appendix A for the search strategy. For SBRT (and any other new interventions, we subsequently determined merited inclusion), the literature was searched electronically for citations from January 1, 1990, through April 2013. The search will be updated at the time the draft is posted for peer review by AHRQ. We searched the following databases:

- MEDLINE<sup>®</sup>
- EMBASE<sup>®</sup>
- Cochrane Controlled Trials Register

## **Data Abstraction and Data Management**

Literature search results were transferred to EndNote<sup>®</sup> and subsequently into Distiller for study screening.

### **Review of titles and abstracts**

We developed data collection forms for abstract review, full-text review, and data extraction. Using the study-selection criteria for screening titles and abstracts, each citation was marked as eligible for review as full-text article or ineligible for full-text review. Two CER team members performed the initial title and abstract screen. A training set of 25–50 article selections were examined initially to assure uniform application of screening criteria. Full-text review was performed if it was unclear whether the study-selection criteria were satisfied. Reasons for study exclusions at the title and abstract screening phase were not noted. To be excluded, a study must have been independently excluded by both team members. Discrepancies were decided by consensus opinion; a third reviewer was consulted if necessary.

### **Full-text review**

Full-text articles were reviewed in the same fashion against the PICOTS to determine their inclusion in the systematic review. The reason for excluding an article retrieved in full-text was recorded in the Distiller database. Although an article could be excluded for multiple reasons, only the principal reason identified was noted.

### **Data abstraction**

For studies that met the inclusion criteria, we abstracted data into tables created in the Systematic Review Data Repository, with elements defined in an accompanying data dictionary. A training set of five articles was abstracted by one team member and reviewed by the Team Lead to ensure consistency. Each article included was abstracted by a single reviewer. A second reviewer assessed the data extraction against the original articles for quality control. Identified differences in data coding between the abstractor and reviewer were resolved by consensus.

The data elements abstracted included the following:

- Patient characteristics, including:
  - Age (excluding pediatric patients, 18 years or younger)
  - Sex
  - Race/ethnicity
  - Tumor location
  - Tumor stage
- Treatment characteristics, including:
  - Type of RT (e.g., photons, electrons, protons)
  - Total RT dose
  - Fractionation schedule

- Imaging methods used to guide RT (e.g., CT, implanted fiducials, bony landmarks) and the frequency of imaging to assess therapy (e.g., daily, weekly, monthly)
- Other prior or concurrent treatment modalities (e.g., systemic chemotherapy)
- Number of prior lines of treatment
- Outcome assessment
  - Identified final outcome (see PICOTS and Analytical Framework)
  - Identified intermediate outcomes (see PICOTS and Analytical Framework)
  - Adverse event response criteria
  - Followup frequency and duration
  - Data analysis details, including:
    - Statistical analyses (statistical test/estimation results)
    - Summary measures
    - Sample variability measures
    - Precision of estimate
    - p-values
  - Regression modeling techniques
    - Model type
    - Candidate predictors and methods for identifying candidates
    - Univariate analysis results
    - Selected predictors and methods for selecting predictors
    - Testing of assumptions
    - Inclusion of interaction terms
    - Multivariable model results
    - Discrimination or validation methods and results
    - Calibration or “goodness-of-fit” results

## Evidence tables

The same abstraction tables were used for all studies. The dimensions of each evidence table may vary by Key Question, but the tables contain common elements such as author, year of publication, sample size, study type, intervention(s), and comparator(s). We report outcome data in strata according to prognostic or other patient-related factors (e.g., tumor stage) provided they were reported separately or could be inferred from the study in question.

## Assessment of Methodological Risk of Bias (Quality or Limitations) of Individual Studies

In adherence to the *Methods Guide*,<sup>17</sup> the general approach to grading the quality or limitations of individual comparative studies was performed by using a United States Preventive Services Task Force (USPSTF) method (Appendix B).<sup>19</sup> Individual study quality assessment accounted for the following study elements:

- Number of participants and flow of participants through steps of study
- Treatment-allocation methods (including concealment)

- Use of blinding
- Study design (prospective vs. retrospective)
- Use of an independent outcome assessor

The quality of the abstracted studies was assessed independently by two investigators. Discordant quality assessments were resolved with input from a third reviewer, if necessary.

## Data Synthesis

The qualitative synthesis emphasized comparative studies sorted by specific head-to-head comparisons of interventions, specific treatment settings, patient characteristics, specific outcomes, and status relative the evidence hierarchy and study quality assessment.

## Grading the SOE for Individual Comparisons and Outcomes

Studies were assessed for relevance against target populations, interventions of interest, and outcomes of interest. The system used for rating the strength of the overall body of evidence is outlined in the recently updated (2013) chapter from the *Methods Guide*<sup>17</sup> and is based on a system developed by the GRADE Working Group.<sup>20</sup> This system explicitly addresses the following domains: study limitations, directness, consistency, precision, and reporting bias. Additional (optional) domains, including strength of association (magnitude of effect), dose-response association, and plausible confounding, could be addressed if appropriate. Table 1 describes the four required and three optional domains and their scores and applications.

**Table 1. SOE rating domains: required and optional**

Domain name	Domain type	Domain definition and elements	Domain score and application
Study limitations	Required	This domain reflects the degree to which included studies for a given outcome have high likelihood of protection against bias (i.e., good internal validity), assessed through two main elements: <ul style="list-style-type: none"> <li>• Study design: Whether included studies are RCTs or other designs such as nonexperimental or observational studies.</li> <li>• Study conduct: Considers aggregation of ratings of risk of bias of the individual studies under consideration.</li> </ul>	Score as one of three levels, separately by type of study design: <ul style="list-style-type: none"> <li>• Low level of study limitations</li> <li>• Medium level of study limitations</li> <li>• High level of study limitations</li> </ul>
Directness	Required	Directness relates to: <ul style="list-style-type: none"> <li>• Whether evidence links interventions directly to a health outcome of specific importance for the review, and</li> <li>• Whether the comparisons are based on head-to-head studies.</li> </ul> <p>The EPC should specify the comparison and outcome for which the SOE grade applies.</p> <p>Evidence may be indirect in several situations such as:</p> <ul style="list-style-type: none"> <li>• The outcome being graded is considered</li> </ul>	Score as one of two levels: <ul style="list-style-type: none"> <li>• Direct</li> <li>• Indirect</li> </ul> <p>If the domain score is indirect, the EPC should specify what type of indirectness accounts for the rating.</p>

		<p>intermediate (i.e., laboratory test results) in a review that is focused on clinical health outcomes (i.e., morbidity, mortality).</p> <ul style="list-style-type: none"> <li>Data do not come from head-to-head comparisons but rather from two or more bodies of evidence to compare interventions A and B (e.g., studies of A vs. placebo and B vs. placebo, or studies of A vs. C and B vs. C but not direct studies of A vs. B).</li> <li>Data are available only for proxy respondents (e.g., from family members or nurses) instead of directly from patients.</li> </ul> <p>Indirectness always implies that more than one body of evidence is required to link interventions to the most important health outcome.</p>	
Consistency	Required	<p>Consistency is the degree to which included studies find either the same direction or similar magnitude of effect. The EPC can assess this through two main elements:</p> <ul style="list-style-type: none"> <li>Direction of effect: Effect sizes have the same sign (i.e., are on the same side of no effect or a minimally important difference).</li> <li>Magnitude of effect: The range of effect sizes is similar. The EPC may consider the overlap of confidence intervals when making this evaluation.</li> </ul> <p>The importance of direction versus magnitude of effect will depend on the key question and EPC judgments.</p>	<p>Score as one of three levels:</p> <ul style="list-style-type: none"> <li>Consistent</li> <li>Inconsistent</li> <li>Unknown (e.g., single study)</li> </ul> <p>Single-study evidence bases (including mega-trials) cannot be judged with respect to consistency. In that instance, use “Consistency unknown (single study).”</p>
Precision	Required	<p>Precision is the degree of certainty surrounding an effect estimate with respect to a given outcome, based on the sufficiency of sample size and number of events. Several caveats must be considered in determining the precision of a body of evidence.</p> <ul style="list-style-type: none"> <li>A body of evidence will generally be imprecise if the optimal information size is not met. Optimal information size refers to the minimum number of patients (and events when assessing dichotomous outcomes) needed for an evidence base to be considered adequately powered.</li> <li>If an EPC performed a meta-analysis, then it may also consider whether the confidence interval crossed a threshold for a minimally important difference.</li> <li>If meta-analysis is infeasible or inappropriate, the EPC may consider the narrowness of the range of confidence intervals or the significance level of p-values in the individual studies in the evidence base.</li> </ul>	<p>Score as one of two levels:</p> <ul style="list-style-type: none"> <li>Precise</li> <li>Imprecise</li> </ul> <p>A precise estimate is one that would allow users to reach a clinically useful conclusion (e.g., treatment A is more effective than treatment B).</p>
Reporting bias	Required	<p>Reporting bias results from selectively publishing or reporting research findings bases on the favorability of direction or magnitude of effect. It includes:</p> <ul style="list-style-type: none"> <li>Study publication bias (i.e., nonreporting</li> </ul>	<p>Score as one of two levels:</p> <ul style="list-style-type: none"> <li>Suspected</li> <li>Undetected</li> </ul> <p>Reporting bias is suspected</p>

		<p>of the full study)</p> <ul style="list-style-type: none"> <li>• Selective outcome reporting bias (i.e., nonreporting or incomplete reporting of unplanned outcomes)</li> <li>• Selective analysis reporting bias (i.e., reporting one or more favorable analyses for a given outcome while not reporting other, less favorable analyses).</li> </ul> <p>Assessment of reporting bias for individual studies depends on many factors, including availability of study protocols, unpublished study documents, and patient-level data. Detecting such bias is likely with access to all relevant documentation and data pertaining to a journal publication, but such access is rare.</p> <p>Because methods to detect reporting bias in observational studies are less certain, this guidance does not require EPCs to assess it for such studies.</p>	<p>when:</p> <ul style="list-style-type: none"> <li>• Testing for funnel plot asymmetry demonstrates a substantial likelihood of bias, and/or</li> <li>• A qualitative assessment suggests the likelihood of missing studies, analyses, or outcomes data that may alter the conclusions from the reported evidence.</li> </ul> <p>Undetected reporting bias includes all alternative scenarios.</p>
Dose-response association	Optional	<p>This association, either across or within studies, refers to a pattern of a larger effect with greater exposure (dose, duration, adherence)</p>	<p>This domain should be considered when studies in the evidence base have noted levels of exposure.</p> <p>Score as one of two levels:</p> <ul style="list-style-type: none"> <li>• Present: Dose-response pattern observed.</li> <li>• Undetected: No dose-response pattern observed (dose-response relationship not present or could not be determined).</li> </ul>
Plausible confounding that would decrease observed effect	Optional	<p>Occasionally, in an observational study, plausible confounding would work in the direction opposite that of the observed effect. Had these confounders not been present, the observed effect would have been even larger than the one observed.</p>	<p>This domain should be considered when plausible confounding exists that would decrease the observed effect.</p> <p>Score as one of two levels:</p> <ul style="list-style-type: none"> <li>• Present: Confounding factors that would decrease the observed effect may be present and have not been controlled for.</li> <li>• Absent: Confounding factors that would decrease the observed effect are not likely to be present or have been controlled for.</li> </ul>
Strength of association (magnitude of effect)	Optional	<p>Strength of association refers to the likelihood that the observed effect is large enough that it could not have occurred solely as a result of bias from potential confounding factors.</p>	<p>This additional domain should be considered when the effect size is particularly large.</p> <p>Score as one of two levels:</p> <ul style="list-style-type: none"> <li>• Strong: Large effect size</li> </ul>

			<p>that is unlikely to have occurred in the absence of a true effect of the intervention.</p> <ul style="list-style-type: none"> <li>• Weak: Small enough effect size that it could have occurred solely as a result of bias from confounding factors.</li> </ul>
--	--	--	---

RCT = randomized controlled trial; EPC = ; Evidence-based Practice Center; SOE = strength of evidence.

Grading a body of evidence involves consideration of the type of studies included in the review. For assessing a clinical outcome, RCT evidence is considered the best evidence, based purely on study design. In the EPC grading system, a body of evidence including RCTs is assigned a provisional SOE grade of “high”. This may change, however, after assessment of study limitations based on how the RCTs were conducted, and other domains such as directness, consistency, and precision.

By contrast, evidence from observational studies is assumed to pose a greater risk of having study limitations because of the typically higher risk of bias attributable to a lack of randomization and inability to control for critical confounding factors. This type of evidence is generally assigned a provisional initial SOE grade of “low.” The latter may be moved up to “moderate” when study limitations are graded as low or medium, based on controls for risk of bias through study conduct or analysis. The initial SOE for observational study evidence may also be initially graded as “moderate” for certain outcomes such as important harms or for certain key questions when it is deemed at less risk for study limitations secondary to a lower risk of bias related to potential confounding.

A few real-world examples of grading evidence are illustrative of the literature encountered on this topic. In synthesizing a body of evidence represented by a single RCT rated as good quality and multiple nonrandomized comparative studies of lower quality (e.g., primarily poor), we would start with the findings from the “best available evidence” (the good quality RCT) and a high initial SOE. The study limitation domain in this instance would initially be rated as low. If the RCT and nonrandomized studies report results in opposite directions of effect, the body of evidence could be rated as having unknown consistency, thus reducing the overall strength by one level. Concluding unknown consistency is based on lack of confirmation for the direction and would be justified particularly if biases and confounding in nonrandomized studies do not have a predictable direction. However, if the differences are less dramatic and could be explained by bias in a predictable direction, then it may be considered consistent. Direct head-to-head comparisons of an intervention and comparator that report on an important health outcome lead to a rating of direct on the directness domain. In a qualitative synthesis of this hypothetical body of evidence, insufficient size (compared with the optimal information size) of the RCT would render the aggregate results imprecise on the precision domain, reducing strength by at least one level. According to EPC convention, the path through all required domains would take the strength from high through two reductions to a final strength of low.

A second example would comprise a body of observational (nonrandomized) comparative evidence that included multiple studies. Even if direct results are consistent and precise, this example would have a starting study limitations grade of high and SOE

of low. If all studies were deemed to be poor quality and poorly conducted, the body of evidence could be downgraded further to insufficient. However, application of the optional domains, particularly magnitude of effect in favor of an intervention, could raise the strength one level to low or, perhaps, moderate if sufficiently robust.

The overall SOE grade is classified into four categories, as shown in Table 2. Specific outcomes and comparisons to be rated depend on the evidence found in the literature review. The grade rating was made by independent reviewers, and disagreements were resolved by consensus adjudication.

We report a summary of key outcomes for each Key Question in a table that lists the major outcomes, the study design and number of studies of each type plus number of subjects, the findings, and the direction and magnitude of effect where applicable. The overall SOE grade for each outcome is specifically reported in this table.

**Table 2. Overall SOE categories and criteria for assignment**

Grade	Definition	Criteria for assignment
High	We are very confident that the estimate of effect lies close to the true effect for this outcome.	The body of evidence has few or no deficiencies. We believe that the findings are stable.
Moderate	We are moderately confident that the estimate of effect lies close to the true effect for this outcome.	The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	We have limited confidence that the estimate of effect lies close to the true effect for this outcome.	The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Insufficient	We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome.	No evidence is available or the body of evidence has unacceptable deficiencies, precluding judgment.

## Assessing Applicability

We assessed applicability of findings with the AHRQ Comparative Effectiveness *Methods Guide* using the PICOTS (Population, Intervention, Comparator, Outcome, Timing, Setting) framework.<sup>17,21</sup> Included studies were assessed for relevance against target populations, interventions and comparators of interest, and outcomes of interest. We anticipated that results would be applicable only to the specialized populations of interest by Key Question.

# Results

## Overview

In this section, we first report our literature search results and PRISMA diagram, which depicts the flow of articles through the review according to our screening and inclusion criteria. We then provide an overview of the design, patients, and study limitations (risk of bias) of all included studies, including relevant studies from CER No. 20. We then lay out the results for each Key Question, starting with an overview of the relevant current studies, key bulleted points of information, and a synthesis of the evidence when possible. In the results, we did not incorporate formal data synthesis (e.g., meta-analysis) because there was only one randomized trial involving the interventions of interest for treatment of head and neck cancer and the nonrandomized studies were highly heterogeneous and of “poor” quality according to the USPSTF criteria. Finally, we lay out in tabular format the conclusions and evidence base from CER No. 20 and those from this update to qualitatively integrate the findings of both.

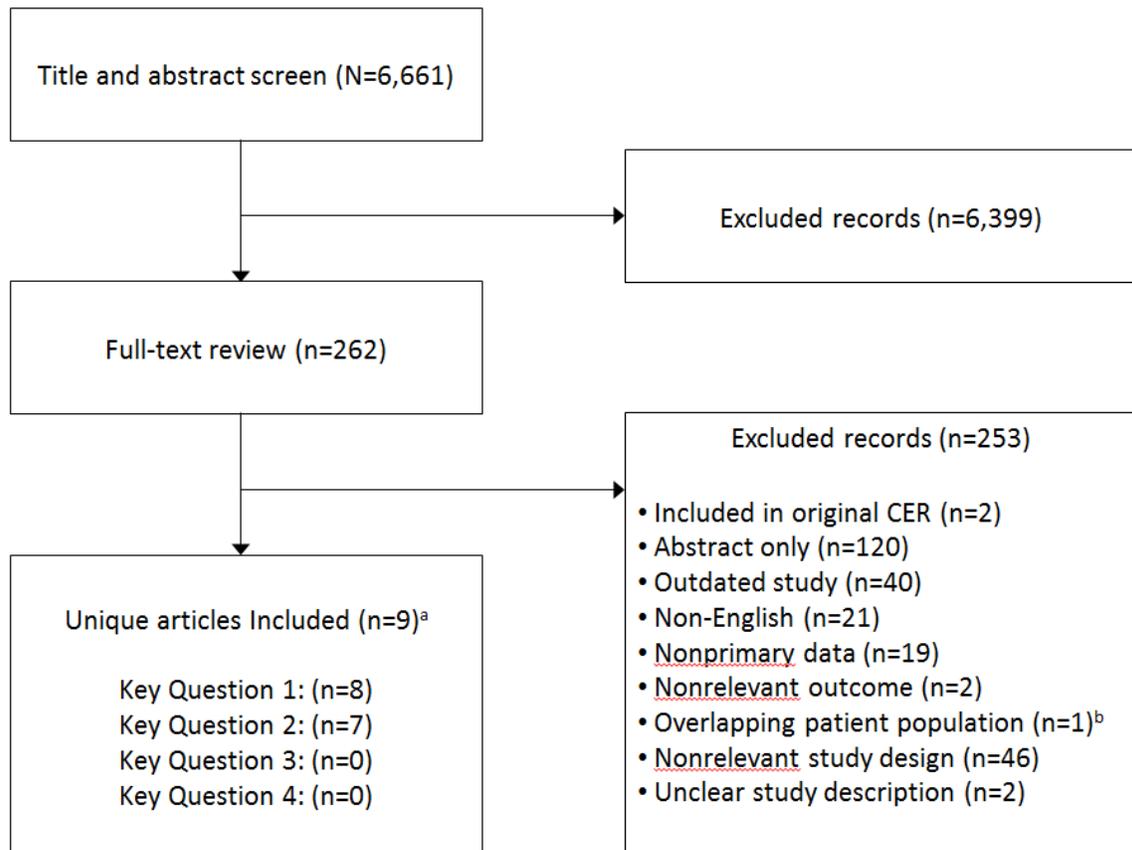
## Results of Literature Searches

### Electronic Search

A medical librarian searched MEDLINE<sup>®</sup>, EMBASE<sup>®</sup>, and the Cochrane Controlled Trials Registry for English-language articles. The overall search was performed for a period dating 12 months before the final literature search for CER No. 20 (September 28, 2009) through April 2013. For SBRT, the literature was searched for the period January 1, 1990, through April 2013.

In our initial literature search for this update, we identified 6,661 unique titles and screened 262 in full text. Of the latter, nine (total N=1,072) met the CER inclusion criteria.<sup>22-30</sup> See Appendix C for the list of 253 excluded full-text screened articles (with reasons for exclusion). The flow of articles through the screening and study selection process is shown in the PRISMA diagram (Figure 2).

**Figure 2. PRISMA diagram for disposition of literature search results**



<sup>a</sup> Six studies addressed both Key Questions 1 and 2.

<sup>b</sup> Overlapping patient population refers to the studies in which the same patients were included in more than one study. In this case, only one study was included to avoid oversampling. Decision to include a study was based on the clarity in reporting relevant patients and/or outcomes.

## Grey Literature (Publication Bias)

The study selection criteria for this update stipulate exclusion of abstracts or other non-peer-reviewed or non-full-length studies. Therefore we did not include any information based on comprehensive searches of meeting abstracts. We examined the bibliographies of all papers screened in full text to identify peer-reviewed articles the electronic search may have missed.

We accessed the Web site ClinicalTrials.gov to identify ongoing phase 3 RCTs that would meet the criteria for inclusion based on our protocol. We identified two phase 3 RCTs of conformal RT in head and neck cancer that are recruiting patients. The first trial (NCT01893307) is designed to compare IMRT and PBRT in the treatment of oropharyngeal cancer. The primary outcome is the incidence of any late-onset (>90 days) grade 3 toxicity during the 2 years after completion of RT. The second RCT (NCT01216800) is designed to compare the effects of IMRT and 3DRT on auditory function (hearing) when used as adjuvant therapy in patients who have undergone surgical resection of parotid tumors. After a MEDLINE search of the NCT number(s) and title(s), we did not find any published results; it is unknown whether any data have been

reported. Examination of a Scientific Information Packet from one manufacturer of RT equipment did not yield additional published evidence to add to this update.

## **Description of All Included Studies**

We identified nine studies that met the inclusion criteria for this CER update. All are generally described in this section; other details and results specific to a particular Key Question are considered in the relevant subsections to follow.

## **Study Limitations**

According to the USPSTF criteria for assessing the risk of bias of individual studies, the RCT of Gupta was rated “fair,” whereas the eight nonrandomized studies were rated “poor.” The rationale for the ratings is provided in Table B-1.

We assigned a “fair” USPSTF rating to the RCT of Gupta primarily because the study was not double-blinded, particularly its outcomes assessments. The investigators did not report an intention-treat analysis, but this is moot because they reported a 97 percent followup rate in each of two study arms. Gupta reported aggregated survival results in patients with tumors in different sites. However, the distribution of tumor sites and characteristics between arms was similar. Overall, the two study arms were statistically similar and comparable.

The eight nonrandomized studies were retrospective database analyses, one of which used a historical comparator group. All of the included nonrandomized studies reported results in aggregate, mixing outcomes achieved in heterogeneous groups who may not have received the same treatment(s). Overall, these eight studies were poorly designed, executed, and reported.

## **Study Design and Patient Characteristics**

Table 3 provides a high-level view of the studies included in this update. For comparative purposes, Table 3 also depicts the studies from CER No. 20 that compared 3DRT and IMRT and reported on clinical outcomes covered herein. We address applicable evidence in more detail in the Discussion section, relating the results and conclusions to those of this update.

In total, for the update, 3DRT and IMRT were compared in eight studies, including one small (N=60) RCT.<sup>28</sup> One study compared 3DRT and SBRT<sup>30</sup>; none compared IMRT and SBRT. As in CER No. 20, no evidence was identified on PBRT. Study details are summarized in Table B-2.

Overall, similar to what we identified for CER No. 20, the body of studies in the update is heterogeneous in terms of tumor site and stage, treatment setting, and treatment intent (e.g., curative vs. palliative or recurrent). Patients were generally in their mid-fifties, with males predominating across studies. Tumor sites included the hypopharynx, larynx, nasal sinus, nasopharynx, oral cavity, and oropharynx. Four studies involved patients with single tumor sites. The majority of patients across studies had locally advanced (stage III and IV) cancer, although small proportions of patients had stage I or II disease.

The treatment settings included concurrent chemoradiotherapy (CCRT); RT with or without concurrent chemotherapy (CCT); CCRT with or without surgery; and adjuvant

postoperative RT. Where it appears in all tables throughout this report, the term RT ± CCT refers to treatment regimens in which all patients received RT, but not all received concurrent CT. This is distinct from the setting of CCRT, in which all patients were reported to have received RT and CT concurrently. We did not abstract information on specific chemotherapy regimens or surgical procedures; they are beyond the scope of this review. As summarized in Table B-3, ionizing radiation was delivered by 3DRT or IMRT to a total dose of 60–74 Gy using conventional fractionation schedules, which are typical of 3DRT and IMRT (30–35 fractions, 2 Gy per fraction for 5–7 weeks); SBRT was delivered in a similar total dose but in five single fractions. We did not abstract or report on RT protocols in detail because they also are beyond the proposed scope of the review.

**Table 3. Design and characteristics of studies included in the CER No. 20 Update and CER No. 20**

Investigator (year)	Comparison	Total no. patients	RCT	Non-RCT	Mixed tumor setting	Single tumor setting	CCRT	RT ± CCT	CCRT ± surgery	Postop RT	RT ± CCT± surgery	rRT ± CCT	USPSTF study quality
<b>Update</b>													
Gupta [2012] <sup>28</sup>	3DRT vs. IMRT	60	•		•		•						Fair
Al-Mamgani (2013) <sup>24</sup>		204		•		•	•						Poor
Lambrech (2013) <sup>29</sup>		245		•	•		•						Poor
Al-Mamgani (2012) <sup>22</sup>		176		•		•		•					Poor
Chen (2012) <sup>25</sup>		155		•	•			•					Poor
Al-Mamgani (2012) <sup>23</sup>		82		•		•			•				Poor
Dirix (2010) <sup>26</sup>		81		•	•					•			Poor
Guan (2013) <sup>27</sup>		59		•	•						•		Poor
Ozyigit (2011) <sup>30</sup>	3DRT vs. SBRT	51		•		•						•	Poor
<b>CER No. 20</b>													
Nutting (2011) <sup>31</sup>	3DRT vs. IMRT	84	•			•		•					Good
Chao (2001) <sup>32</sup>		41		•	•						•		Poor
Marchal (2004) <sup>33</sup>		87		•	•						•		Poor
Chen (2007) <sup>34</sup>		68		•	•						•		Poor

Fang (2007) <sup>35</sup>	85		•		•		•					Poor
Golen (2007) <sup>36</sup>	40		•	•			•					Poor
Hodge (2007) <sup>37</sup>	195		•		•		•					Poor
Rades (2007) <sup>38</sup>	44		•		•					•		Poor
Fang (2008) <sup>39</sup>	203		•		•		•					Poor
Gomez (2008) <sup>40</sup>	42		•		•					•		Poor
Palazzi (2008) <sup>41</sup>	137		•		•					•		Poor
Rusthoven (2008) <sup>42</sup>	87		•		•		•					Poor
Vergeer (2008) <sup>43</sup>	141		•		•					•		Poor
Langendijk (2009) <sup>44</sup>	529		•		•					•		Poor

CCRT = concurrent chemoradiotherapy; CCT = concurrent chemotherapy; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; RCT = randomized controlled trial; RT = radiotherapy; rRT = reirradiated radiotherapy; 3DRT = three-dimensional conformal radiotherapy; USPSTF = United States Preventive Services Task Force.

# **Key Question 1: Comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding adverse events and QOL**

## **Overview**

Tables 4 and 5 depict key acute and late toxicity outcomes reported by each relevant study; a blank cell in any table means that the toxicity was not reported in that study.

Not all outcomes were collected in each study. Four studies reported no acute toxicities,<sup>23, 25, 27, 30</sup> and three reported no late toxicities.<sup>25, 27, 30</sup> Only the study by Chen 2012 reported QOL evidence according to RT modality.<sup>25</sup>

Patients in all studies, except that of Dirix 2010,<sup>26</sup> received chemotherapy as part of treatment; those treated by Dirix received postoperative RT. In general, investigators did not adjust results to account for chemotherapy-associated toxicities independently of RT-associated toxicities, which complicates interpretation of toxicity evidence for many adverse events (e.g., mucositis). This is somewhat ameliorated by our focus on studies in which chemotherapy regimens are similar between study arms, thus potentially isolating the effect of the RT modality on such outcomes. However, as we show in Tables 4 and 5, toxicity outcomes were inconsistently reported across studies. For this reason, we focused this update, as we did CER No. 20, on those toxicities prominently associated with RT in the head and neck: dysphagia, salivary gland function, and xerostomia. We also only consider toxicities of grade 2 or greater according to accepted criteria, such as those of the Radiation Therapy Oncology Group (RTOG) or the NCI Common Terminology Criteria for Adverse Events (CTCAE). Grades greater than 2 are those that have direct impact on patient outcomes and can adversely affect treatment delivery.

**Table 4. Summary of key reported acute (<90 days posttreatment) comparative toxicity outcomes**

Study (year)	Study design (USPSTF rating)	RT modalities (n)	Treatment setting	Tumor setting	Dermatitis	Dysphagia	Mucositis	Nausea	Pain	Salivary glands	Weight loss	Other
Gupta (2012) <sup>28</sup>	RCT (fair)	3DRT (28) IMRT (32)	CCRT	Hypopharynx Larynx Oropharynx	•	•	•			•	•	
Al-Mamgani (2013) <sup>24</sup>	Comparative Retrospective (poor)	3DRT (65) IMRT (139)	CCRT	Oropharynx	•	•	•		•			
Lambrecht (2013) <sup>29</sup>	Comparative Retrospective (poor)	3DRT (135) IMRT (110)	CCRT	Hypopharynx Larynx Nasopharynx Oral cavity Oropharynx		•	•					Erythema
Al-Mamgani (2012) <sup>22</sup>	Comparative Retrospective (poor)	3DRT (62) IMRT (114)	RT ± CCT	Hypopharynx	•	•	•		•			Neutropenic fever Intercurrent infection Severe malaise
Chen (2012) <sup>25</sup>	Comparative Retrospective (poor)	3DRT (71) IMRT (84)	RT ± CCT	Hypopharynx Larynx Nasopharynx Oropharynx Unknown primary								
Al-Mamgani (2012) <sup>23</sup>	Comparative Retrospective (poor)	3DRT (25) IMRT (57)	CCRT ± surgery	Paranasal sinus								
Dirix (2010) <sup>26</sup>	Comparative Prospective (IMRT) Retrospective (3DRT) (poor)	3DRT (41) IMRT (40)	Postoperative RT	Nasal cavity Paranasal sinus	•	•	•		•	•		Smell Taste Fatigue Conjunctivitis Dry eye Tearing Alopecia Tinnitus Serous otitis Blurred vision
Guan	Comparative Retrospective	3DRT (16) IMRT (43)	RT ± CCT ± surgery	Nasal cavity Paranasal sinus								

(2013) <sup>27</sup>	(poor)											
Ozyigit (2011) <sup>30</sup>	Comparative Retrospective (poor)	3DRT (27) SBRT (24)	rRT ± CCT	Nasopharynx								

CCRT = concurrent chemoradiotherapy; CCT = concurrent chemotherapy; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; RCT = randomized controlled trial; RT = radiotherapy; rRT = reirradiated radiotherapy; SBRT = stereotactic body radiotherapy; 3DRT = three-dimensional conformal radiotherapy; USPSTF = United States Preventive Services Task Force.

**Table 5. Summary of key reported late (>90 days posttreatment) comparative toxicity outcomes**

Study (year)	Study design (USPSTF Rating)	RT Modalities (n)	Treatment Setting	Tumor Setting	Dysphagia	Mucositis	Pain	Skin	Salivary glands	Subcutaneous	Xerostomia	Other
Gupta (2012) <sup>28</sup>	RCT (fair)	3DRT (28) IMRT (32)	CCRT	Hypopharynx Larynx Oropharynx					•	•	•	
Al-Mamgani (2013) <sup>24</sup>	Comparative Retrospective (poor)	3DRT (65) IMRT (139)	CCRT	Oropharynx	•	•	•	•		•	•	
Lambrecht (2013) <sup>29</sup>	Comparative Retrospective (poor)	3DRT (135) IMRT (110)	CCRT	Hypopharynx Larynx Nasopharynx Oral cavity Oropharynx	•						•	
Al-Mamgani (2012) <sup>22</sup>	Comparative Retrospective (poor)	3DRT (62) IMRT (114)	RT ± CCT	Hypopharynx	•	•	•	•		•	•	Cartilage necrosis Esophagus
Chen (2012) <sup>25</sup>	Comparative Retrospective (poor)	3DRT (71) IMRT (84)	RT ± CCT	Hypopharynx Larynx Nasopharynx Oropharynx Unknown primary								
Al-Mamgani (2012) <sup>23</sup>	Comparative Retrospective (poor)	3DRT (25) IMRT (57)	CCRT ± surgery	Paranasal sinus	•						•	Osteoradio-necrosis Nasolacrimal duct Stenosis Ectropion Entropion Blindness Trismus Deafness
Dirix (2010) <sup>26</sup>	Comparative Prospective (IMRT) Retrospective (3DRT) (poor)	3DRT (41) IMRT (40)	Postoperative RT	Nasal cavity Paranasal sinus		•	•	•	•			Dry eye syndrome Neuropathy

Guan (2013) <sup>27</sup>	Comparative Retrospective (poor)	3DRT (16) IMRT (43)	RT ± CCT ± surgery	Nasal cavity Paranasal sinus								
Ozyigit (2011) <sup>30</sup>	Comparative Retrospective (poor)	3DRT (27) SBRT (24)	rRT ± CCT <sup>1</sup>	Nasopharynx								Cranial neuropathy Carotid blow-out syndrome Brain necrosis Trismus

CCRT = concurrent chemoradiotherapy; CCT = concurrent chemotherapy; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; RCT = randomized controlled trial; RT = radiotherapy; rRT = reirradiated radiotherapy; SBRT = stereotactic body radiotherapy; 3DRT = three-dimensional conformal radiotherapy; USPSTF = United States Preventive Services Task Force.

<sup>a</sup> The study involved reirradiation of recurrent head-and-neck cancer tumors.

## Key Points

- New comparative evidence assessed in this update strengthens the conclusion from CER No. 20 that the risk of grade 2 or higher late xerostomia is significantly lower in patients treated with IMRT than 3DRT.
- Evidence remains insufficient to draw relative conclusions on adverse events other than late xerostomia (e.g., mucositis, dysphagia, skin toxicities, osteoradionecrosis of the jaw) that alter conclusions of CER No. 20.
- The results of primary interest for this Key Question comprise acute (<90 days) and late (>90 days) radiation-associated dysphagia, salivary gland dysfunction, xerostomia, and QOL.
- Posttreatment toxicities were reported inconsistently across studies, precluding comparisons within the body of evidence. We are uncertain whether the limited evidence on RT-associated toxicities overall reflects their absence or that the investigators did not systematically collect or report them.
- The best quality evidence comprises one small (N=60), fair quality RCT (Gupta 2012) in which 3DRT and IMRT were compared in the setting of concurrent CRT to treat patients with cancer of the hypopharynx, larynx, and oropharynx. Key findings of this study relevant to Key Question 1 pertained to late xerostomia and salivary gland dysfunction.
- One nonrandomized, poor quality study of 3DRT versus SBRT did not report on primary outcomes for Key Question 1.
- One nonrandomized, poor quality study reported QOL outcomes related to treatment with 3DRT or IMRT.

## Qualitative Synthesis

In Table 6, we aggregate new evidence related to Key Question 1 on toxicities actually reported in studies according to the intervention comparison, treatment setting, and timeframe (acute vs. long-term). We identified no evidence from patients stratified according to tumor site(s), so we did not include tumor information in this table. Although we collected evidence on lesser NCI CTCAE or RTOG grades, as shown in Tables B-4 and B-5, here we present grade 2 or higher toxicities, which are likely to adversely impact patient management, hospitalization, and survival outcomes. The last two columns of Table 6 show reported proportions for each toxicity and any statistically significant results by study if so achieved.

### RT-Associated Toxicities

Results from one nonrandomized study show a statistically significant lower rate of acute dysphagia (49 percent vs. 84 percent, respectively,  $p=0.04$ ) with IMRT compared with 3DRT in the setting of concurrent CRT.<sup>24</sup> The RCT of Gupta 2012<sup>28</sup> showed a lower rate of acute dysphagia with 3DRT (0 percent) than IMRT (9.5 percent), although the difference was nonsignificant ( $p=0.21$ ). Significantly reduced rates of late dysphagia were reported in single studies of IMRT compared with 3DRT in the setting of concurrent CRT<sup>24</sup> or RT with or without concurrent CT.<sup>22</sup> Two individual studies showed

a reduced rate of acute salivary gland dysfunction with IMRT compared with 3DRT in the setting of concurrent CRT<sup>28</sup> or postoperative RT,<sup>26</sup> respectively.

As shown in Table 6, all three studies of IMRT compared with 3DRT in the setting of concurrent CRT showed statistically significant reduction in late xerostomia.<sup>24, 28, 29</sup> The rate of late xerostomia also was significantly lower with IMRT than with 3DRT in single studies in the setting of RT with or without concurrent CT,<sup>22</sup> or postoperative RT,<sup>23</sup> respectively.

### **RT-Associated QOL**

One nonrandomized study reported QOL evidence on IMRT versus 3DRT in the setting of RT with or without concurrent CT. Chen et al. reported on mean QOL scores using the University of Washington Quality of Life validated, self-administered tool.<sup>25</sup> In this study, the salivary gland domain was the only specific component of this score wherein significant differences were observed between the IMRT and 3DRT groups at both 1 and 2 years ( $p < 0.001$  at both points). Other domains (pain, appearance, activity, recreation, swallowing, chewing, speech, shoulder, taste, saliva, mood, anxiety) showed no differences according to RT modality. At 1 year after completion of RT, the global QOL was rated as “very good” or “outstanding” among 51 percent of patients treated with IMRT compared with 41 percent of those treated with 3DRT ( $p = 0.11$ ). However, at 2 years, the corresponding percentages were 73 percent and 49 percent, respectively ( $p < 0.001$ ), showing a benefit of IMRT. Multivariate analysis showed no effect on QOL scores of age, sex, radiation intent, radiation dose, T stage, primary site, or use of concurrent CT and neck dissection. The use of IMRT was the only variable associated with improved QOL ( $p < 0.01$ ).

**Table 6. Key Question 1: Evidence synthesis for key reported comparative grade 2 or higher toxicity outcomes**

Intervention	Comparator	Treatment setting	Outcome	Timeframe	Number of studies (number of patients)	Reported rates across studies (%)	Individual study statistically significant results (p-value)
3DRT	IMRT	CCRT	Dysphagia	Acute	Three studies <sup>24, 28, 29</sup> (N=509)	3DRT: 0–84% IMRT: 9.5–76%	Only 1 study showed a statistically significant benefit of IMRT vs. 3DRT 3DRT: 84% IMRT: 49% (p=0.04) <sup>24</sup>
				Late	Two studies <sup>24, 29</sup> (N=707)	3DRT: 30%, 34% IMRT: 20%, 38%	Only 1 study showed a statistically significant benefit of IMRT vs. 3DRT (grade ≥2) 3DRT: 30% IMRT: 20% (p=0.04) <sup>24</sup>
		RT ± CCT	Dysphagia	Acute	One study <sup>22</sup> (N=176)	3DRT: 47% IMRT: 36%	Not significant
				Late	One study <sup>22</sup> (N=176)	3DRT: 10% IMRT: 1%	p=0.02
		CCRT ± surgery	Dysphagia	Late	One study <sup>23</sup> (N=82)	3DRT: 12% IMRT: 5%	Not significant
		Postoperative RT	Dysphagia	Acute	One study <sup>26</sup> (N=81)	3DRT (any grade): 34% IMRT (grade 2): 7.5%	p=0.003
3DRT	IMRT	CCRT	Salivary glands	Acute	One study <sup>28</sup> (N=60)	3DRT (grade 2): 89% IMRT (grade 2): 59%	p=0.03
		Postoperative RT	Salivary glands	Acute	One study <sup>26</sup> (N=81)	3DRT (any grade): 83% IMRT (grade 2): 0.0%	p<0.001
3DRT	IMRT	CCRT	Xerostomia	Late	Three studies <sup>24, 28, 29</sup> (N=509)	3DRT (grade >2): 49–77% IMRT (grade >2): 23–33%	All three studies showed statistically significant benefit of IMRT vs. 3DRT: p=0.001, p=0.002, p<0.001
		RT ± CCT	Xerostomia	Late	One study <sup>22</sup> (N=176)	3DRT (grade 2): 24% IMRT (grade 2): 11%	p=0.009
		CCRT ± surgery	Xerostomia	Late	One study <sup>23</sup> (N=82)	3DRT (grade >2): 16% IMRT (grade >2): 7%	Not significant

		Postoperative RT	Xerostomia	Late	One study <sup>26</sup> (N=81)	3DRT (any grade): 34% IMRT (grade 2): 0.0%	p=0.03
--	--	------------------	------------	------	-----------------------------------	---	--------

CCRT = concurrent chemoradiotherapy; CCT = concurrent chemotherapy; IMRT = intensity-modulated radiotherapy; RT = radiotherapy; 3DRT = three-dimensional conformal radiotherapy.

## **Strength of Evidence for Key Question 1**

To evaluate the SOE, we used an approach specifically developed by the AHRQ EPC program and referenced in the *Methods Guide*. This approach is based on a system initially described by the GRADE Working Group. It explicitly addresses four required domains: risk of bias, directness, consistency, and precision, as outlined in the Methods section.

Table 7 shows the SOE for new evidence on the comparative effects on QOL and toxicities of 3DRT, IMRT, SBRT, and PBRT in the treatment of head and neck cancer patients.

The evidence we identified for this update supports an SOE rating of “moderate” for the comparison of 3DRT and IMRT in the treatment setting of CCRT, showing a benefit of IMRT in significantly reducing the incidence of late grade 2 or higher xerostomia. Two other studies showed a statistically significant reduction in the incidence of late grade 2 or higher xerostomia in two other treatment settings (RT with or without CCT, postoperative RT). New evidence on any other RT-associated toxicity is insufficient to form conclusions on a benefit or harm of 3DRT compared with IMRT.

**Table 7. SOE for Key Question 1: Adverse effects and QOL**

Intervention	Comparator	Treatment setting	Outcome	Evidence base (number of patients)	Study limitations (risk of bias)	Directness	Consistency	Precision	Overall SOE
3DRT	IMRT	CCRT	Late xerostomia	Three studies including one RCT <sup>24, 28, 29</sup> (N=509)	Moderate  One “fair” quality small RCT (n = 60, Gupta 2012) plus two “poor” quality non-randomized studies result in a “moderate” study limitations rating	Direct  All three studies directly compared IMRT and 3DRT.	Consistent  All three studies showed a statistically significant reduction of late grade >2 xerostomia with IMRT compared with 3DRT (p=0.001, p<0.002, p<0.001)	Precise	Moderate  The body of evidence comprises one “fair” quality RCT, for a provisional SOE of “high”. We downgraded the SOE one level based on the “moderate” risk of bias of the body of evidence. Although the Gupta trial was relatively small, its statistically significant result coupled with similar findings of the much larger nonrandomized evidence merits an overall rating of precise.
		RT ± CCT	Late xerostomia	One study <sup>22</sup> (N=176)	High	Direct	Unknown	Imprecise	Insufficient
		CCRT ± surgery	Late xerostomia	One study <sup>23</sup> (N=82)	High	Direct	Unknown	Imprecise	Insufficient
		Postoperative RT	Late xerostomia	One study <sup>26</sup>	High	Direct	Unknown	Imprecise	Insufficient

			(N=81)						
		CCRT	Acute dysphagia	Three studies, including one RCT <sup>24, 28, 29</sup> (N=509)	Moderate  One “fair” quality small RCT (n = 60, Gupta 2012) plus two “poor” quality non-randomized studies result in a “moderate” study limitations rating.	Direct	Inconsistent  One study non-randomized study showed a statistically significant reduction with IMRT (49%) compared with 3DRT (84%). <sup>24</sup> The other non-RCT showed a directionally same but nonsignificant effect that favored IMRT over 3DRT. Gupta 2012 showed a lower but also nonsignificant rate difference of acute dysphagia with 3DRT (0%) compared with IMRT (9.5%) (p=0.21).	Imprecise  The Gupta RCT only included 60 cases, compared with 449 for the other 2 studies. It was likely not sufficiently powered to detect slight changes in rates of adverse effects, particularly in the face of much larger, albeit “poor” quality non-RCT evidence.	Insufficient  A “high” provisional SOE based on the Gupta RCT was reduced three levels for three reasons: (1) inconsistent rating; (2) imprecise rating based on the small size of the Gupta RCT and its nonsignificant result; and (3) the two non-randomized studies were of “poor” quality, heterogeneous, and subject to a high risk of bias, thus increasing the risk of bias to “moderate” for the body of evidence.
			Late dysphagia	Two studies, not including the Gupta RCT <sup>24, 29</sup> (N=707)	High  Two “poor” quality, non-randomized studies comprise the body of evidence.	Direct	Inconsistent  One study showed a statistically significant effect (p=0.03) of IMRT compared with 3DRT, with the second study showing a	Precise	Insufficient  The two nonrandomized studies were “poor” quality and heterogeneous, with high risk of bias that compromises the value of their

							reduction, albeit nonsignificant reduction.		results.
		RT ± CCT	Acute dysphagia	One study <sup>22</sup> (N=176)	High	Direct	Unknown	Imprecise	Insufficient
			Late dysphagia	One study <sup>22</sup> (N=176)	High	Direct	Unknown	Imprecise	Insufficient
			QOL	One study <sup>25</sup> (N=155)	High	Direct	Unknown	Imprecise	Insufficient
		CCRT ± surgery	Late dysphagia	One study <sup>23</sup> (N=82)	High	Direct	Unknown	Imprecise	Insufficient
		Postoperative RT	Acute dysphagia	One study <sup>26</sup> (N=81)	High	Direct	Unknown	Imprecise	Insufficient
		CCRT	Acute salivary gland dysfunction	One study <sup>28</sup> (N=60)	High	Direct	Unknown	Imprecise	Insufficient
		Postoperative RT	Acute salivary gland dysfunction	One study <sup>26</sup> (N=81)	High	Direct	Unknown	Imprecise	Insufficient
	SBRT PBRT	Any setting	Any outcome	No evidence	No evidence	No evidence	No evidence	No evidence	No evidence
IMRT	SBRT PBRT	Any setting	Any outcome	No evidence	No evidence	No evidence	No evidence	No evidence	No evidence

CCT = concurrent chemotherapy; CCRT = concurrent chemoradiotherapy; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; RCT = randomized controlled trial; RT = radiotherapy; SBRT = stereotactic body radiotherapy; SOE = strength of evidence; 3DRT = three-dimensional conformal radiotherapy.

## **Key Question 2: Comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding tumor control and patient survival**

### **Overview**

Table 8 depicts key oncologic outcomes reported by each relevant study; a blank cell in each table means the outcome was not reported in that study. Not all outcomes were collected in each study. Three studies, including the Gupta RCT, reported data on overall survival, local control (no evidence of primary tumor), or locoregional control (no evidence of primary tumor or regional metastatic spread) among patients treated with IMRT compared with 3DRT in the setting of concurrent CRT.<sup>24, 28, 29</sup> Overall survival was reported in one study of IMRT versus 3DRT in the setting of postoperative RT.<sup>26</sup> Overall survival was also reported in the study of 3DRT versus SBRT in the setting of RT with or without concurrent CT.<sup>30</sup> Other oncologic outcomes were inconsistently reported across the body of studies, as shown in Table 8.

**Table 8. Summary of key reported comparative oncologic outcomes**

Study (year)	Study design (USPSTF rating)	RT modalities (n)	Treatment setting	Tumor setting	Overall survival	Cancer-specific survival	Disease-free survival	Local control	Loco-regional control	Distant control	Other
Gupta (2012) <sup>28</sup>	RCT (fair)	3DRT (28) IMRT (32)	CCRT	Hypopharynx Larynx Oropharynx	•				•		
Al-Mamgani (2013) <sup>24</sup>	Comparative Retrospective (poor)	3DRT (65) IMRT (139)	CCRT	Oropharynx	•	•	•	•			
Lambrecht (2013) <sup>29</sup>	Comparative Retrospective (poor)	3DRT (135) IMRT (110)	CCRT	Hypopharynx Larynx Nasopharynx Oral cavity Oropharynx	•				•		
Al-Mamgani (2012) <sup>22</sup>	Comparative Retrospective (poor)	3DRT (62) IMRT (114)	RT ± CCT	Hypopharynx							•
Chen (2012) <sup>25</sup>	Comparative Retrospective (poor)	3DRT (71) IMRT (84)	RT ± CCT	Hypopharynx Larynx Nasopharynx Oropharynx Unknown primary							•
Al-Mamgani (2012) <sup>23</sup>	Comparative Retrospective (poor)	3DRT (25) IMRT (57)	CCRT ± surgery	Paranasal sinus				•			
Dirix (2010) <sup>26</sup>	Comparative Prospective (IMRT) Retrospective (3DRT) (poor)	3DRT (41) IMRT (40)	Postoperative RT	Nasal cavity Paranasal sinus	•		•	•		•	
Guan (2013) <sup>27</sup>	Comparative Retrospective	3DRT (16) IMRT (43)	RT ± CCT ± surgery	Nasal cavity Paranasal					•		•

	(poor)			sinus							
Ozyigit (2011) <sup>30</sup>	Comparative Retrospective (poor)	3DRT (27) SBRT (24)	rRT ± CCT	Nasopharynx	•	•		•			

3DRT = three-dimensional conformal radiotherapy; CCT = concurrent chemotherapy; CCRT = concurrent chemoradiotherapy; IMRT = intensity-modulated radiotherapy; RCT = randomized controlled trial; RT = radiotherapy; rRT = reirradiated radiotherapy; SBRT = stereotactic body radiotherapy; USPSTF = United States Preventive Services Task Force.

## Key Points

- As we found in CER No. 20, comparative evidence assessed in this update was insufficient to draw relative conclusions on any oncologic outcomes.
- The results of primary interest for this Key Question comprise overall survival, local control, and locoregional control.
- The key oncologic outcomes were not reported universally across studies, so we could not make comparisons across a larger body of evidence.
- The best quality evidence comprises one small (N=60), fair quality RCT (Gupta 2012) in which 3DRT and IMRT were compared in the setting of concurrent CRT to treat patients with cancer of the hypopharynx, larynx and oropharynx.
- Two additional nonrandomized, poor quality studies reported on the key oncologic outcomes with 3DRT and IMRT in the setting of concurrent CRT among patients with cancer of the hypopharynx, larynx, oral cavity, nasopharynx, and oropharynx.
- One study of 3DRT versus SBRT reported overall survival and local control in the setting of RT with or without concurrent CRT among patients with nasopharyngeal cancer. However, 22 percent of unidentified patients in the 3DRT arm received concurrent high-dose rate brachytherapy so the oncologic outcomes are not included in this synthesis.

## Qualitative Synthesis

In Table 9, we have aggregated new evidence related to Key Question 2 on oncologic outcomes actually reported in studies according to the intervention comparison, treatment setting, and timeframe. We identified no evidence from patients stratified according to tumor site(s) so have not included tumor information in this table. Further, we did not identify any evidence on differences in oncologic outcomes related to the HPV status of patient tumors. The last two columns of Table 9 show reported proportions for each outcome and statistically significant results if attained.

In general, evidence on tumor control and survival outcomes is sparse. Table 9 shows that no statistically significant differences were reported for overall survival, local control, or locoregional control among studies of 3DRT versus IMRT in any setting compiled there. The only statistically significant oncologic result we found was in disease-free survival with IMRT compared with 3DRT in the postoperative adjuvant setting for paranasal sinus cancer (72 percent vs. 60 percent,  $p=0.02$ ). All abstracted data are shown in detail in Table B-6.

**Table 9. Key Question 2: Evidence synthesis for key reported comparative oncologic outcomes**

Intervention	Comparator	Treatment setting	Outcome	Number of studies (number of patients)	Reported rates across studies (%)	Individual study statistically significant results (p-value)
3DRT	IMRT	CCRT	Overall survival	Three studies <sup>22, 28, 29</sup> (N=509)	<u>3 years</u> (2 studies) 3DRT: 88%, 61% IMRT: 80%, 64% <u>5 years</u> (1 study) 3DRT: 43% IMRT: 47%	No statistically significant difference in overall survival was reported in any study.
			Local control	One study <sup>24</sup> (N=204)	<u>5 years</u> 3DRT: 74% IMRT: 82%	No statistically significant difference in local control was reported.
			Locoregional control	Two studies <sup>28, 29</sup> (N=305)	<u>3 years</u> 3DRT: 71%, 71% IMRT: 68%, 70%	No statistically significant difference in locoregional control was reported in either study.
			Disease-free survival	One study <sup>24</sup> (N=204)	<u>5 years</u> 3DRT: 58% IMRT: 60%	No statistically significant difference in disease-free survival was reported.
		Postoperative RT	Overall survival	One study <sup>26</sup> (N=81)	<u>2 years</u> 3DRT: 73% IMRT: 89%	No statistically significant difference in overall survival was reported.
			Local control	One study <sup>26</sup> (N=81)	<u>2 years</u> 3DRT: 67% IMRT: 76%	No statistically significant difference in local control was reported.
			Disease-free survival	One study <sup>26</sup> (N=81)	<u>2 years</u> 3DRT: 60% IMRT: 72%	p=0.02

CCRT = concurrent chemoradiotherapy; IMRT: intensity-modulated radiotherapy; RT: radiotherapy; 3DRT: three-dimensional conformal radiotherapy.

## **Strength of Evidence for Key Question 2**

To evaluate the SOE, we used an approach specifically developed by the AHRQ EPC program and referenced in the *Methods Guide*. This approach is based on a system initially described by the GRADE Working Group. It explicitly addresses four required domains: risk of bias, directness, consistency, and precision, as outlined in the Methods section.

Table 10 shows the SOE for the comparative effects of 3DRT, IMRT, SBRT, and PBRT on oncologic outcomes in the treatment of head and neck cancer patients. The criteria we used to arrive at the SOE ratings are outlined in the Methods section of the report. Details on how the SOE ratings were determined are summarized in Table 10.

We determined that new evidence, including one “fair” quality RCT (Gupta 2012), is insufficient to support a conclusion on the relative effect of IMRT and 3DRT on overall survival or locoregional control rates in the setting of CCRT. New evidence is insufficient to form conclusions on the effect of any other RT modality comparison for any oncologic outcome in any other setting we identified in this review.

**Table 10. SOE for Key Question 2: Oncologic outcomes**

Intervention	Comparator	Setting	Outcome	Evidence base (number of patients)	Risk of bias	Directness	Consistency	Precision	Overall SOE
3DRT	IMRT	CCRT	Overall survival	Three studies including the Gupta 2012 RCT <sup>24, 28, 29</sup> (N=509)	Moderate  One “fair” quality small RCT (n = 60, Gupta 2012) plus two “poor” quality non-randomized studies result in a provisional “moderate” study limitations rating.	Direct  All three studies directly compared IMRT and 3DRT.	Consistent  All three studies showed no statistically significant difference between 3DRT and IMRT in rate of overall survival at 2 or 5 years.	Imprecise  The Gupta 2012 RCT was likely not sufficiently powered to detect slight changes in rates of overall survival with IMRT compared with 3DRT, particularly in the face of much larger, albeit “poor” quality non-RCT evidence.	Insufficient  A “high” provisional SOE based on the Gupta RCT was reduced three levels for three reasons: (1) imprecise rating based on the small size of the Gupta RCT and its nonsignificant result; (2) the two nonrandomized studies were of “poor” quality, heterogeneous, and subject to a high risk of bias, thus increasing the risk of bias to “moderate” for the body of evidence; and, (3) the relative larger size of these 2 studies compared to Gupta, accounting for 88% of all patients in the body of evidence, obscure the findings of the latter, resulting in

									an overall SOE rating of “insufficient”.
			Locoregional control	Two studies including the Gupta 2012 RCT <sup>28, 29</sup> (N=305)	Moderate  One “fair” quality RCT (Gupta 2012) and a much larger “poor” quality non-randomized study result in a “moderate” study limitations rating.	Direct  Both studies directly compared IMRT and 3DRT.	Consistent  Both studies showed no statistically significant difference between 3DRT and IMRT in rate of overall survival at 2 or 5 years.	Imprecise  The Gupta 2012 RCT is was likely not sufficiently powered to detect slight changes in rates of locoregional control with IMRT compared with 3DRT, particularly in the face of much larger, albeit “poor” quality non-RCT evidence.	Insufficient  A “high” provisional SOE based on the Gupta RCT was reduced three SOE levels basically as outlined above for overall survival. Note the patients in the nonrandomized study comprised more than 80% of the evidence base, obscuring Gupta’s results.
			Local control	One study <sup>24</sup> (N=204)	High	Direct	Unknown	Imprecise	Insufficient
			Disease-free survival	One study <sup>24</sup> (N=204)	High	Direct	Unknown	Imprecise	Insufficient
		Postoperative RT	Overall survival	One study <sup>26</sup> (N=81)	High	Direct	Unknown	Imprecise	Insufficient
			Local control	One study <sup>26</sup> (N=81)	High	Direct	Unknown	Imprecise	Insufficient
			Disease-free survival	One study <sup>26</sup> (N=81)	High	Direct	Unknown	Imprecise	Insufficient

	SBRT PBRT	Any setting	Any outcome	No evidence					
IMRT	SBRT PBRT	Any setting	Any outcome	No evidence					

CCT = concurrent chemotherapy; CCRT = concurrent chemoradiotherapy; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; RCT = randomized controlled trial; RT = radiotherapy; SBRT = stereotactic body radiotherapy; SOE = strength of evidence; 3DRT = three-dimensional conformal radiotherapy.

### **Key Question 3: Comparative effectiveness of 3DRT, IMRT, SBRT, or PBRT for specific patient and tumor characteristics**

#### **Key Points**

- In CER No. 20, no comparative studies addressed these issues. In this update, we did not identify any new evidence that specifically addressed Key Question 3.
- Therefore insufficient evidence exists to form conclusions about the comparative effects or SOE on 3DRT, IMRT, SBRT, or PBRT based on specific patient and tumor characteristics.

### **Key Question 4: Comparative effectiveness of 3DRT, IMRT, SBRT, or PBRT because of differences in user experience, treatment planning, treatment delivery, and target volume delineation**

#### **Key Points**

- In CER No. 20, no comparative studies addressed these issues. In this update, we did not identify any new evidence that specifically addressed Key Question 4.
- Therefore insufficient evidence exists to form conclusions about the comparative effects or SOE on 3DRT, IMRT, SBRT, or PBRT based on specific patient and tumor characteristics.

## Discussion

### **CER Update Strength of Evidence Relative to CER No. 20 Findings**

Table 11 provides a summary of the conclusions we drew for the relevant interventional comparisons for each Key Question in CER No. 20 and in this update. Because 2DRT and SBRT are not commonly addressed in CER No. 20 and the update, they are not included in Table 11. Moderate strength evidence from the update shows a reduction of the incidence of late grade 2 or higher xerostomia with IMRT compared with 3DRT, which strengthens the conclusion on this toxicity and comparison from CER No. 20. Evidence in the update is insufficient to show a difference between IMRT and 3DRT in overall survival or locoregional tumor control rates. We found no new evidence to alter any conclusions of CER No. 20 for any other toxicity or oncologic outcomes or comparisons.

**Table 11. Comparison of relevant CER No. 20 and update conclusions**

Key question	Comparison	Clinical outcome	CER No. 20 total evidence base	CER No. 20 conclusions	CER No. 20 update total evidence base	CER No. 20 update conclusions	Cumulative update conclusions (action needed)
Key Question 1: What is the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding adverse events and QOL?	3DRT vs. IMRT	Grade $\geq 2$ late xerostomia	One good quality RCT and six poor quality non-RCTs	Moderate SOE shows significant reduction in incidence	One fair quality RCT, two poor quality non-RCTs	Moderate SOE shows significant reduction in incidence	Raises SOE to “high” (no further study required)
		Other RT-associated grade $>2$ toxicities (e.g., acute or late dysphagia, salivary gland dysfunction, swallowing function)	Variously, one good quality RCT, 13 poor quality non-RCTs	Insufficient evidence to draw conclusions	Variously, one good quality RCT, eight poor quality non-RCTs	Insufficient evidence to draw conclusions	Insufficient evidence to draw conclusions (further study required)
		QOL	Three poor quality non-RCTs				
	3DRT or IMRT vs. PBRT	Any	No evidence identified	No evidence identified, insufficient	No evidence identified	No evidence identified, insufficient	
Key Question 2: What is the comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT regarding tumor control and patient survival?	3DRT vs. IMRT	Overall survival, locoregional control	Variously, one good quality RCT, six poor quality non-RCTs	Insufficient evidence to draw conclusions	One fair quality RCT, three poor quality non-RCTs	Insufficient evidence to draw conclusions	
	3DRT or IMRT vs. PBRT	Any	No evidence identified	No evidence identified, insufficient	No evidence identified	No evidence identified, insufficient	
Key Question 3: Are there differences in comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT	3DRT or IMRT vs. PBRT	Any	No evidence identified	No evidence identified, insufficient	No evidence identified	No evidence identified, insufficient	

for specific patient and tumor characteristics?							
Key Question 4: Is there variation in comparative effectiveness of 3DRT, IMRT, SBRT, and PBRT because of differences in user experience, treatment planning, treatment delivery, and target volume delineation?	3DRT or IMRT vs. PBRT	Any	No evidence identified	No evidence identified, insufficient	No evidence identified	No evidence identified, insufficient	

CCT = concurrent chemotherapy; CCRT = concurrent chemoradiotherapy; CER = Comparative Effectiveness Review; CT = computed tomography; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; QOL = quality of life; RCT = randomized controlled trial; RT = radiotherapy; SBRT = stereotactic body radiotherapy; SOE = strength of evidence; 3DRT = three-dimensional conformal radiotherapy.

## **Applicability of the Findings**

In general, applicability assessment would depend on a body of evidence sufficient to form new conclusions about the comparative outcomes of 3DRT, IMRT, SBRT, and PBRT in treatment of head and neck cancer. However, comparative evidence that meets study selection criteria for this CER update is sparse for 3DRT, IMRT, and SBRT, and nonexistent for PBRT. In the absence of sufficient evidence, additional factors may be considered in making a treatment decision. Those could include relative convenience and cost, issues outside the scope of this CER.

In preparing this update, we considered the interventions that we included in CER No. 20 and whether all remained applicable to current radiation oncology practice. In particular, we examined the role of conventional opposed beam 2DRT and brachytherapy in modern radiation oncology practice. Our conclusion, based on the current literature and input from our TEP members, was that 2DRT is no longer in use in the U.S. for definitive treatment of head and neck cancer, thus we excluded it from the update. Further, although brachytherapy can be used in select cases as a means of dose-escalation in conjunction with external beam irradiation for head and neck cancer<sup>1,2</sup> this practice has become uncommon because sufficient dose escalation can often be achieved in these cases with a noninvasive approach (e.g., conformal RT). Brachytherapy alone is very rarely employed, except in small (T1) tumors of the nasal vestibule, lip, or oral cavity, which are relatively uncommon (1 percent to perhaps 5 percent of all cases).<sup>3-7</sup> Therefore, because use of brachytherapy alone for primary management of head and neck malignancies has limited applicability in modern head and neck radiation oncology practice, we did not seek evidence of it for this CER; we focused instead on RT modalities that are used as the sole RT intervention for a given presentation of head and neck cancer.

## **Key Questions 1 and 2**

The degree to which the evidence presented in this report is applicable to clinical practice is a function of the similarity between populations in the included studies and the patient population that receives clinical care in diverse settings. It also is related to the relative availability of the interventions. Because of the overall weakness of evidence for Key Questions 1 and 2, we have primarily limited comments to the relevance of the PICOTS elements, a practical and useful structure to review the applicability in a systematic manner (Table 12).

**Table 12. Summary of applicability of evidence for Key Questions 1 and 2**

PICOTS Domain	Applicability of evidence
Populations	<ul style="list-style-type: none"> <li>Overall patients included in the evidence base of this CER update are typical of the head and neck cancer population treated with RT based on age, sex, and tumor characteristics.</li> </ul>
Interventions	<ul style="list-style-type: none"> <li>3DRT, IMRT, and SBRT represent different technological approaches to the delivery of conformal photon RT. The major advantage of these interventions compared with traditional wide-field 2DRT is the ability to deliver tightly focused cytotoxic radiation by delineating the shape and size of the tumor using a CT-based or other imaging planning system.</li> <li>3DRT represents a minimum technical standard for delivery of conformal RT. It involves static fields with a fixed shape, modified by compensators (wedges and segments). 3DRT is widely available.</li> <li>IMRT offers beam strength attenuation through a multileaf collimator (tungsten), with dynamic field shapes for each beam angle. IMRT is not as widely available as 3DRT and requires a higher level of inverse planning effort and quality assurance.</li> <li>SBRT is a hypofractionated technique administered in five or fewer fractions; 3DRT and IMRT typically deliver radiation in many more fractions than SBRT.</li> <li>SBRT is not as widely available as 3DRT or IMRT, but its use is growing in other settings such as non-small-cell cancer. The institutional programmatic requirements for SBRT are similar to those for IMRT.</li> <li>Comparative evidence for PBRT is unavailable.</li> </ul>
Comparators	<ul style="list-style-type: none"> <li>See above for Interventions.</li> </ul>
Outcomes	<ul style="list-style-type: none"> <li>The major beneficial health outcomes in this CER are overall survival and late toxicities, particularly xerostomia.</li> <li>Overall survival is the primary outcome of interest for any cancer intervention study.</li> <li>Local control is of interest to patients because it measures the effectiveness of an intervention in disease control. On local failure, patients enter into a new category centered on systemic chemotherapy. This is a perilous position for typically medically frail patients.</li> </ul>
Timing	<ul style="list-style-type: none"> <li>The relevant periods occur from the time of treatment through followup over months (palliation) or years (overall survival).</li> </ul>
Setting	<ul style="list-style-type: none"> <li>The evidence for Key Questions 1 and 2 is mostly international, primarily obtained in tertiary institutional settings. More sophisticated interventions such as IMRT and SBRT require an institutional commitment to quality assurance and ongoing training that may be difficult to achieve in smaller community-based centers.</li> <li>We did not collect or analyze information to examine these issues.</li> </ul>

CCT = concurrent chemotherapy; CCRT = concurrent chemoradiotherapy; CER = Comparative Effectiveness Review; CT = computed tomography; IMRT = intensity-modulated radiotherapy; PBRT = proton-beam RT; RCT = randomized controlled trial; RT = radiotherapy; SBRT = stereotactic body radiotherapy; 3DRT = three-dimensional conformal radiotherapy.

## Key Questions 3 and 4

The current evidence base for Key Questions 3 and 4 is nonexistent based on our literature review. Therefore we cannot assess the applicability to clinical practice.

## Findings in Relationship to What Is Already Known

Our updated systematic literature search and review revealed no relevant evidence-based guidelines we could compare with our findings for any of the key questions.

## **Limitations of Current Review and Evidence Base**

The primary limitation for all key questions here is lack of well-designed and conducted comparative trials. Although the body of evidence we identified was more substantial for 3DRT and IMRT than SBRT, and nonexistent for PBRT, we have significant concerns about interstudy heterogeneity, with variability in RT dose, schedule of treatment, concurrent treatments, patient selection criteria, tumor size and location, and so forth. In a systematic review in general, heterogeneous evidence makes it very difficult to assess the benefits and harms of any intervention, particularly evidence drawn from nonrandomized trials. In this CER update, the sparse new evidence we identified limited additional comparative assessment among the interventions. We therefore believe further careful study of the interventions considered in this CER is needed, particularly in the settings of Key Question 1 or 2 to establish optimal technical protocols and patient selection criteria, perhaps standardizing and comparing them across institutions. These data and methods could, in theory, be applied to the design and conduct of comparative studies, as outlined in the Research Gaps section below.

## **Research Gaps**

The primary research gap we identified is a continuing lack of evidence from well-executed comparative studies (randomized or otherwise) to draw conclusions on the relative clinical benefits and harms of the RT interventions used in patients with head and neck cancer. We also identified some feasibility issues associated with the interventions that are potential impediments to the type of rigorous comparative studies we suggest are necessary to determine their comparative effectiveness. In this section, we first describe characteristics of ideal comparative studies we believe are needed to compare these technologies. Some potential impediments to such studies are discussed subsequently in this section.

## **Lack of Clinical Trial Evidence on RT Interventions for Head and Neck Cancer**

We suggest that further prospective studies are needed to properly evaluate the relative clinical benefits and harms of the technologies evaluated in this CER, taking into account the potential impediments we discuss below. Ideally, comparative studies in this setting would incorporate the following:

- To assure comparability of patients and to minimize bias, standardized patient selection criteria would be used that involve consultation, including a head and neck surgeon, medical oncologist, and radiation oncology specialist. Key factors to consider include comorbidity status, age, performance status, tumor size, and tumor location.
- Standardized intervention protocols with training and quality assurance programs within and across participating institutions are necessary for the best study. For RT, key factors would include the imaging and planning method, immobilization method, dose, and fractionation schedule for comparisons of different modalities (e.g., 3DRT, IMRT, SBRT, PBRT).

- Prespecified followup criteria and methods—in particular, notation of systemic therapy—are key considerations in study design. Systemic therapy is a key concern because it is difficult to discern the effects of an intervention with systemic therapy from that achieved with the intervention alone. Is the effectiveness a function of the systemic therapy, the intervention, or the combination?
- Rigorous and standardized reporting is needed to account for all patients and treatments received. We urge that rigorous methods be used to conduct RCTs, particularly intention-to-treat analysis and adjustment of survival data to account for all patients based on their treatment plans.
- Primary outcomes would include overall survival, cancer-specific survival, and local control. Prespecified systematic collection of adverse events using validated criteria (e.g., CTCAE) is necessary to permit accurate assessment of relative benefits and risks of the interventions.
- As alluded to in the Introduction of this report, the potential impact of tumor tissue HPV positivity on oncologic outcomes and management of such patients has been increasing in importance. Studies are needed to identify reduced intensity therapies that still yield satisfactory oncologic outcomes in HPV-positive cases.

## **Potential Impediments to Comparative Studies of RT Interventions for Head and Neck Cancer**

The general dissemination of conformal RT technologies into community clinical practice is a potential impediment to comparative study of those technologies. We acknowledge that randomized studies of 3DRT versus IMRT or PBRT may be very difficult to recruit and conduct, based on technical and potential ethical issues related to perceptions of unequal clinical benefit among the interventions. This CER supports a conclusion that RT-associated adverse events—in particular late xerostomia—are lessened with IMRT compared with 3DRT. However, we maintain that current evidence is insufficient to support a view that clinical oncologic outcomes achieved with any of the technologies are relatively superior or inferior. Clinical evidence from comparative studies is needed to establish the standard of care for head and neck cancer patients.

## **Summary and Conclusions**

Key questions in CER No. 20 asked whether any of the RT modalities under consideration (2DRT, 3DRT, IMRT, PBRT) is more effective than the others:

- in reducing normal tissue toxicity and adverse events, and improving QOL
- in improving local tumor control, time to disease progression, and survival
- when used in certain anatomic locations or patient subpopulations
- whether there is more variation in patient outcomes with any modality secondary to user experience, treatment planning, or target volumes

The main finding of CER No. 20 was that late grade 2 or higher xerostomia was reduced and QOL domains related to xerostomia were improved in patients treated with

IMRT compared with those who received either 3DRT or 2DRT. Evidence was insufficient to draw relative conclusions on survival or tumor control; adverse events other than late xerostomia (e.g., mucositis, dysphagia, skin toxicities, osteoradionecrosis of the jaw); whether patient and tumor characteristics affected relative outcomes; or whether physician experience and treatment characteristics affected relative clinical outcomes such as survival or treatment-associated adverse events.

Moderate strength evidence from the update shows a reduction of the incidence of late grade 2 or higher xerostomia with IMRT compared with 3DRT, which strengthens the conclusion on this toxicity and comparison from CER No. 20. Evidence in the update is insufficient to show a difference between IMRT and 3DRT in overall survival or locoregional tumor control rates. We found no new evidence to alter any conclusions of CER No. 20 for any other toxicity or oncologic outcomes or comparisons.

## References

1. Erickson BA, Wilson JF. Nasopharyngeal brachytherapy. *Am J Clin Oncol*. 1993 Oct;16(5):424-43. PMID: 8213626.
2. Esche BA, Haie CM, Gerbaulet AP, et al. Interstitial and external radiotherapy in carcinoma of the soft palate and uvula. *Int J Radiat Oncol Biol Phys*. 1988 Sep;15(3):619-25. PMID: 3138213.
3. Jorgensen K, Elbrond O, Andersen AP. Carcinoma of the lip. A series of 869 cases. *Acta Radiol Ther Phys Biol*. 1973 Jun;12(3):177-90. PMID: 4749173.
4. Marchese MJ, Nori D, Anderson LL, et al. A versatile permanent planar implant technique utilizing iodine-125 seeds imbedded in gelfoam. *Int J Radiat Oncol Biol Phys*. 1984 May;10(5):747-51. PMID: 6735761.
5. Marcus RB, Jr., Million RR, Mitchell TP. A preloaded, custom-designed implantation device for stage T1-T2 carcinoma of the floor of mouth. *Int J Radiat Oncol Biol Phys*. 1980 Jan;6(1):111-3. PMID: 7364600.
6. Mendenhall NP, Parsons JT, Cassisi NJ, et al. Carcinoma of the nasal vestibule. *Int J Radiat Oncol Biol Phys*. 1984 May;10(5):627-37. PMID: 6429094.
7. Mendenhall NP, Parsons JT, Cassisi NJ, et al. Carcinoma of the nasal vestibule treated with radiation therapy. *Laryngoscope*. 1987 May;97(5):626-32. PMID: 3573913.
8. Jemal A, Bray F, Center MM, et al. Global cancer statistics. *CA: A Cancer Journal for Clinicians*. 2011;61(2):69-90.
9. Pfister DG, Ang K-K, Brizel DM, et al. Head and Neck Cancers. *Journal of the National Comprehensive Cancer Network*. 2011 June 1, 2011;9(6):596-650.
10. Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *Journal of Clinical Oncology*. 2011 November 10, 2011;29(32):4294-301.
11. Sturgis EM, Ang KK. The epidemic of HVP-associated oropharyngeal cancer is here: Is it time to change our treatment paradigms? *Journal of the National Comprehensive Cancer Network*. 2011;9(6):665-73.
12. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med*. 2010 Jul 1;363(1):24-35. PMID: 20530316.
13. Burri M, Bevan A, Roach III M. Advances in radiation therapy: conventional to 3D, to IMRT, to 4D, and beyond. *CA Cancer J Clin*. 2005;55:117-34.
14. Chavaudra J, Bridier A. [Definition of volumes in external radiotherapy: ICRU reports 50 and 62]. *Cancer Radiother*. 2001 Oct;5(5):472-8. PMID: 11715299.
15. Lee N, Puri DR, Blanco AI, et al. Intensity-modulated radiation therapy in head and neck cancers: An update. *Head & Neck*. 2007;29(4):387-400.
16. Siddiqui F, Raben D, Lu JJ, et al. Emerging applications of stereotactic body radiation therapy for head and neck cancer. *Expert Rev Anticancer Ther*. 2011 Sep;11(9):1429-36. PMID: 21929316.
17. Helfand M, Balshem H. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. August 2009. Rockville, MD. <http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayProduct&productID=318>.
18. Tsertsvadze A, Maglione M, Chou R, et al. Updating comparative effectiveness reviews: current efforts in AHRQ's Effective Health Care Program. *J Clin Epidemiol*. 2011 Nov;64(11):1208-15. PMID: 21684114.

19. Harris RP, Helfand M, Woolf SH, et al. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med.* 2001;20(3 Suppl):21-35. PMID: 11306229
20. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—Agency for Healthcare Research and Quality and the Effective Health-Care Program. *J Clin Epidemiol.* 2010;63:513-23. PMID: 19595577
21. Brown P, Brunnhuber K, Chalkidou K, et al. How to formulate research recommendations. *BMJ.* 2006 Oct 14;333(7572):804-6. PMID: 17038740.
22. Al-Mamgani A, Mehilal R, van Rooij PH, et al. Toxicity, quality of life, and functional outcomes of 176 hypopharyngeal cancer patients treated by (chemo)radiation: the impact of treatment modality and radiation technique. *Laryngoscope.* 2012 Aug;122(8):1789-95. PMID: 22833307.
23. Al-Mamgani A, Monserez D, Rooij P, et al. Highly-conformal intensity-modulated radiotherapy reduced toxicity without jeopardizing outcome in patients with paranasal sinus cancer treated by surgery and radiotherapy or (chemo)radiation. *Oral Oncol.* 2012 Sep;48(9):905-11. PMID: 22584070.
24. Al-Mamgani A, van Rooij P, Verduijn GM, et al. The impact of treatment modality and radiation technique on outcomes and toxicity of patients with locally advanced oropharyngeal cancer. *Laryngoscope.* 2013 Feb;123(2):386-93. PMID: 23404489.
25. Chen AM, Farwell DG, Luu Q, et al. Intensity-modulated radiotherapy is associated with improved global quality of life among long-term survivors of head-and-neck cancer. *Int J Radiat Oncol Biol Phys.* 2012 Sep 1;84(1):170-5. PMID: 22300572.
26. Dirix P, Vanstraelen B, Jorissen M, et al. Intensity-modulated radiotherapy for sinonasal cancer: Improved outcome compared to conventional radiotherapy. *International Journal of Radiation Oncology Biology Physics.* 2010;78(4):998-1004.
27. Guan X, Wang X, Liu Y, et al. Lymph node metastasis in sinonasal squamous cell carcinoma treated with IMRT/3D-CRT. *Oral Oncology.* 2013;49(1):60-5.
28. Gupta T, Agarwal J, Jain S, et al. Three-dimensional conformal radiotherapy (3D-CRT) versus intensity modulated radiation therapy (IMRT) in squamous cell carcinoma of the head and neck: a randomized controlled trial. *Radiother Oncol.* 2012 Sep;104(3):343-8. PMID: 22853852.
29. Lambrecht M, Nevens D, Nuyts S. Intensity-modulated radiotherapy vs. parotid-sparing 3D conformal radiotherapy. Effect on outcome and toxicity in locally advanced head and neck cancer. *Strahlenther Onkol.* 2013 Mar;189(3):223-9. PMID: 23319256.
30. Ozyigit G, Cengiz M, Yazici G, et al. A retrospective comparison of robotic stereotactic body radiotherapy and three-dimensional conformal radiotherapy for the reirradiation of locally recurrent nasopharyngeal carcinoma. *International Journal of Radiation Oncology Biology Physics.* 2011;81(4):e263-e8.
31. Nutting CM, Morden JP, Harrington KJ, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. *Lancet Oncol.* 2011 Feb;12(2):127-36. PMID: 21236730.
32. Chao KS, Deasy JO, Markman J, et al. A prospective study of salivary function sparing in patients with head-and-neck cancers receiving intensity-modulated or three-dimensional radiation therapy: initial results. *Int J Radiat Oncol Biol Phys.* 2001 Mar 15;49(4):907-16. PMID: 11240231.
33. Marchal C, Lapeyre M, Beckendorf V, et al. [Preliminary results of the assessment of intensity modulated radiotherapy (IMRT) for prostatic and head and neck tumors (STIC 2001)]. *Cancer Radiother.* 2004 Nov;8 Suppl 1:S121-7. PMID: 15679257.
34. Chen AM, Daly ME, Bucci MK, et al. Carcinomas of the paranasal sinuses and nasal cavity treated with radiotherapy at a single institution over five decades: are we making improvement? *Int J Radiat Oncol Biol Phys.* 2007 Sep 1;69(1):141-7. PMID: 17459609.

35. Fang FM, Tsai WL, Chen HC, et al. Intensity-modulated or conformal radiotherapy improves the quality of life of patients with nasopharyngeal carcinoma: comparisons of four radiotherapy techniques. *Cancer*. 2007 Jan 15;109(2):313-21. PMID: 17154159.
36. Golen M SK, Wygoda A, et al. The influence of radiation technique on xerostomia in head and neck cancer patients: prospective study. *Rep Pract Oncol Radiother*. 2007;12(5):253-60.
37. Hodge CW, Bentzen SM, Wong G, et al. Are we influencing outcome in oropharynx cancer with intensity-modulated radiotherapy? An inter-era comparison. *Int J Radiat Oncol Biol Phys*. 2007 Nov 15;69(4):1032-41. PMID: 17967300.
38. Rades D, Fehlauer F, Wroblewski J, et al. Prognostic factors in head-and-neck cancer patients treated with surgery followed by intensity-modulated radiotherapy (IMRT), 3D-conformal radiotherapy, or conventional radiotherapy. *Oral Oncol*. 2007 Jul;43(6):535-43. PMID: 17005437.
39. Fang FM, Chien CY, Tsai WL, et al. Quality of life and survival outcome for patients with nasopharyngeal carcinoma receiving three-dimensional conformal radiotherapy vs. intensity-modulated radiotherapy-a longitudinal study. *Int J Radiat Oncol Biol Phys*. 2008 Oct 1;72(2):356-64. PMID: 18355980.
40. Gomez DR, Hoppe BS, Wolden SL, et al. Outcomes and prognostic variables in adenoid cystic carcinoma of the head and neck: a recent experience. *Int J Radiat Oncol Biol Phys*. 2008 Apr 1;70(5):1365-72. PMID: 18029108.
41. Palazzi M, Tomatis S, Orlandi E, et al. Effects of treatment intensification on acute local toxicity during radiotherapy for head and neck cancer: prospective observational study validating CTCAE, version 3.0, scoring system. *Int J Radiat Oncol Biol Phys*. 2008 Feb 1;70(2):330-7. PMID: 17881148.
42. Rusthoven KE, Raben D, Ballonoff A, et al. Effect of radiation techniques in treatment of oropharynx cancer. *Laryngoscope*. 2008 Apr;118(4):635-9. PMID: 18176348.
43. Vergeer MR, Doornaert PA, Rietveld DH, et al. Intensity-modulated radiotherapy reduces radiation-induced morbidity and improves health-related quality of life: results of a nonrandomized prospective study using a standardized follow-up program. *Int J Radiat Oncol Biol Phys*. 2009 May 1;74(1):1-8. PMID: 19111400.
44. Langendijk JA, Doornaert P, Rietveld DH, et al. A predictive model for swallowing dysfunction after curative radiotherapy in head and neck cancer. *Radiother Oncol*. 2009 Feb;90(2):189-95. PMID: 19167120.

## **Definition of Terms**

Terms are defined in the text, tables, and figures.

## Summary of Protocol Amendments

Date	Section	Original protocol	Revised protocol	Rationale
January 10, 2014	IV. Methods: Grading the Strength of Evidence (SOE) for Individual Comparisons and Outcomes	Please refer to section IV(F), p. 14: Grading the Strength of Evidence (SOE) for Individual Comparisons and Outcomes	Please refer to section IV(F), p. 14: Grading the Strength of Evidence (SOE) for Individual Comparisons and Outcomes	We performed a total rewrite based on input from the TOO and AHRQ personnel to make explicit the process to be used for grading the SOE, based on the updated chapter in the Methods Guide (2013).
January 10, 2014	IV. Methods: P. 10	"We will include only randomized controlled trials (RCTs) and non-randomized comparative studies (observational, case-control, and cohort studies) of populations, comparisons, interventions, and outcomes that are part of the PICOTS (see above). We will exclude non-comparative studies from this CER,"...	"We will include only full-length reports—excluding conference abstracts and other non-peer reviewed articles—describing final results of randomized controlled trials (RCTs) and non-randomized comparative studies (observational, case-control, and cohort studies) of populations, comparisons, interventions, and outcomes that are part of the PICOTS (see above). We will exclude non-comparative studies from this CER,"...	To make explicit study selection criteria that include only full-length, peer-reviewed evidence

### Review of Key Questions

For all EPC reviews, key questions were reviewed and refined as needed by the EPC with input from Key Informants and the TEP to assure that the questions were specific and explicit about what information is being reviewed. In addition, the key questions were posted for public comment and finalized by the EPC after review of the comments.

### Key Informants

Key Informants are the end-users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the key questions for research that will inform health care decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high-priority research gaps and needed new research. Key Informants were not involved in

analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals were invited to serve as Key Informants and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

### **Technical Experts**

Technical Experts comprise a multidisciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes as well as identifying particular studies or databases to search. They were selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, study questions, design, and/or methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not conduct analyses, contribute to the writing of the report, or review the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals were invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

### **Peer Reviewers**

Peer Reviewers were invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. Peer review comments on the preliminary draft of the report were considered by the EPC in preparation of the final draft of the report. Peer Reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments were documented and will, for CERs and Technical Briefs, be published three months after the publication of the report.

Potential reviewers must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than \$10,000. Peer Reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

### **EPC Team Disclosures**

EPC core team members must disclose any financial conflicts of interest greater than \$1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest, which cumulatively total greater than \$1,000, will usually disqualify EPC core team investigators.

### **Role of the Funder**

This project was funded under Contract No. 290-2007-10058 from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.